Motivated Beliefs and Biases

By

Zeeshan Samad

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

March 31, 2021

Nashville, Tennessee

Approved:

Myrna H. Wooders, Ph.D. (Chair)
Andrew F. Daughety, Ph.D.
Gregory C. Leo, Ph.D.
Bradley A. Malin, Ph.D.
Jennifer F. Reinganum, Ph.D.

ACKNOWLEDGEMENTS

A very special thank you to my committee, Professors Myrna Wooders, Greg Leo, Andrew Daughety, Jennifer Reinganum, and Brad Malin. This dissertation would not have been possible without their constant support, guidance, and wisdom. I am particularly grateful to Professor Wooders for the invaluable time she invested towards helping me become a better writer, researcher, and person; to Professor Leo for willing to meet me nearly every day as I developed my experiment designs; to Professors Daughety and Reinganum for staying on my committee and actively helping me even after formally retiring; and to Professor Malin for introducing me to the world of interdisciplinary research and encouraging me to develop habits of a systematic researcher. This dissertation benefited greatly from financial support from GetPrecise, an NIH Center of Excellence in Ethics Research, and from Professor Wooders' research funds that she graciously spent to fund my experiments.

Thank you to my parents, my wife Rabia, and my boys Ismael and Yahya, for their infinite support, encouragement, and love.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Appendices

CHAPTER 1

Self-Deception: Adopting False Beliefs for a Favorable Self-View

Abstract

This paper demonstrates how people deceive themselves into thinking of themselves as altruistic. I present a lab experiment in which subjects need to decide whether to behave altruistically or selfishly in an ambiguous environment. Due to the nature of ambiguity in this environment, those who are pessimistic have a legitimate reason to behave selfishly, even if they are inherently altruistic. For people who are inherently selfish but like to think of themselves as altruistic, this environment can serve as a scapegoat for selfish behavior. That is, by falsely claiming to be pessimistic, individuals can behave selfishly without damaging their self-image of being altruistic. Through two seemingly unrelated experimental tasks, I elicit subjects' adopted beliefs and true beliefs about the same probability. I find that selfish subjects adopt beliefs that are systematically more pessimistic beliefs than their true beliefs, whereas altruistic subjects adopt beliefs that are in alignment with their true beliefs. The most plausible explanation for why only selfish subjects manipulate their beliefs is that selfish behavior damages their self-image and belief manipulation helps them mitigate that damage; altruistic subjects, by contrast, have no such need for belief manipulation because their behavior does not damage their self-image.

1.1 Introduction

We like to think of ourselves as good people and, at the same time, like to act in ways that conflict with our definition of being good. As these conflicting desires coexist, we look for ways that can allow us to *think* of ourselves as good without *truly* being good.

This paper examines how people maintain the self-image of an altruistic person even as they deliberately choose a selfish action over an altruistic one.[1] To do this, I present a static model in which an individual faces a trade-off between the self-image of an altruistic person and a higher monetary payoff, both of which she desires. This model depicts a situation in which an individual needs to decide whether to behave altruistically or selfishly (e.g., while deciding whether or not to give a dollar to a panhandler), where behaving altruistically is a surefire way of obtaining an altruistic self-image but is also costly. The conflicting desires for behaving selfishly and for being an altruistic person create an unpleasant tension called *cognitive dissonance*.[2] According to traditional cognitive dissonance theory, an individual may resolve this dissonance by changing either her behavior or her self-concept (Festinger, 1957). In the current context, this would correspond to either behaving altruistically or

[1] Throughout the paper, I take "good" to mean altruistic. This is mostly for the sake of concreteness, and the same arguments can apply to any other "good" characteristic traits (e.g., honesty, intelligence, well-mannered, good-looking, etc.).

The idea that people derive utility from an altruistic self-image is well-established. For example, several experimental studies (Ahmed and Salas, 2011; Battigalli et al., 2013; Lambarraa and Riener, 2015; Mazar et al., 2008; Shariff and Norenzayan, 2007; Tonin and Vlassopoulos, 2013) show that people are more willing to give money when their self-image is at stake, compared to when their self-image is not at stake. As for *why* people want to think of themselves as altruistic, Benabou and Tirole (2002, p. 872) propose, "First, people may just derive utility from thinking well of themselves, and conversely find a poor self-image painful. Second, believing – rightly or wrongly – that one possesses certain qualities may make it easier to convince others of it. Finally, confidence in his abilities and efficacy can help the individual undertake more ambitious goals and persist in the face of adversity."

[2] The description of cognitive dissonance in this paper owes a great deal to the excellent works by Akerlof and Dickens (1982) and Rabin (1994). For example, Akerlof and Dickens (1982, p.308) write, "cognitive dissonance reactions stem from peoples' view of themselves as 'smart, nice people.' Information that conflicts with this image tends to be ignored, rejected, or accommodated by changes in other beliefs."

accepting the self-image of a selfish person. However, as this paper shows, there is another way of resolving cognitive dissonance, namely belief manipulation, which allows people to behave selfishly and also maintain the self-image of an altruistic person. This way involves coming up with a situational excuse that can justify selfish behavior. This justification, in turn, frees people to behave selfishly without holding their inherent altruism responsible for their behavior. For example, an individual who is contemplating whether to give money to a panhandler can manipulate her belief about the panhandler's likelihood of being a drug addict, because such a belief can justify not giving money in this particular instance while allowing her to believe that she would have given money under other circumstances.

First, I present a theoretical model in which an individual decides whether to behave altruistically or selfishly in an ambiguous environment, i.e., an environment where her action gets implemented with an *unknown* probability, say p. While the individual does not know the actual value of p, she has a prior belief about it, say $\tilde{p}$. If she believes this probability to be substantially large (specifically $\tilde{p} \geq \overline{p}$, for some self-created threshold $\overline{p}$), then she finds it optimal to behave altruistically. I show that such an individual can further increase her utility by manipulating her belief about p to something less than $\overline{p}$, as doing this allows her to maintain the self-image of an altruistic person (more precisely, the self-image of someone who behaves altruistically whenever $\tilde{p} \geq \overline{p}$) without incurring the cost of behaving altruistically. That is, the model predicts that in an ambiguous environment, an individual can maintain the self-image of an altruistic person by manipulating her beliefs instead of behaving altruistically.

Next, I present a lab experiment based on this theoretical model. In the experiment, subjects are endowed with 100 tokens and asked if they would like to donate half of them (i.e., 50 tokens) to a charity. If they choose to donate, the charity receives 120 tokens (i.e., 2.4 times the amount donated) with probability p and 0 tokens with probability

$1 - p$. Subjects do not know the actual value of p, but are told that p will be randomly selected from the following list of eleven numbers: 0%, 10%, 20%, ... , 100%. The most important feature of this experiment is that it elicits subjects' beliefs about p in two different ways. In the first way, subjects are directly asked about their belief about p, without being offered a monetary reward for accuracy or honesty. The un-incentivized nature of this question is crucial because it gives subjects an opportunity to deliberately falsify their beliefs. The second way of eliciting subjects' beliefs about p is indirect, and hidden inside a seemingly unrelated task. In this task, subjects choose whether they would like to receive an ambiguous lottery, specifically one that pays 120 tokens with probability p, or an unambiguous lottery that pays 120 tokens with a known probability (e.g., 50%). Each subject does this for several different known probabilities, in a price list format. The row in which they switch reveals their beliefs about p. Since the second belief elicitation method is incentive compatible while the first one is not, I refer to the beliefs elicited using the second way as *true* beliefs about p and beliefs elicited using the first way as *adopted* beliefs about p.

Comparing subjects' adopted and true beliefs shows that selfish subjects (i.e., subjects who chose not to donate) adopt systematically more pessimistic (lower) beliefs about p than their true beliefs, whereas altruistic subjects adopt their true beliefs. This is because pessimistic beliefs can serve as a valid excuse to behave selfishly, and only selfish subjects have an incentive to look for an excuse that can allow them to blame their selfish behavior on something other than their own selfishness. This result is also consistent with the model's prediction that subjects have an incentive to maintain the self-image of an altruistic person by manipulating their beliefs. Overall, I find that about 35% of subjects who maintain the self-image of an altruistic person maintain it by manipulating beliefs whereas the remaining 65% maintain it by donating 50 tokens.

The rest of the paper is structured as follows. Section 1.2 shows how this paper contributes to the literature. Section 1.3 presents a theoretical model in which selfish subjects obtain self-image utility deceptively (i.e., by manipulating their beliefs) whereas altruistic subjects obtain self-image utility while adopting their true beliefs. Sections 1.4 and 1.5 present, respectively, the design and results of the laboratory experiment. Section 1.6 discusses other possible reasons that might explain such a systematic discrepancy in adopted and true beliefs, and Section 1.7 concludes.

1.2 Related Literature

Since James Andreoni's seminal paper on the theory of warm glow (Andreoni, 1990), it has become increasingly evident that a genuine concern for others (i.e., pure altruism) is not the only, and perhaps not even the dominant, reason that explains why people behave altruistically. Instead, a more significant reason appears to be that people want to be *viewed* as altruistic, by others (Andreoni and Bernheim, 2009; Ariely et al., 2009; Bénabou and Tirole, 2006; Harbaugh, 1998) and by themselves (Benabou and Tirole, 2002). People who are motivated by self-image concerns can obtain utility from the self-image of an altruistic person either by behaving altruistically or by deceiving themselves.

This paper is most closely related to studies that depict various ways in which people make excuses for their selfish behavior. For example, Di Tella et al. (2015) and Andreoni and Sanchez (2019) conduct experiments in which they show that subjects who refuse to give money to others are the ones who claim that their potential recipient is likely to be a selfish person (and, therefore, not deserving of money). Most interestingly, they do not *truly* think so, indicating that the real reason for refusing to give money is that they are selfish, but they do not want to admit that. Another experiment by Exley (2016) shows that subjects who refuse to donate to a charity claim to be more risk-averse than they truly are, because in her experiment setting, risk aversion creates a disincentive to donate money, even for purely altruistic subjects.

Another group of papers shows that people take pains to avoid situations in which they may be asked to give money (Andreoni et al., 2017; Broberg et al., 2007; Dana et al., 2006, 2007; Lazear et al., 2012). If they are doing this because it allows them to continue thinking of themselves as altruistic without having to actually give, then this is a way of deceiving themselves for the sake of a favorable self-image, which is precisely the topic of the current paper. However, this paper presents another mechanism of self-deception, namely belief manipulation, which contrasts with the avoidance mechanism presented in this group of paper.

The theoretical model of my paper builds upon the frameworks developed by Akerlof and Dickens (1982) and Rabin (1994). Akerlof and Dickens (1982) model a situation where workers in a hazardous industry can choose to believe that their work is safe. There is a benefit of such a belief, but also a cost. The benefit, which is purely psychological, is that it prevents unpleasant feelings (e.g., being constantly afraid of a hazard, or doubting their decision of taking this job), and the cost is greater exposure to risk, because this belief motivates workers to stop using safety equipment, which increases the probability of an accident. They model this as a two-period problem in which safety equipment becomes available only in the second period and workers are forced to face the possibility of an accident in the first period. Because the fact that workers continue to work in this industry creates cognitive dissonance, workers choose to believe that the job is safe. In the second period, when safety equipment becomes available, workers do not purchase it because by then they believe their work to be safe. They show that workers will choose to believe their job is safe if the psychological cost of fear (which they can now avoid) is greater than the monetary cost of safety equipment plus the perceived cost of an accident. Rabin

(1994) models a situation in which people can choose whether to engage in a morally questionable activity (e.g., buying fur products and thus contributing to an animal-rights violation). People believe that there is a morally legitimate level of activity and anything beyond that level would create cognitive dissonance. Moreover, they can choose, at a cost, this moral threshold level of activity. Letting x be the level of activity a person chooses and letting y be a morally legitimate level of activity, a person's utility function is $U(x) - D(x - y) - C(y)$, where $U(x)$ represents material utility derived from the activity, $D(x - y)$ represents cognitive dissonance for engaging in an excess amount of activity, and $C(y)$ represents psychological costs of changing one's belief about y.

In the experiment presented in the current paper, subjects are given an opportunity to manipulate their beliefs about a probability. This aspect makes this paper related to some other studies that give subjects an opportunity to manipulate their perceptions about certain probabilities (Andreoni and Sanchez, 2019; Dana et al., 2007; Exley and Kessler, 2019; Exley, 2016; Haisley and Weber, 2010; Regner and Matthey, 2017).[3] Dana et al. (2007) conduct a series of modified dictator games that they call *moral wiggle room experiments*. Their main finding, which is also confirmed by later replications (e.g., Feiler, 2014; Larson and Capra, 2009), is that subjects are more likely to behave selfishly when they have an opportunity to blame their selfish behavior on something else (e.g., lack of information, other individuals, or uncertainty). Haisley and Weber (2010) show that subjects are ambiguity-averse for themselves but ambiguity-seeking for others. They find that dictators are more likely to behave selfishly when a recipient's payoff depends on an ambiguous lottery than on a lottery with a known probability. Moreover, when asked to estimate ambiguous payoffs of recipients, dictators suddenly adopt greater optimism,

[3]In all of these experiments, subjects interact anonymously with each other. However, their choices are not anonymous to the experimenter, which means that subjects may be partially motivated by a favorable *social* image in addition to self-image.

demonstrating an ability to manipulate their beliefs about their own ambiguity aversion. Exley (2016) provides another example of a dictator game in which the recipient is a charity. In this work, she investigates whether subjects' behavior is consistent with the hypothesis that individuals behave altruistically in order to obtain the self-image of an altruistic person and find excuses to not donate money. She shows that subjects who are not very risk-averse themselves, pretend to be extremely risk-averse when it can become an excuse to not donate money. Using multiple price lists under various scenarios, she shows that subjects who exhibit a low level of risk aversion when their own payoffs are not affected by a risky donation, suddenly start exhibiting high levels of risk aversion when their own payoffs *are* affected by a donation. In a related study by Exley and Kessler (2019), subjects maintain the self-image of an altruistic person by damaging another aspect of their self-image: their cognitive abilities. In particular, this study shows that subjects behave as if they suffer from cognitive limitations that prevent them from understanding the decision problem correctly. However, they do this only when such cognitive limitations can serve as an excuse for behaving selfishly, and do not exhibit the same cognitive limitations when their payoff is not affected. Regner and Matthey (2017) also give experiment subjects an opportunity to manipulate beliefs about a probability, but using trust games. They show that reciprocating subjects take up excuses to reduce their transfer amounts. Specifically, when there is a probability that their transfer could fail, they return less money.[4]

At the same time, altruism is not the only desirable trait for which people have been shown to deceive themselves. For instance, an important paper by Quattrone and Tversky (1984) shows that subjects deceptively obtain the self-image of a healthy person. In their

[4]It should be noted that the precise motivation for making this excuse is unclear. This is because actions in a trust game can be motivated by various traits, such as altruism, fairness, reciprocity, trustworthiness, or trust in others. It is not clear if subjects are motivated to think of themselves as, say, more trustworthy or more altruistic.

experiment, subjects are first asked to keep their hand submerged in cold water until they could no longer tolerate the pain. This is followed by a debriefing, in which subjects are told that a certain inborn heart condition can be diagnosed by the effect of exercise on tolerance to cold. About half of the subjects are told that having a bad heart would increase cold tolerance, while the other half are told the opposite. (They even back this up with charts showing lifespans of people who have this heart condition and who do not.) Having absorbed this information, subjects are put on an Exercycle for one minute, after which they repeat the same cold water tolerance test. Interestingly, both groups of subjects show changes in tolerance in the direction that is indicative of a healthy heart. In effect, they cheat on their own diagnosis. In another study, Gneezy et al. (2016) show that people deceptively obtain the self-image of an honest person. They ask their subjects to recommend either option A or option B to another experiment participant, where one of the two options benefits them personally. They observe that subjects are more likely to recommended the option that benefits them personally. However, in a post-experiment survey, the same subjects claim to have made recommendations sincerely, i.e., with only the other person's interest at heart. They point out an important policy implication of this finding, that physicians should not be given financial benefits from recommending certain procedures. Mijović-Prelec and Prelec (2010) show that people like to think of themselves as someone who is good at making predictions and are willing to obtain that self-image deceptively. They show their experimental subjects a set of symbols (specifically Korean characters) and, for each symbol, they ask their subjects if the symbol is more masculine or feminine (there is a correct answer). However, before each symbol is shown, subjects are asked to anticipate whether the next symbol would be masculine or feminine. They find that subjects tend to conform to their predictions – in a treatment where subjects are not asked to anticipate the next symbol, the same subjects make different classifications.

Mijović-Prelec and Prelec propose that the reason subjects conform to their predictions is that it lets them believe that they made the correct guess.

Lastly, the current paper relates to a group of papers that shows that there is a strong correlation between the ability to persuade oneself and the ability to persuade others. Since deception is a special case of persuasion, this means that people who seek to acquire the skill of effective persuasion have an incentive to deceive themselves. For example, Smith et al. (2017) show that, when incentivized to persuade others about a neutral notion (e.g., "Mark is a nice guy"), subjects end up believing that notion themselves in the process of persuading others. Similarly, Schwardmann and Van der Weele (2019) show that an attempt to persuade others about their own abilities makes subjects more confident in themselves. Because people get utility from holding positive beliefs about themselves, they propose that this belief manipulation may be subconsciously intentional, calling it *strategic self-deception*. Bentley et al. (2019) show that people are more effective at persuading others about viewpoints that they sincerely believe themselves; and manipulate their beliefs when they have an incentive to become effective persuaders.

1.3 Model

1.3.1 Base Environment

An individual decides how much money to donate to a charity. For every x dollars she donates, the charity receives $y = rx$ dollars, for some $r \geq 1$. The individual is endowed with some money that is normalized to 2 dollars, and has an altruism level that is measured by a random variable $a \in [0, 1]$, so that the most selfish person is the one with $a = 0$ and higher values of a represent more altruistic individuals.[5] Moreover, a is a continuous

[5]The assumption $a \in [0, 1]$ suggests that people value others' consumption up to as much as they value their own consumption, not more. While this makes sense, this is not a technical requirement for the model

random variable with a probability distribution function $f(a)$ and cumulative distribution function $F(a) = \int_0^a f(x)dx$ s.t. $F(0) = 0$ and $F(1) = 1$.

An individual's utility function comprises of two additively separable components, *altruism utility,* $u_a(rx, a)$, and *consumption utility,* $u_c(2-x)$, where the former increases with x and the latter decreases with x. Both u_a and u_c are increasing and concave. This is formalized by the definition below.

Definition 1. An individual's utility function is $U(rx, a) = u_a(rx, a) + u_c(2 - x)$, where $u_a(y, a)$ satisfies properties 1-4, and $u_c(y)$ satisfies properties 5-6 below.

1. $u_a(0, a) = 0 \quad \forall a$, i.e., a donation of 0 results in no altruism utility, regardless of an individual's altruism level.

2. $u_a(y, 0) = 0 \quad \forall y$, i.e., the most selfish individual is defined as someone who obtains zero utility from donating money.

3. $u_a'(y) \geq 0$ and $u_a''(y) \leq 0$, i.e., altruism utility increases with the donation amount, but at a decreasing rate.

4. For any $y > 0$, $\frac{\partial u_a(y,a)}{\partial a} \geq 0$ and $\frac{\partial^2 u_a(y,a)}{\partial a^2} \leq 0$, i.e., more altruistic individuals get more utility than less altruistic individuals from donating the same amount.[6]

5. $u_c'(x) > 0$ and $u_c''(x) < 0$, i.e., consumption utility is strictly increasing and strictly concave.

6. $u_c(0) = 0$, i.e., an individual who keeps zero dollars gets zero consumption utility.

since all propositions can also hold for $a \in [0, \infty)$.

[6] This is what will motivate more altruistic people to donate more money. In the absence of this assumption, there would be no behavioral difference between altruistic and selfish individuals.

Lemma 1. There exists a cost function $c(x) = u_c(2) - u_c(2-x)$ s.t. an individual with an altruism level of a will choose x^* that solves $u_a'(rx^*, a) = c'(x^*)$. Moreover, $c(x)$ is strictly increasing and convex, with $c(0) = 0$.

Proof. The FOC of $U(rx, a) = u_a(rx, a) + u_c(2 - x)$ is

$$\begin{aligned} u_a'(rx, a) - u_c'(2 - x) &= 0 \\ u_a'(rx, a) &= c'(x) \end{aligned}$$

Since $u_c(x)$ is strictly increasing and concave, $u_c(2-x)$ is strictly decreasing and concave, and $-u_c(2 - x)$ is strictly increasing and convex.

Proposition 1. Let $x_0 = \arg\max_x U(x, a_0)$ and $x_1 = \arg\max_x U(x, a_1)$. Then $x_0 < x_1$ if and only if $a_0 < a_1$.

Proof. This proof will first show (i) If $x_0 < x_1$ then $a_0 < a_1$ and then show (ii) If $a_0 < a_1$ then $x_0 < x_1$. For simpler notation, let $u_a(x_i, a_0) =: u_0(x_i)$ and $u_a(x_i, a_1) =: u_1(x_i)$.

(i). Let $x_0 < x_1$ but suppose $a_0 \geq a_1$. Given $\frac{\partial u}{\partial a} \geq 0$, for any x,

$$u_0(x) \geq u_1(x)$$

Since $u_0(0) = u_1(0) = 0$, the above inequality can be written as:

$$\begin{aligned} \frac{u_0(x) - u_0(0)}{x - 0} &\geq \frac{u_1(x) - u_1(0)}{x - 0} \\ \Rightarrow u_0'(x) &\geq u_1'(x) \end{aligned} \tag{1.1}$$

We also know (from Proposition 1.3.1) that $c'(x_0) = u_0'(x_0)$ and $c'(x_1) = u_1'(x_1)$.

Moreover, because $c''(x) < 0$, we have

$$
\begin{aligned}
& c'(x_0) < c'(x_1) \\
& \Rightarrow u_0'(x_0) < u_1'(x_1) \\
& \Rightarrow u_0'(x_1) \leq u_0'(x_0) < u_1'(x_1) \quad \text{(since } u_0''(x) \leq 0 \text{ and } x_0 < x_1) \\
& \Rightarrow u_0'(x_1) < u_1'(x_1)
\end{aligned}
$$

which contradicts 1.1. Therefore, if $x_0 < x_1$ then $a_0 < a_1$.

(ii). Conversely, let $a_0 < a_1$, and suppose that $x_0 \geq x_1$. Given $c''(x) > 0$, we have

$$c'(x_0) > c'(x_1) \tag{1.2}$$

Since $\frac{\partial u}{\partial a} \geq 0$, then for all x,

$$u_0(x) \leq u_1(x)$$

Given $u_0(0) = u_1(0) = 0$, this can be rewritten as:

$$
\begin{aligned}
& \frac{u_0(x) - u_0(0)}{x - 0} \leq \frac{u_1(x) - u_1(0)}{x - 0} \\
& \Rightarrow u_0'(x) \leq u_1'(x)
\end{aligned}
$$

In particular, this will also hold for $x = x_0$, i.e.,

$$u_0'(x_0) \leq u_1'(x_0)$$

Given that $u_1''(x) < 0$, $x_0 \geq x_1$ implies that $u_1'(x_0) \leq u_1'(x_1)$, and thus:

$$u_0'(x_0) \leq u_1'(x_0) \leq u_1'(x_1)$$
$$\Rightarrow u_0'(x_0) \leq u_1'(x_1)$$

Given the optimization conditions $u_0'(x_0) = c'(x_0)$ and $u_1'(x_1) = c'(x_1)$, the above inequality implies:

$$c'(x_0) \leq c'(x_1)$$

which contradicts 1.2. Therefore, if $a_0 < a_1$ then $x_0 < x_1$.

Example 1. Suppose that, for some $b \in [0, 1]$, an individual has the following altruism and consumption utility functions.[7]

$$u_a(rx, a) = a(rx)^b$$
$$u_c(x) = x^b$$

The first order condition is:

[7]For example, if $b = 0.5$ and $r = 1$, the net utility function is $U(x, a) = \sqrt{x} + a\sqrt{x}$. Note that having the same parameter value, b, in both functions suggests that individuals think that others are similar to them in the way they derive utility from money, which seems like a reasonable assumption and has been adopted by several notable models of altruism (e.g., Bénabou and Tirole, 2006; Levine, 1998). The difference between less altruistic and more altruistic individuals is that more altruistic people assign a greater relative weight to others people's utility.

$$u'_a(rx) = c'(x)$$

$$abr^b x^{b-1} = b(2-x)^{b-1}$$

$$\frac{r^b a}{x^{1-b}} = \frac{1}{(2-x)^{1-b}}$$

$$\left(\frac{2-x}{x}\right)^{1-b} = \frac{1}{r^b a}$$

$$\frac{2-x}{x} = \frac{1}{r^{\frac{b}{1-b}} a^{\frac{1}{1-b}}}$$

$$x^*(a) = \frac{2r^{\frac{b}{1-b}} a^{\frac{1}{1-b}}}{1 + r^{\frac{b}{1-b}} a^{\frac{1}{1-b}}}$$

To illustrate this for some simple parameter values, suppose $b = 0.5$ and $r = 1$. Then $x^*(a) = \frac{2a^2}{1+a^2}$, which means that an individual with $a = 1$ (i.e., someone who values a charity's payoff as much as her own) will donate $x^* = 1$ dollar (i.e., half her endowment). Similarly, someone with $a = 0.5$ will donate $x^* = 0.4$ (i.e., one-fifth of her endowment). Figure 1.1 illustrates this example by presenting indifference curves and budget constraints for altruism levels $a = 1$ and $a = 0.5$.

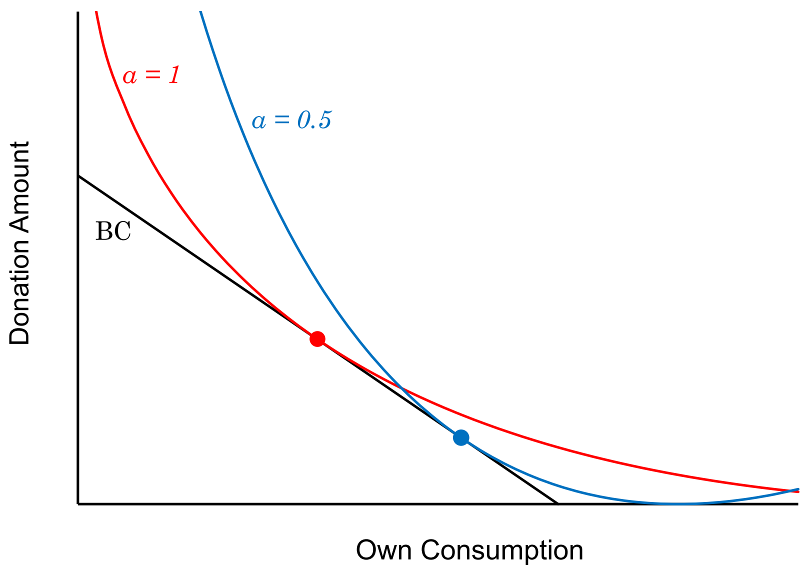
In the example above, the relationship $x^*(a) = \frac{2r^{\frac{b}{1-b}} a^{\frac{1}{1-b}}}{1+r^{\frac{b}{1-b}} a^{\frac{1}{1-b}}}$ shows that x^* is strictly increasing for all $a \in (0, 1)$.[8] However, it is also possible that $x^* = 0$ for a range of values of a. And, as shown by the corollary below, this range is continuous, i.e., $(\underline{a}, \overline{a}) \subseteq [0, 1]$.

Corollary. An individual's donation choice $x(a)$ is a function of a s.t. $x'(a) \geq 0$. Moreover, $\exists\ 0 \leq \underline{a} < \overline{a} \leq 1$ s.t. $x'(a) > 0 \quad \forall a \in (\underline{a}, \overline{a})$.

Proof. Suppose that $\bar{x}$ is the largest observed donation amount (i.e., $x^*(a = 1) = \bar{x}$). We already know that the smallest donation is zero (by assumption). Define $\underline{a}$ as the largest

[8]An easy way to see this is that the numerator increases twice as fast as the denominator for any $r \geq 1$ and $0 \leq b \leq 1$. Alternatively, one can verify that $\frac{\partial x}{\partial a} > 0$.

Figure 1.1: Example of Indifference Curves



value of a for which $u'(0, a) = c'(0)$, and define $\bar{a}$ as the smallest value of a for which $u'(\bar{x}, a) = c'(\bar{x})$. This corollary states that individuals with very low levels of altruism (i.e., $a \in [0, \underline{a}]$) will donate zero dollars in this environment and individuals with very high levels of altruism (i.e., $a \in [\bar{a}, 1]$) will donate $\bar{x}$ dollars, which is also the amount that the most altruistic person donates.

The fact that $f(a)$ is continuous guarantees that there are at least two values $\underline{a} \in [0, 1]$ and $\bar{a} \in [0, 1]$ such that $\underline{a} < \bar{a}$. The rest follows from Proposition 1.3.1, i.e., $\underline{a} < \bar{a} \implies x(\underline{a}) < x(\bar{a})$, $0 \leq \underline{a} \implies x(0) \leq x(\underline{a})$, and $\bar{a} \leq 1 \implies x(\bar{a}) \leq x(1)$.

1.3.2 Uncertainty Environment

So far we have considered a *base environment* where an individual chooses how many dollars to donate. Now consider an *uncertainty environment* in which an individual faces a binary decision, denoted by $x \in \{0, 1\}$ of whether (or not) to donate some fixed amount. For simplicity, let us fix that amount to half of her endowment. Letting endowment remain normalized to 2 dollars, the individual needs to decide whether to donate 1 dollar.[9]

If she donates (i.e., $x = 1$), the charity receives either $\bar{y}$ dollars, with some p, or 0 dollars, with probability $1 - p$. Let us also assume that $\bar{y}$ is large enough so that the expected value of this outcome is greater than the amount donated, i.e., $\bar{y}p > 1$ (to prevent donating in this environment from being trivially unattractive). If she does not donate, the charity receives nothing. Let $x = 1$ represent an individual's decision to donate and let $x = 0$ represent her decision to not donate. Using the same altruism and consumption utility functions as earlier, an individual's *net expected utility* is

$$\mathbb{E}U(x|a, p, \bar{y}) = xpu_a(\bar{y}, a) + u_c(2 - x) \tag{1.3}$$

where values of $p, \bar{y}$, and a are known, and the functions $u_a(\cdot)$ and $u_c(\cdot)$ have the properties mentioned in Definition 1.3.1.

Lemma 2. An individual's optimal choice is $x^* = 1$ if and only if $p \geq \frac{C}{u_a(\bar{y},a)}$, where $C = u_c(2) - u_c(1)$ represents dis-utility associated with losing 1 dollar.

Proof. For any $a, p, \bar{y}$, first suppose $p \geq \frac{C}{u_a(\bar{y},a)}$. Then:

$$p \cdot u_a(\bar{y}, a) + u_c(1) \geq u_c(2)$$
$$\implies \mathbb{E}U(x = 1) \geq \mathbb{E}U(x = 0)$$
$$\implies x^* = \arg\max_x \mathbb{E}U(x, p) = 1$$

Conversely, suppose $x^* = \arg\max_x \mathbb{E}U(x, p) = 1$. Then:

[9]Thus, the notation $x \in \{0, 1\}$ also allows us to interpret x as the actual donation amount.

$$
\begin{aligned}
&\mathbb{E}U(1) \geq \mathbb{E}U(x') \quad \forall x' \neq 1 \\
\implies &\mathbb{E}U(1) \geq \mathbb{E}U(0) \\
\implies &p \cdot u_a(\bar{y}, a) + u_c(1) \geq u_c(2) \\
\implies &p \geq \frac{C}{u_a(\bar{y}, a)}
\end{aligned}
$$

It may help to think of a world where $\bar{y}$ and C are fixed, so that environments are different only due to p and individuals are different only due to a. This would also help in seeing that extreme cases are trivial. For example, the optimal decision for someone who is extremely selfish ($a \to 0$) is, trivially, to not donate. Similarly, in an environment where the probability of transfer is very small ($p \to 0$), the optimal decision for any type of individual is to not donate. Therefore, the most interesting applications of this model are for regular, as opposed to extreme, environments.

The following two lemmas present some results for regular environments, while mentioning what a regular environment technically means. Lemma 1.3.2 shows that in an environment where the most altruistic person (i.e., someone with $a = 1$) finds it optimal to donate, there must be someone who is indifferent between donating and not donating, so that those who are more altruistic than the indifferent type will donate while those who are less altruistic will not.[10] Lemma 1.3.2 shows that every individual has some threshold probability, so that they find it optimal to donate in environments with a greater probability and to not donate in environments with a lower probability. Together, these lemmas direct the reader to think about regular situations, i.e., where some people prefer

[10] Throughout the framework, I assume that when indifferent between donating and not donating, an individual donates.

to donate and others do not, and about regular people, i.e., who would donate in very favorable conditions but not donate in unfavorable conditions.

Lemma 3. For any p' and a' s.t. $p' > \frac{C}{u_a(\bar{y},a')}$, $\exists a^* < a'$ s.t. $x^*(p',\bar{y},a) = 0 \quad \forall a < a^*$ and $x^*(p',\bar{y},a) = 1 \quad \forall a \geq a^*$.

Proof. Suppose $p'u_a(\bar{y},a') > C$, i.e., for given p' and $\bar{y}$, an individual of altruism level a' strictly prefers to donate.[11] We also know, by assumption, that $u_a(y,0) = 0 \forall y$, i.e., an individual with $a = 0$ prefers not to donate, and that $C > 0$. Since $\frac{\partial u_a}{\partial a} > 0$, there must be some value $0 < a^* < a'$ s.t. $0 = u_a(\bar{y},0) < p'u_a(\bar{y},a^*) = C < p'u_a(\bar{y},a')$. That is, an individual with an altruism level a^* is indifferent between donating and not donating, individuals with $a \geq a^*$ prefer to donate, and individuals with $a < a^*$ prefer not to donate.

Lemma 4. For any individual who strictly prefers to donate when $p = 1$, $\exists p^* \in (0,1)$ s.t. $x^* = 0 \quad \forall p < p^*$ and $x^* = 1 \quad \forall p \geq p^*$.

Proof. Let a' be the altruism level of an individual who strictly prefers to donate when $p = 1$. From Lemma 1.3.2, this implies $u_a(\bar{y},a') - C > 0$. Let $g(p) = pu_a(\bar{y},a') - C$, observing that $g(p)$ is strictly increasing and one-to-one in p, and, therefore has an inverse for all $p \in \mathbb{R}$. Since $g(0) = -C < 0 < g(1)$ and $g^{-1}(0)$ exists, define $p^* = g^{-1}(0) = \frac{C}{u_a(\bar{y},a')}$. That is, $g(p^*) = 0$. Therefore, an individual of altruism level a' will be indifferent between donating and not donating at $p = p^*$, will strictly prefer to donate when $p > p^*$, and will strictly prefer to not donate when $p < p^*$.

[11] Note that this proposition is stated in general terms, i.e., for all $u_a(\cdot)$ that have the required properties, not just $u_a = au_c$

Definition 2. An individual's *probability threshold*, interpreted as the lowest probability of transfer at which she is willing to make a donation, is $p^* = \frac{C}{u_a(\bar{y},a)}$. Given that a is exogenous, and p^* is a one-to-one transformation of a, p^* is also exogenous.

Proposition 2. In the uncertainty environment, an individual's optimal strategy, conditional on p, is: $x^* = \begin{cases} 0 & \text{if } p < p^* \\ 1 & \text{if } p \geq p^* \end{cases}$, where $p^* = \frac{C}{u_a(\bar{y},a)}$.

Proof. Suppose $p \geq p^*$. Then $p \geq \frac{C}{u_a(\bar{y},a)} \implies pu_a(\bar{y},a) > C \implies \mathbb{E}U(1|a,p,\bar{y}) \geq \mathbb{E}U(0|a,p,\bar{y}) \implies \arg\max_x \mathbb{E}U(x|a,p,\bar{y}) = 1$. Therefore, $p \geq p^* \implies x^* = 1$

Conversely, suppose $p < p^*$. Then $p < \frac{C}{u_a(\bar{y},a)} \implies pu_a(\bar{y},a) < c(1) \implies \mathbb{E}U(1|a,p,\bar{y}) < \mathbb{E}U(0|a,p,\bar{y}) \implies \arg\max_x \mathbb{E}U(x|a,p,\bar{y}) = 0$. Therefore, $p < p^* \implies x^* = 0$

Example 2. Suppose an individual is endowed with 2 dollars and has the opportunity to donate $x \in \{0, 1\}$ dollars from it to a charity. The charity receives, with an equal likelihood, either $2.4x$ dollars or 0 dollars. Suppose individuals have utility functions of the form $u_c(y) = \sqrt{y}$ and $u_a(y) = a\sqrt{y}$. What is the probability threshold of an individual with an altruism level of a?

From Definition 1.3.2, $p^* = \frac{C}{u_a(\bar{y},a)} = \frac{\sqrt{2}-\sqrt{1}}{a\sqrt{2.4}} \approx \frac{0.267}{a}$. So, more altruistic individuals will have lower probability thresholds, with the most altruistic probability threshold being $p^* = 0.267$. By Proposition 1.3.2, individuals with a probability threshold of $p^* \leq 0.5$ (or equivalently with altruism level $a \geq .534$) will make a donation.

1.3.3 Ambiguity Environment

Now consider a last donation environment, namely *ambiguity environment*, which is different from the uncertainty environment in only one way: the probability p is no longer a constant value, but a random variable that has a continuous distribution $f_P(p)$. In this environment, an individual forms a belief about p and makes a donation decision based on that belief. Denoting $\tilde{p}$ as an individual's belief about p, an individual's choice consists of two parameters, $x \in \{0, 1\}$ and $\tilde{p} \in [0, 1]$, or in simpler notation, $(x, \tilde{p})$.

Self-image. The altruism utility function, $u_a(y, a)$, is intended to contain utility from both pure as well as impure altruism. In particular, it contains *self-image utility*, $s(p^*)$, which is the utility that an individual gets from believing that she has a probability threshold of p^*.[12] To capture the notion that individuals get greater image utility from thinking of themselves as more altruistic, i.e., having a lower p^*, define $s(p^*)$ as a decreasing function of p^*, i.e., $s'(p^*) < 0$.

Further, to incorporate the idea that self-image concerns are important motivators for altruistic behavior, assume that $s(p^*)$ is a large part of u_a. Or, for simplicity, it may help to imagine that altruism utility is almost entirely made up of self-image utility, i.e., $u_a \approx s(p^*)$. That is, an individual donates when $p \geq p^*$ not because she cares about donating, but because she cares about believing that she is the kind of person who donates whenever $p \geq p^*$.

True beliefs. Individuals know $f_P(p)$ and also have a given way (e.g., by taking the mean of the distribution) of using this distribution to form an initial, or *true*, belief about p. Thus, define an individual's *true belief* as $p_t = \int_0^1 p f_P(p) dp$.[13] Note that since p_t is calculated using a given function, p_t is essentially exogenous. That is, an individual does not get to choose her true belief; instead, nature chooses it for her.

[12] Recall that this is equivalent to saying that an thinks of herself as having an level of a, since $p^*(a)$ is a one-to-one and decreasing function of a.

[13] It does not matter what the actual function for converting f_P to p_t is, as long as it is consistent and given. For example, instead of defining an individual's true belief as the mean of the distribution, we could have defined it as the median or mode (or even the 75th percentile) of the distribution, without having to change anything that follows.

If an individual is unable to adopt any other belief (e.g., if she cannot come up with any excuse that might justify a self-serving action), then she he has no choice but to adopt her true belief, i.e., $\tilde{p} = p_t$. And given that p_t is exogenous, an individual's decision problem in this case is identical to the one in the uncertainty environment. Therefore, her optimal strategy would be

$$(x, \tilde{p}) = \begin{cases} (1, p_t) & \text{if } p_t \geq p^* \\ (0, p_t) & \text{if } p_t < p^* \end{cases}$$

If she chooses $x = 0$ when $p_t \geq p^*$, then she loses $s(p^*)$ because she can no longer claim that she is the kind of person who makes a donation whenever $p_t \geq p^*$. However, and most importantly, she does *not* lose $s(p^*)$ from choosing $x = 0$ when $p_t < p^*$. This is because she can still claim that she would have donated if $p_t \geq p^*$. Her belief about having a probability threshold of p^* is only challenged when $p_t \geq p^*$ and not when $p_t < p^*$.

More generally, an individual's net utility in this environment is

$$U(x, \tilde{p}) = \begin{cases} u_c(2 - x) + s(p^*) & \text{if } \tilde{p} < p^* \\ u_c(2 - x) & \text{if } \tilde{p} \geq p^* \end{cases}$$

Separating this for each $x \in \{0, 1\}$ gives us:

$$U(0, \tilde{p}) = \begin{cases} u_c(2) + s(p^*) & \text{if } \tilde{p} < p^* \\ u_c(2) & \text{if } \tilde{p} \geq p^* \end{cases}$$

$$U(1, \tilde{p}) = \begin{cases} u_c(1) + s(p^*) & \text{if } \tilde{p} < p^* \\ u_c(1) + s(p^*) & \text{if } \tilde{p} \geq p^* \end{cases}$$

Manipulated beliefs. Imagine that an individual, who thinks of herself as someone who would always donate if asked, is asked to donate (say, by a panhandler). Taking some help from observable clues (e.g., "this panhandler looks rather smartened up") and some from imagination (e.g., "panhandling is probably a very lucrative business"), the individual forms a belief about the probability that her donation will actually go to someone deserving. If she is able to convince herself that this probability is rather low (and specifically lower than her own threshold p^*), then she can choose not to donate and still maintain the self-image of an altruistic person, obtaining utility of $s(p^*)$.

In this situation, an individual spontaneously comes up with an excuse to justify her selfish action. Given the limited time, and therefore cognitive effort, that the individual has at her disposal, I assume that there is a small, fixed cost of manipulating beliefs, say $\varepsilon > 0$.[14] Instead, the more binding constraint for the individual is the number of excuses that she can come up with in this time frame. To represent this constraint, suppose that $\tilde{p}$ must be chosen from a finite set $\mathcal{P}$ that contains n elements, each of which represents a belief about p that the individual can possibly adopt. Since an individual's true belief is always available to her, the set $\mathcal{P}$ always contains at least one element, p_t. If an individual is able to manipulate her beliefs, then $n > 1$. The number of manipulated beliefs available to an individual, $n - 1$, can substantially differ across individuals and across situations.

Definition 3. Denote p_m as the most pessimistic belief among all manipulated beliefs that an individual can come up with. That is, $p_m = \min\{\mathcal{P} \backslash \{p_t\}\}$.

Recall that all previous results still hold. In particular, Proposition 2 implies that for any $x \in \{0, 1\}$, $\tilde{p} \neq p_t \iff U(x, \tilde{p}) > U(x, p_t)$. In terms of the self-image function, this means $s(p^*) > C$.

[14] This is a departure from prior models of belief manipulation (Akerlof and Dickens, 1982; Rabin, 1994) where the cost of manipulating one's belief increases with the size of manipulation.

Proposition 3. Let p_m be the lowest possible manipulated belief. Then an individual's optimal strategy is

$$(x, \tilde{p}) = \begin{cases} (0, p_t) & \text{if } p_m < p_t < p^* \\ (0, p_m) & \text{if } p_m \leq p^* < p_t \\ (1, p_t) & \text{if } p^* < p_m < p_t \end{cases}$$

Proof. (i) Suppose $p_t < p^*$. Since the individual believes that she is the type of person who makes a donation whenever $p \geq p^*$, she is not liable to make a donation in order to maintain a self-image of $s(p^*)$. Thus, she will choose $(0, p_t)$ since that has the lowest cost.

(ii) Suppose $p_m \leq p^* < p_t$. In this case, if an individual holds on to her true belief, then she will find it optimal to donate, obtaining a net utility of $s(p^*) + u_c(1)$. If she manipulates her belief to p_m, or any $\tilde{p} < p^*$, then she can obtain a greater net utility, equal to $s(p^*) + u_c(2) - \varepsilon$. Therefore, in this case, the optimal action is $(0, \tilde{p}) \forall \tilde{p} \in [p_m, p^*)$.

(iii) Suppose $p^* < p_m < p_t$. In this case, an individual finds it optimal to donate regardless of whether she manipulates her belief top_m or holds on to her true belief. However, by holding on to her true belief, she at least saves on the effort cost of manipulation, obtaining a net utility of $s(p^*) + u_c(1)$. Therefore, her optimal choice in this case is $(1, p_t)$.

Example 3. Suppose $f_P(p) \sim U[0, 1]$ so that $p_t = 0.5$. Also suppose an individual can manipulate her belief by a maximum of 0.1, so that$p_m = 0.4$. Lastly, she is endowed a probability threshold of $p^* = 0.5$.

The individual has four possible choices, i.e., $(x, \tilde{p}) \in \{(0, p_t), (1, p_t), (0, p_m), (1, p_m)\}$.

The utility from each of these choices is:

$$U(0, p_t) = u_c(2)$$

$$U(1, p_t) = u_c(1) + s(p^*)$$

$$U(0, p_m) = u_c(2) + s(p^*) - \varepsilon$$

$$U(1, p_m) = u_c(1) + s(p^*) - \varepsilon$$

Thus, the optimal solution is $(x, \tilde{p})^* = (0, p_m)$.

Conclusion. This framework can be used to separate selfish individuals into those who do not desire a self-image of an altruistic person, those who desire and obtain the self-image of an altruistic person, and those who do desire a self-image of an altruistic person but do not obtain it.[15] In terms of p^*, there are three types of individuals: (i) $p_m \leq p_t < p^*$, i.e., those who are content with being selfish; (ii) $p^* < p_m \leq p_t$, i.e., those who do not manipulate their beliefs, either because they are motivated by pure altruism or they are not able to manipulate their beliefs; (iii) $p_m \leq p^* < p_t$, i.e., those who manipulate their beliefs solely to maintain the self-image of an altruistic person.

1.4 Experiment Design

I design a laboratory experiment that most closely resembles the parameters from Examples 1.3.2 and 1.3.3. That is, subjects are asked whether to donate 50 tokens from an endowment of 100 tokens. If they make this donation, a charity receives 120 tokens with a probability p, or 0 tokens with probability $1 - p$. Subjects do not know the actual value of p but know that it is uniformly distributed between 0 and 1. Through two separate tasks,

[15] To be more precise, an "altruistic person" in the ambiguity environment is someone who is altruistic enough to donate 50 tokens in an environment where the charity receives 120 tokens with probability $\tilde{p}$.

namely *donation task* and *raffle task*, I elicit subjects' manipulated beliefs, p_m, and true beliefs, p_t.

1.4.1 Implementation

I conduct this experiment in the form of an online, computerized survey. I recruit subjects through Amazon's micro-employment platform called Mechanical Turk. In this platform, any adult (i.e., over 18) who has a social security number can register as a *worker* and perform small tasks for modest compensation. My survey is available to workers located in the United States, on a first-come-first-serve basis. Amazon acts as an intermediary between workers and *requesters* (i.e., experimenters). This assures subjects that they will receive any payments promised to them, and also ensures their anonymity. Requesters are not provided any identifiable information about workers, and can only identify workers by an alphanumeric ID, of 14-20 characters long, assigned to them by Amazon. Amazon takes several measures to prevent fraudulent activity (e.g., one person cannot create multiple worker accounts; workers need to continuously exhibit traits of human intelligence, etc.), and takes a sizable (40%) commission for providing this service.[16]

To ensure that subjects will be able to comprehend the experiment instructions, they must correctly answer some qualifying questions in order to participate in the experiment. Those who pass the qualifying questions are guaranteed a participation payment of $4. In addition to this, they earn money during the experiment in terms of *tokens*, which are converted to US dollars at the end of the experiment at a conversion rate of $0.02/token or $0.12/token, chosen randomly. Each subject earns at least 50 tokens, and at most 220 tokens, from the experiment tasks. This means that the minimum possible earning is $5.00,

[16]For more details on these measures, see https://blog.mturk.com/important-updates-on-mturk-marketplace-integrity-worker-identity-and-requester-tools-to-manage-206e4e90da0c.

and maximum possible earning is \$30.40.[17] Given that the survey takes about 20 minutes, these are considerably generous amounts. Subjects make all choices through their own computers, and cannot go back to a previous page at any point in the survey. The actual survey questions, including screenshots of the interface seen by subjects, are shown in Appendix A.1.

1.4.2 Donation Task

Immediately after passing the qualifying questions, subjects are taken to the first task, titled *donation task*, and are given 100 tokens (equivalent to \$2 or \$12, depending on the exchange rate chosen for the subject). Subjects are first asked to choose a charity from a list of 11 popular charities, each of which supports a different cause.[18] After choosing a charity, they need to decide whether or not they want to donate 50 tokens (i.e., \$1 or \$6, depending on the randomly selected exchange rate) to that charity.[19] Being limited to two choices forces subjects into a situation where they must either take an action that is clearly altruistic or an action that is clearly selfish. In particular, it would be very difficult

[17]These are subjects' net earnings, i.e., excluding Amazon's commission.

[18]I let subjects choose a charity so that they are not able to use the pre-specification of charity as an excuse for not donating. That is, a subject might prefer to donate towards, say, cancer research instead of poverty relief. Such a preference would provide a subject with an excuse to not donate, allowing her to think of herself as generous without making a donation. Therefore, allowing subjects to choose their own charity reduces the number of possible excuses available to them.

[19]The greatest advantage of a binary choice is that it allows me to easily separate altruistic and selfish subjects. At the same time, to control for subjects who might turn a binary choice set itself into an excuse – e.g., by thinking that they would have donated, say, 30 tokens if they could have and now must resort to donating 0 – I include a question in the post-experiment questionnaire that asks them if they would have donated another amount if they were not limited to a binary choice. I find that some subjects do claim that they would have made a donation if they could have donated, say, 20 tokens; however, all of these subjects also manipulate their beliefs about the probability p, suggesting that self-deceivers find multiple ways to deceive themselves while non-deceivers do not come up with any way to deceive themselves.

An interesting extension of this work would be to give subjects more choices (i.e., letting them donate any number of tokens) and then observing if having a range of choices gives subjects more avenues of self-deception, such as manipulating the definition of altruism. For example, some individuals may be able to maintain a self-image of an altruistic person by donating a very small portion of their endowments (and subsequently thinking, "I am obviously altruistic because otherwise I would have donated nothing.").

to view a “selfish” action as a “slightly less altruistic” action.

Subjects who donate 0 tokens get to keep their entire endowment of 100 tokens, and subjects who donate 50 tokens get to keep the remaining 50 tokens of their endowment. If a subject donates, then with a probability p, the charity receives 120 tokens (which is 2.4 times the donation amount), and with the remaining probability $1-p$, the charity receives nothing.[20] The value of p is unknown, but subjects are told that it will be chosen randomly, and with equal likelihood, from these numbers: $\{0\%, 10\%, 20\%, 30\%, 40\%, 50\%, 60\%, 70\%, 80\%, 90\%, 100\%\}$. Along with choosing a donation amount, subjects may also provide an email address where they would receive a donation receipt directly from the charity.[21] I categorize subjects who donate as *altruistic*, and subjects who do not donate as *selfish*.

On the same screen (i.e., before submitting their choice of donating), subjects are asked to guess the value that p will take. Let us refer to this response as subjects' *adopted belief*s. A subject can either adopt her true belief or some other, false, belief. To avoid nudging subjects towards adopting their true beliefs, I do not reward (or penalize) them on the accuracy of their guess. Keeping this question deliberately un-incentivized question is a crucial aspect of this experiment, particularly because it allows subjects to adopt a

[20] A couple of weeks prior to running this experiment, I conducted a survey on Amazon Mechanical Turk in which I asked subjects, in a price-list format, to choose between a sure outcome of 50 tokens (for themselves) and a lottery that pays x tokens with a probability of 0.5, for $x \in \{60, 70, 80, 90, 100, 110, 120, 130, 140, 150\}$. The median response of that study was 120 tokens, suggesting that an average subject would be indifferent between this donation environment and a simpler environment where a donation of 50 tokens results in the charity receiving 50 tokens with certainty. While this opens the doors for another possible extension of comparing these results with another treatment in which a charity receives 50 tokens for sure (and, therefore, belief manipulation is not possible), for the current purposes it does not matter what number of tokens the charity is paid with a probability of p (i.e., it does not matter if 50 tokens gets converted to a lottery that pays 120 tokens with probability p or 220 tokens with probability p as long as all subjects are given the same numbers).

[21] The ability to obtain proof of donation prevents subjects from using their skepticism as an excuse to rationalize selfish behavior, e.g., by saying “Who knows if the charity will even get the money. How do I know the experimenter will not keep 120 tokens with probability p for himself?”.

pessimistic belief about p, even if they are not truly pessimistic. By adopting a pessimistic belief, a subject is able to behave selfishly and still maintain the self-image of an altruistic person (or, more precisely, the image of a person who is altruistic under a favorable environment but not under an unfavorable environment). Therefore, selfish subjects – and particularly those selfish subjects who get utility from thinking of themselves as altruistic – have an incentive to adopt pessimistic beliefs, whereas altruistic subjects do not.

If the non-incentivized nature of this question causes subjects to not think carefully, or respond carelessly (perhaps even randomly), then the response will only be noisy (due to misstatements in both directions), but not affect the *average adopted belief*. However, if selfish subjects systematically adopt more pessimistic beliefs than altruistic subjects, then one is hard-pressed to think of any other explanation, except that subjects do not want to donate money but also want to think of themselves as altruistic.

1.4.3 Raffle Task

After submitting their choices in the donation task, subjects are taken to a *raffle task*, in which they are given the same lottery that would be given to the charity, i.e., 120 tokens with probability p. This lottery is in addition to the participation fee and earnings from the donation task. Subjects are told that the same value of p that would determine the charity's payoff in the donation task will be used to determine their payoff in the raffle task. They still do not know the actual value of p, but are reminded that it is uniformly distributed between 0 and 1. Moreover, since subjects are not informed about the raffle task until they actually begin the task itself, their previous choices – particularly their adopted belief about p – cannot possibly be influenced by anything included in this task.

In the raffle task, subjects are asked how they would like to receive these additional earnings. In particular, they have a choice between keeping this ambiguous lottery or replacing it with an unambiguous (but still uncertain) lottery, i.e., where the probability of receiving 120 tokens is *known*. Subjects make this choice in a *multiple price list* format,[22] i.e., they choose between this ambiguous lottery and an outside lottery for 11 different outside lotteries, making a total of 11 decisions, one of which is is randomly chosen for determining their payoff.[23]

Each of 11 decision rows consists of two options, where the first option is always a lottery that pays 120 tokens with probability p and 0 tokens with probability $1 - p$, i.e., the ambiguous lottery. The second option is a lottery that pays 120 tokens with a *known* probability and 0 tokens with one minus that probability. The only difference between one row and another row is the value of this probability. In the first decision row, the known probability is 100%, which means that the second option pays 120 tokens with certainty (see Figure A.9 in the Appendix). With each descending row, the known probability in the second option falls by 10%. In the last decision row, the known probability is 0%, i.e., the second option pays 0 tokens with certainty. Clearly, the outside lottery is better in the first decision row and the ambiguous lottery is better in the last decision row.[24] In all other rows, subjects should choose the first option only if they believe that p is greater than the known probability given in that row. Therefore, the row in which a subject switches from the second option to the first option reveals the subject's true belief about the value of p.

Belief Elicitation Method. The real purpose of the raffle task is to elicit subjects' *true beliefs* about p, so these can be later compared with subjects' adopted beliefs in

[22]See Andersen et al. (2006) for further explanation and discussion of the multiple price list method.

[23]This is a fairly well-established method of eliciting true beliefs about a probability (Andreoni and Sanchez, 2019; Schlag et al., 2015; Schotter and Trevino, 2014).

[24]Previous experiments involving price lists report that some subjects switch multiple times between the two options presented (Holt and Laury, 2002; Jacobson and Petrie, 2009; Meier and Sprenger, 2010). Since multiple switch points can indicate subject confusion and are difficult to rationalize, it is generally accepted to use a framing device to avoid confusion and clarify the decision process (Andreoni and Sprenger, 2011; Exley, 2016). Following this, I explain in the instructions that the choices in the first and last rows involve certain outcomes and also pre-highlight the clearly better choices in those rows.

the donation task. In order to elicit true beliefs, subjects need to have an incentive to respond truthfully, that is, the belief elicitation method should be a proper scoring rule. However, not all proper scoring rules (including the quadratic scoring rule) would work with self-deception because subjects may have an incentive to remain consistent with their previous response, and that incentive may be stronger than the monetary incentive to be truthful. Therefore, I elicits subjects' true beliefs about p *indirectly* through another task that appears to have a completely different purpose. As Andreoni and Sanchez (2019, p. 6) explain, "When subjects have private incentives to mislead us or themselves on their true beliefs, the QSR or any other devise that asks directly for beliefs can be expected to elicit biased reports from subjects, even if it is a proper scoring rule. We instead must derive true beliefs by masking them in another task which, without the subject's awareness, will indirectly reveal beliefs." [25] This is exactly what the raffle task does. Based on questionnaire responses and informal discussions, this strategy appears to have been successful.[26] This is still safely distant from deception of subjects because subjects *can* figure out the true intention of this task by thinking a bit deeper – but do not. Instead, they think that the only purpose of this task is to determine which lottery should be used to pay subjects some additional money.

1.4.4 Questionnaire

Lastly, subjects complete a short questionnaire that collects some basic demographic information, such as gender, age, race, religiousness, education, and income. In addition,

[25] They also mention, "this method is superior to the QSR since it is valid beyond the case of risk neutrality." (Andreoni and Sanchez, 2019, p.7)

[26] I conducted a pilot version of this experiment with students of the undergraduate course *ECON 3893: Experimental Economics* taught by Prof. Myrna Wooders at Vanderbilt University. A post-experiment discussion revealed that none of the students were able to figure out the true intention of the raffle task. Given that these were students experimental economics, a more general audience is even more unlikely to make the connection between the raffle task and their true beliefs about p.

this questionnaire gives subjects the opportunity to provide an explanation for their action. For example, the first question asks, "What was your primary motivation to donate 50 [or 0] tokens?" Although I do not include these qualitative variables in the formal analysis, I use them to check if subjects are trying to come up with other types of excuses to justify not donating.

1.4.5 Hypotheses

After eliciting each subject's adopted belief and true belief, I compare the two. If there is a disparity between them, then there are two possible explanations for it. First, it could be that a subject deliberately adopts a false belief in the donation task, or, in other words, manipulates her belief. If this is the case, then someone who does not donate would have an incentive to adopt a more pessimistic belief about p than her true belief, whereas someone who donates would have an incentive to adopt a more optimistic belief than her true belief. The second possibility is that a subject responds randomly (or carelessly) when asked to adopt a belief, in which case her response would not indicate her adopted belief at all. Instead, it would be a meaningless random variable that, by virtue of being random, could be either more optimistic or more pessimistic than her true belief.

If it turns out that selfish subjects systematically adopt pessimistic beliefs while altruistic subjects systematically adopt optimistic beliefs, then it points towards belief manipulation. However, if it turns out that selfish subjects adopt both pessimistic and optimistic beliefs (and so do altruistic subjects), then it is difficult to claim deliberate belief manipulation. Recall that subjects themselves are the only observers of their actions and, therefore, any deliberate belief manipulation must be motivated by a desire for a self-image of an altruistic person, as opposed to a social image of an altruistic person.

Hypothesis 1.

Among subjects who do not adopt their true beliefs, only selfish subjects (and selfish subjects only) will adopt pessimistic beliefs, whereas only altruistic subjects will adopt optimistic beliefs. Using the same notation as in the model, this hypothesis states that if $\tilde{p} \neq p_t$, then $\tilde{p} < p_t \iff x = 0$ and $\tilde{p} > p_t \iff x = 1$, which is broadly consistent with Proposition 1.3.3

Hypothesis 2. A significant proportion of subjects have the willingness and ability to manipulate their beliefs. In terms of the notation used in the model, this means that the inequality $p_m \leq p^* < p_t$ holds for a non-negligible range of values of p^*. More precisely, $\exists \varepsilon > 0$ s.t. $F_P(p_t) - F_P(p_m) > \varepsilon$.

1.5 Results

In this section, I first describe my data set and then provide an overview of the results. In the following subsections, I explain each individual result in more detail along with showing each result's robustness to various methodologies.

1.5.1 Data

I recruited 110 workers from Amazon Mechanical Turk, with the only restriction that they must be located in the US. Out of these 110 participants, only 70 correctly answered the qualifying questions and, therefore, were able to proceed to the actual tasks. I dropped another 8 out of these 70 subjects because they switched multiple times in the raffle task.[27] Thus, the final data set contained 62 subjects. In the donation task, 27 of these 62 (44%) donated money while 35 (56%) did not donate money. Based on these donation choices, I assign subjects a *type*, which is either altruistic or selfish, ending up with 27 altruistic

[27] However, the results would have remained the same even if I had kept these 8 observations. I still dropped them, though, because it feels like the more prudent thing to do.

subjects and 35 selfish subjects.

Table 1.1 presents the summary statistics of altruistic and selfish subjects. The first column presents the average values of each variable for selfish subjects, and the second column presents the same for altruistic subjects. To see if any particular characteristics are correlated with selfish or altruistic behavior, I conduct a two-tailed t-test for each variable; the resulting p-values of these t-tests are reported in the third column. The t-test column shows that the variables $\tilde{p}$, *income*, and *age* are significantly different, on average, for selfish subjects and altruistic subjects. That is, these variables appear to be correlated with altruistic behavior, with older people being more likely to be more altruistic, and wealthier people also being more likely to be more altruistic.

Lastly, my overall sample of 62 subjects consists of only 19 (31%) females for some reason. This is notable only because it is different from 50%, which is the proportion of females in most random samples. Nonetheless, it is not concerning because the proportion of females is not significantly different across selfish and altruistic subjects.

1.5.2 Overview of Results

The first variable in Table 1.1, $\tilde{p}$, represents subjects' adopted beliefs and is obtained from their responses in the donation task. The second variable, p_t, represents subjects' true beliefs and is obtained from their responses in the raffle task. On average, selfish subjects truly believe the probability to be 48.29% but adopt the belief that it is 42%. That is, they exaggerate their pessimism, or manipulate their beliefs, by 6.29 percentage points on average. By contrast, altruistic subjects truly believe the probability to be 52.22%, on average, but adopt the belief that it is 55.56%, on average. These subjects manipulate their beliefs too, but by a much smaller amount of 3.34 percentage points and also by exaggerating their *optimism* as opposed to their pessimism. Regardless, this observation

Table 1.1: Summary Statistics

Variable Name	Selfish Subjects	Altruistic Subjects	*p*-value of Difference
$\widetilde{p}$	42.00	55.56	0.000
	(2.082)	(3.082)	
p_t	48.29	52.22	0.256
	(2.070)	(2.841)	
Income	3.171	3.889	0.043
	(0.241)	(0.241)	
Education	14.51	14.89	0.517
	(0.361)	(0.457)	
Religiousness (0-3 scale)	0.543	0.852	0.209
	(0.132)	(0.218)	
Age Category (1-6)	3.600	4.185	0.042
	(0.189)	(0.207)	
Female	0.229	0.407	0.134
	(0.0720)	(0.0964)	
Self-image concern (0-5)	1.914	3.815	0.000
	(0.185)	(0.214)	
Pure Altruism (0-5)	2.571	4.370	0.000
	(0.189)	(0.161)	
Selfishness (0-5)	4.571	3.185	0.000
	(0.138)	(0.245)	

This table reports the average (mean) values of each variable by subject type (namely *altruistic* and *selfish*), with standard errors of the mean in parentheses. The variable $\widetilde{p}$ represents adopted beliefs; p_t represents true beliefs; *income* is a categorical variable with 1: $<$\$20,000, 2: \$20,000-\$35,000; 3: \$35,000-50,000; 4:\$50,000-\$75,000; 5: \$75,000-\$100,000; 6: \$100,000-\$150,000, 7:$\geq$\$150,000; *education* represents a subject's years of education and has been converted from a categorical variable by mapping an individual's highest education level to years of education, so that someone who did not complete high school is assigned a value of 8, a high school graduate is assigned a value of 12, someone with a college degree is assigned 16, and someone with a post-graduate degree is assigned 18; *religiousness* is a self-reported measure of how religious a person is with larger values representing more religiousness; *age* is a categorical variable with 1: <18, 2: 18-25, 3: 25-35, 4: 35-45, 5: 45-55, 6: 55+; *female* is a binary variable for gender with 1: female and 0: male. The last three variables are self-reported and qualitative measures of a subject's motivation to behave altruistically.

suggests that all subjects manipulate their beliefs.

The third column of Table 1.1 (i.e., t-test results) shows that selfish subjects' adopted beliefs are significantly lower (with $p < 0.001$), or pessimistic, than those of altruistic subjects. However, the same is not true for their true beliefs. That is, the true beliefs of selfish subjects and of altruistic subjects are not significantly different ($p = 0.256$). This observation suggests that selfish subjects adopt pessimistic beliefs while altruistic subjects adopt optimistic beliefs, even though both types of subjects truly hold similar beliefs.

These observations are consistent with the hypothesis that individuals like to view themselves as altruistic (whether they truly are altruistic is a separate matter). This is because selfish subjects adopt pessimistic beliefs when pessimistic beliefs can allow them to blame their selfish action on an unfavorable donation environment. By contrast, altruistic subjects do not manipulate beliefs – or if they do, they manipulate them by a much small amount – because they have already proven their altruism to themselves (i.e., obtained the self-image of an altruistic person) by making a donation. The rest of this section shows that it turns out that more concrete analytical methods yield the same conclusions as these cursory observations.

1.5.3 Altruistic Behavior

To look for variables that are correlated with altruistic behavior, I conduct regressions using OLS, fixed effects, and probit specifications, results of which are shown in Table 1.2. The first column of the table presents the results obtained from using an OLS specification in which the dependent variable is the binary variable representing subjects' decision of whether to make a donation. The second column contains results of a regression that includes fixed effects for occupation category and charity chosen. The third column presents results obtained using a probit specification.

Table 1.2: Regression Results (y = *altruistic behavior*)

Variables	OLS	Fixed Effects	Probit
True belief (p_t)	0.00900*	0.00465	0.0271*
	(0.00493)	(0.00580)	(0.0148)
Female	0.247*	0.435**	0.753*
	(0.140)	(0.179)	(0.415)
Religiousness	-0.00315	-0.00260	-0.00724
	(0.0688)	(0.0827)	(0.201)
Years of Education	0.00425	0.0289	0.0265
	(0.0281)	(0.0346)	(0.0861)
Income	0.0840*	0.120**	0.260*
	(0.0467)	(0.0570)	(0.144)
Age	0.116*	0.0227	0.326*
	(0.0597)	(0.0684)	(0.174)
Constant	-0.890*	-0.859	-4.333**
	(0.520)	(0.609)	(1.724)
N	62	62	62

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

In all specifications, the dependent variable is the binary variable indicating the decision of whether to make a donation, and independent variables are: (i) true beliefs about p; (ii) a binary variable indicating if a subject is female; (iii) how religious a subject is, on a scale from 0 to 3; (iv) years of education; (v) income category; and (vi) age group.

In the OLS specification, the variables *true belief*, *female*, *income*, and *age* are statistically significant at the 10% level. Given that all of these coefficients are positive, these results imply that subjects who are (truly) more optimistic are more likely to donate than those who are less optimistic, women are more likely to donate than men are, wealthier people are more likely to donate than those who are less wealthy, and older people are more likely to donate than those who are younger. The second column of the table includes fixed effects within an OLS specification. I include occupation fixed effects to control for the fact that subjects belonging to certain occupation categories (e.g., employed) may be more likely to donate money than subjects belonging to other categories (e.g., unemployed).[28] I include charity fixed effects because some charities (e.g., WWF) appear to be more popular than others (see Figure A.12 in the Appendix for more details), which creates a reason to believe that those who choose a particular charity might be more likely to donate. In this specification, only the variables *female* and *income* are statistically significant. In the probit specification, the variables *true belief*, *female*, *income*, and *age* are statistically significant at the 10% level, similar to the first column.

1.5.4 Belief Manipulation

I define *belief manipulation* as the difference between a subject's adopted belief about p and her true belief about p.[29]

Correlation Between Selfish Behavior and Belief Manipulation. To see if selfish subjects are more likely to be manipulative even when controlling for other subject-level characteristics, I conduct an OLS regression with belief manipulation as the dependent variable. Table 1.3 shows the results of this regression. The first column of the table presents a simple OLS model with the following independent variables: (i) true beliefs about p; (ii) a binary variable indicating subjects' decision of whether or not to donate; (iii) an indicator variable for female; (iv) religiousness (on a scale from 0 to 3); (v) years of education; (vi) income group; and (vii) age group. The second column includes charity

[28] Occupation categories are (i) employed, (ii) unemployed, (iii) student, and (iv) other.

[29] This means that subjects who adopt more pessimistic beliefs have a positive value for belief manipulation.

fixed effects and occupation fixed effects in addition to these independent variables.

As expected, selfish subjects tend to manipulate their beliefs by adopting greater pessimism. This is shown by the negative and statistically significant coefficient of the variable *altruistic behavior* in Table 1.3. Selfish subjects adopt a belief that is significantly more pessimistic than their true belief. Specifically, selfish subjects exaggerate their pessimism by an amount that is 13.15 percentage points greater than the same exaggeration by altruistic subjects (note that altruistic subjects exaggerate their optimism, which means that in terms of pessimism, their exaggeration is actually negative).

In addition to *altruistic behavior*, the variables *true belief*, *female*, and *religiousness* are significant as well. The positive coefficient of *true belief* suggests that subjects who truly hold a more optimistic belief manipulate their beliefs more, most plausibly because they have greater room for adopting pessimism. The negative coefficient of *religiousness* implies that less religious people manipulate their beliefs more. This supports the notion that manipulation is deliberate and has something to do with a person's inherent moral values.

Statistical Tests. Another way to look at the size of belief manipulation is through the bar charts in Figure 1.2. Selfish subjects have a statistically significant difference between their adopted and true beliefs (two-tailed t-test $t = -3.7695, p = 0.0004$, Mann–Whitney $z = -3.756, p = 0.0002$, Pearson $\chi^2 = 14.0144$), whereas altruistic subjects do not. This suggests that selfish subjects manipulate their beliefs while altruistic subjects do not, consistent with our self-image hypothesis (that altruistic subjects do not need to manipulate beliefs because they do not need to). The beliefs truly held by selfish subjects are not significantly different from those held by altruistic subjects, but the beliefs *adopted* by selfish subjects are significantly more pessimistic than those *adopted* by altruistic subjects.

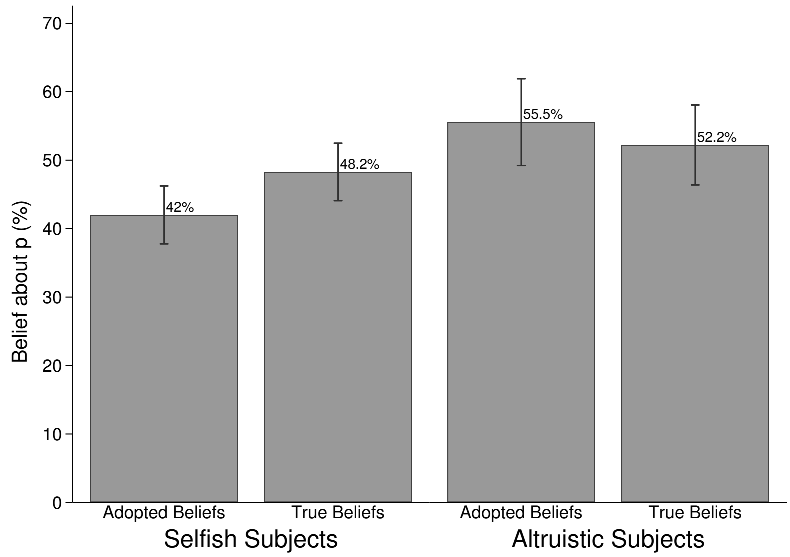
Direction of Manipulation.

Table 1.3: Regression Results ($y=$ *size of manipulation*)

Variables	OLS	Fixed Effects
True belief (p_t)	0.924***	1.041***
	(0.138)	(0.161)
Altruistic behavior	-13.60***	-13.15***
	(3.670)	(4.202)
Female	10.45***	10.44*
	(3.910)	(5.260)
Religiousness (0-3)	-4.657**	-4.698**
	(1.872)	(2.279)
Years of Education	-1.567**	-0.289
	(0.766)	(0.960)
Income Category (1-7)	0.152	-0.834
	(1.308)	(1.650)
Age Category (1-6)	0.663	0.351
	(1.679)	(1.888)
Constant	-18.32	-38.45**
	(14.54)	(17.17)
N	62	62

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

Figure 1.2: Average Values of Adopted Beliefs and True Beliefs

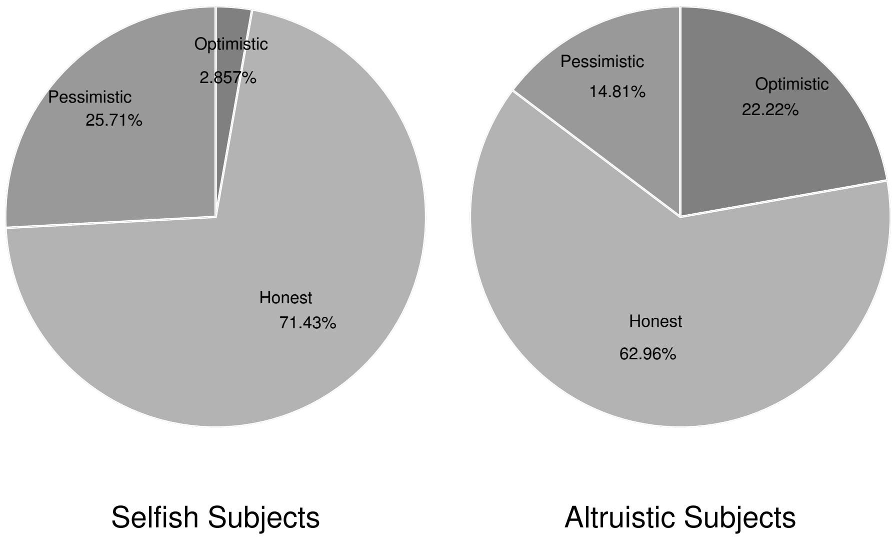


Error bars represent the 95% confidence interval

I code subjects as either *manipulative* or *honest*, depending on whether or not they adopt their true beliefs. More specifically, I code subjects as *manipulative* if their adopted beliefs about p are more than 10 percentage points away from their true beliefs (in either direction). I use this margin of 10 percentage points because subjects can only indicate their beliefs in increments of 10, and this criteria prevents incorrectly coding honest subjects as manipulative. (For example, a subject whose adopted and true beliefs are elicited as 40% and 50% respectively, may have made a rounding error in one of the two elicitation tasks and not deliberately manipulated her belief.)

Therefore subjects are now categorized into types in two different dimensions, i.e., in terms of altruism and in terms of manipulation. Combining both of these categories results in four subject types, (i) AH = altruistic and honest, (ii) AM = altruistic and manipulative, (iii) SH = selfish and honest, and (iv) SM = selfish and manipulative. Based on subjects' actions, the distribution of each of these types is as follows: 17 subjects (27%) are AH, 10 (16%) are AM, 25 (40%) are SH, and 10 (16%) are SM.

Figure 1.3: Proportion of Manipulative Subjects



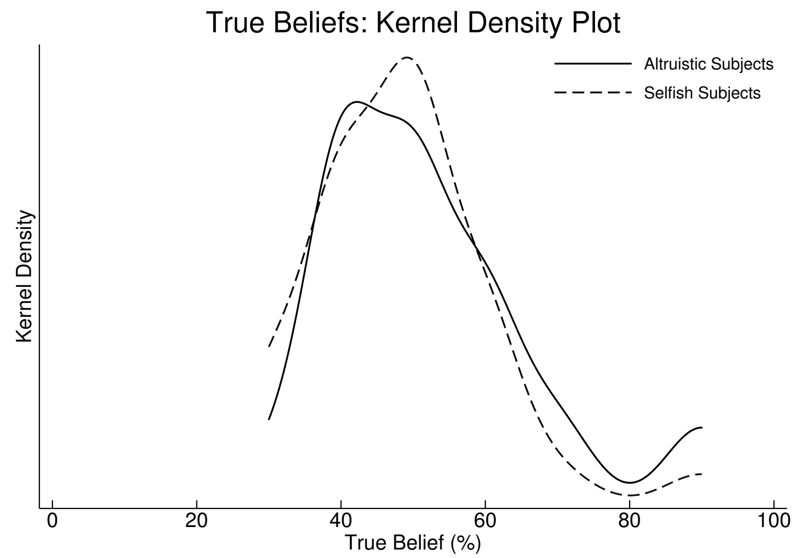
A further split of manipulative subjects into those who adopt more optimistic beliefs and those who adopt more pessimistic beliefs supports the previous result that only selfish subjects deliberately manipulate their beliefs. That is, each altruistic subject and each selfish subject can be either honest, optimistic, or pessimistic. Figure 1.3 shows that selfish and manipulative subjects systematically adopt more pessimistic beliefs while altruistic and manipulative subjects are almost equally likely to adopt either pessimistic or optimistic beliefs. Specifically, 9 out of 10 selfish and manipulative subjects adopt more pessimistic beliefs (i.e., only one adopts a more optimistic belief). By contrast, 6 out of 10 altruistic and manipulative subjects adopt optimistic beliefs while the remaining 4 adopt pessimistic beliefs, which suggests that altruistic subjects are not deliberately manipulative; instead, a discrepancy between their adopted beliefs and true beliefs appears to be due to carelessness.

Kernel Density Plots. Figure 1.4 graphically presents how true beliefs and adopted beliefs differ across selfish and altruistic subjects. The top panel shows kernel density plots of true beliefs held by selfish and altruistic subjects. The distribution of true beliefs seems to be quite similar, with kernel densities very closely aligned. The bottom panel shows similar plots for adopted beliefs. These plots show that adopted beliefs are considerably different across selfish and altruistic subjects, with a clear leftward shift for selfish subjects. This again indicates that selfish subjects systematically adopt more pessimistic beliefs about p, while altruistic subjects adopt beliefs that are both pessimistic and optimistic relative to their true beliefs.

Exogeneity of true beliefs. The analysis above implicitly assumes that true beliefs are exogenous, because only then belief manipulation can be done solely through adopted beliefs. To test for this, I conduct a regression to check for a correlation between adopted and true beliefs. Table 1.4 shows the results of this regression. The coefficient of $\tilde{p}$ is not statistically significant, which indicates a lack of a strong relationship between adopted beliefs and true beliefs and supports the notion that true beliefs are exogenous. This also supports an important feature of the experiment design, that subjects are not able to recognize that the raffle task is related to the adopted belief task.

Distribution of True and Adopted Beliefs. Figure 1.5 presents histograms of adopted and true beliefs, this time separately for manipulative subjects and honest subjects (as opposed to selfish subjects and honest subjects, which is how it is shown in Figures 1.2, 1.3, and 1.4). Recall that honest subjects are defined as those whose true beliefs and adopted beliefs are at most 10 percentage points apart, and manipulative subjects are those whose adopted beliefs are more than 10 percentage points different from their true beliefs. The top panel of Figure 1.5 shows the distribution of adopted beliefs, and the bottom panel shows the distribution of true beliefs. For honest subjects, adopted beliefs (and also true beliefs, since by definition, they are about the same) range from 30% to 70%, with the vast majority at 50%. For manipulative subjects, adopted beliefs vary from 10% to

Figure 1.4: Kernel Density Plots



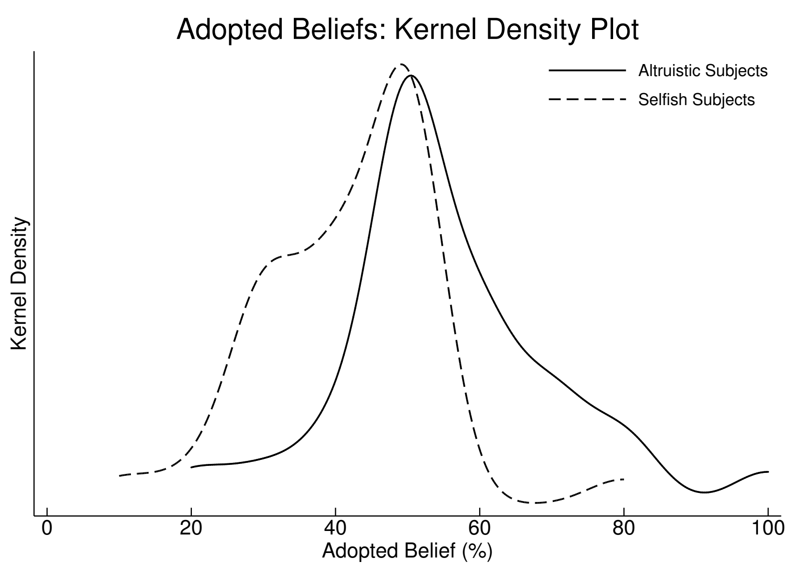


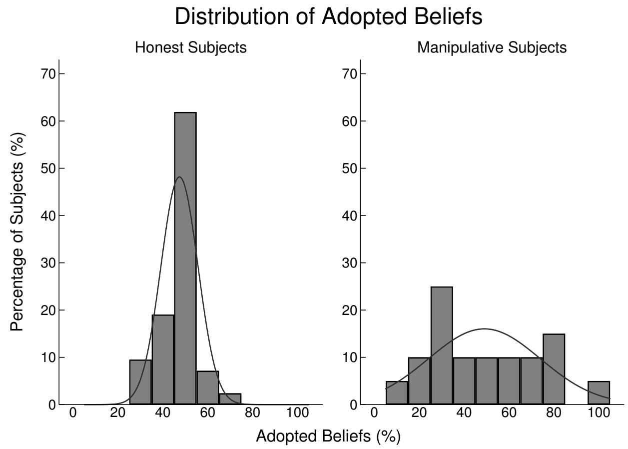
Table 1.4: Regression Results (y = *adopted beliefs*)

Variables	**OLS**	**Fixed effects**
True belief (p_t)	0.198	0.0202
	(0.149)	(0.175)
Female	-7.098*	-4.716
	(4.221)	(5.412)
Religiousness	4.614**	4.664*
	(2.077)	(2.501)
Education	1.625*	0.669
	(0.850)	(1.045)
Income	0.990	2.412
	(1.411)	(1.725)
Age	0.916	-0.0526
	(1.803)	(2.069)
Constant	6.215	27.16
	(15.72)	(18.43)
N	62	62

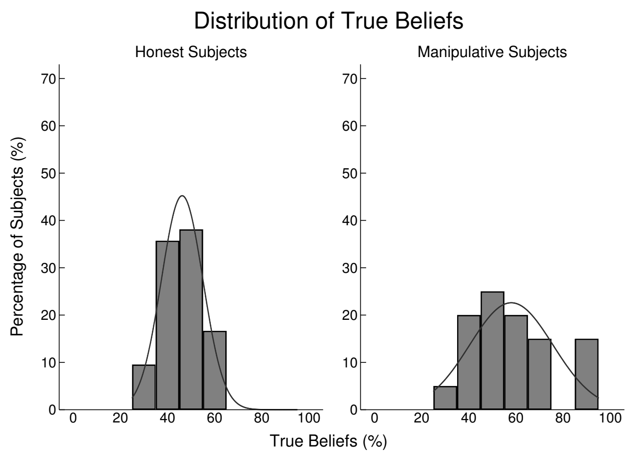
Standard errors in parentheses; * p-value< 0.10; ** p-value < 0.05; *** p-value < 0.01

Figure 1.5: Histograms

(a) Adopted beliefs



(b) True beliefs



100%, and true beliefs range from 30% to 90%. In the top panel, the two distributions are starkly different, whereas, in the bottom panel, the two are considerably less different. This suggests that manipulative subjects' disparity between adopted beliefs and true beliefs is being driven by adopted beliefs, not true beliefs, because these are most different for manipulative and honest subjects.

1.6 Discussion

Cognitive dissonance arises when an individual wants to participate in an activity that is contrary to her beliefs/values/morals. The literature on cognitive dissonance suggests that to reduce cognitive dissonance, one must bring her actions and morals in alignment, which can be done by either changing one's actions or changing one's morals. This paper introduces another mechanism of removing cognitive dissonance: an individual engages in the activity, and convinces herself, by making an excuse, that the activity is not contrary to her morals in this particular instance.

Proportion of manipulative subjects. I find that 32% of subjects manipulate their beliefs while the remaining 68% adopt their true beliefs, consistent with the literature on preferences for honesty (Abeler et al., 2019), . There are several possible explanations for this. First, these subjects may be unable to deceive themselves. For example, as they think about the probability p, they simply think about the question rather than about how their action affects their self-image. As another example, they may be unable to adopt a pessimistic view if perhaps a recent thought reminded them of how pessimistic they truly are. Second, maybe they do not even have a desire for the self-image of an altruistic person. Third, perhaps an individual deceives herself through some other way, instead of manipulating her belief about p. Such an individual would be categorized as honest, but would really be deceptive, just in another way that the experiment is not able to capture.

Other possible explanations for observed discrepancy between adopted and true beliefs. The discrepancy between adopted and true beliefs expressed by manipulative subjects, and particularly by selfish and manipulative (SM) subjects, raises some questions. First, how can we be sure that the raffle task, and not the donation task, is what elicits their true beliefs? This is because, in the raffle task, their own money is at stake. If they do not respond according to their true beliefs in the raffle task, then these subjects are not making an optimal choice. For example, a subject who truly believes that p is likely to be a small number, say 30%, then they would obviously prefer a lottery that has a winning probability of 40% over a lottery that has a winning probability of p). Moreover, given that everyone's (i.e., selfish and altruistic subjects') truly believe that p will take a value of 50%, which is also the actual expected value of p, it is extremely unlikely that they do not understand the raffle task.

Second, what motivates subjects, and particularly manipulative subjects, to adopted false beliefs in the donation task? My explanation for this is that subjects deliberately manipulate a belief that would cast a more forgiving light on them. They do not adopt a belief of something too low, e.g., 0%, because it also needs to be something they themselves find convincing. An individual who manipulates her belief in order to maintain a self-image of an altruistic person may imagine how she might explain this behavior to someone else, e.g., some kind of moral police, who does not know her exact level of pessimism but has some idea (e.g., a range) about it. She claims to be as pessimistic as possible, given this idea.

Third, what if adopted beliefs differ from true beliefs because subjects are forced to think more carefully about their beliefs in the raffle task, and not because they deliberately manipulate their beliefs? While it is certainly plausible that subjects' beliefs may change with more careful consideration, I find that only selfish subjects have a significant difference between their adopted and true beliefs, whereas altruistic subjects do not. If careful consideration is a factor in driving the difference, one should expect to see a significant difference for both altruistic and selfish subjects. Further, in this case, adopted beliefs would be different from true beliefs in both directions – not just in the direction of pessimism

– because there is no reason for carelessness to result in a one-sided difference. That is, some subjects would adopt more pessimistic beliefs than their true beliefs, while others would adopt more optimistic beliefs. The fact that selfish subjects who manipulate beliefs predominantly adopt more pessimistic beliefs about p implies a systematic bias in one direction, and cannot be explained as the result of thinking carefully.

Fourth, does consistency between subjects' adopted and true beliefs imply that they are honest? The answer is not necessarily. Instead, a subject might have chosen an adopted belief randomly simply because of not caring, but the random response happens to align with her true belief. Or even if she deliberately adopted her true belief in this particular instance, it is not necessary that they are always an honest person. It could be that she tries to look for an excuse but is unable to find one, and is ultimately forced to sacrifice either her self-image or her earnings.

Other applications of this framework. This framework shows how people can maintain a favorable self-image without taking an action that is usually required for that self-image, illustrating it with the self-image of an altruistic person. However, this framework can be applied to any favorable self-image, e.g., the self-image of an honest person, of someone who cares for animals/environment, or of someone who is not wasteful, etc. In a situation where an individual's self-image is challenged, i.e., an individual is tempted to act in a way that would damage her self-image, she tries to come up with an excuse that allows her to take that action without damaging her self-image.

For example, suppose that an individual is not able to finish her dinner, but is reluctant to throw away the remaining food because she thinks that would be wasteful and immoral. In other words, throwing the food away would damage her self-image. By coming up with an excuse (e.g., by saying to herself, "*I cannot eat it myself because it is also important for me to control my diet, and I cannot give it to anyone else because no one would* *like to eat this; and perhaps throwing it in the garbage is not so bad because a raccoon might benefit from it.*") she can arrive at a fairly optimal solution. As another example, suppose an individual has a self-image of a healthy person, and particularly defined as someone who eats less than 2000 calories a day. She counts and records her calories after each meal, making sure she does not exceed 2000 calories in a day. However, as she serves some cheesecake to her children, she licks the knife clean (taking in about 200 calories) while adopting the belief that this is such a negligible quantity of cheesecake that she does not need to record it. (She has her cake and eats it too, both figuratively and literally.)

Policy implications. The main conclusion of this paper is that there are several ways, including some very subtle ones, that people can use to reduce cognitive dissonance. An obvious and established way of increasing desirable behavior (such as donations) is to make people feel bad about an undesirable action (e.g., about not donating money). But knowing that individuals can still manage to find an excuse to not donate, and particularly an excuse that allows them to continue thinking of themselves as altruistic, tells us that we need to take measures to reduce the room to make excuses. For example, mentioning the effectiveness of a charity or reducing transaction costs of donating (e.g., "*to make a $10 donation, simply text Donate10 to 1234*") are examples of measures that can reduce the number of available excuses to not donate.

1.7 Conclusion

This study explores whether people are able to convincingly deceive themselves about a belief they hold, and if such deception eliminates their self-created need to behave altruistically. By using an experimental setting that minimizes the effect of social-image and reputation concerns, I provide an environment where the only plausible motivations to donate are pure altruism and self-image of an altruistic person. Within this environment, an excuse for not donating is available for subjects to make. I observe the proportion of subjects who use this excuse and determine if there is a correlation between selfish behavior and making this excuse. I find that there is: subjects who choose *not* to donate are the ones who use this excuse because it helps them justify their selfish behavior. I also show that subjects have no reason to make this excuse, except for a preference for thinking of themselves as more altruistic individuals than they truly are.

CHAPTER 2

Systematically Skewed Beliefs

Abstract

We have a tendency to overestimate the commonness of our own preferences. This paper investigates if this overestimation is limited to our beliefs about a single statistic (such as the median or mean), or if it extends to our beliefs about an entire distribution. In a laboratory experiment, I elicit subjects' beliefs about the distribution of other people's types, and check if subjects systematically skew their beliefs in relation to their own types. Experiment results strongly suggest that they do. In particular, all subjects overestimate the density around their own types. Using experimental data to estimate PDFs of actual and predicted distributions, I show that predicted distributions can be presented as skewed transformations of the actual distribution, where the size and direction of skewness is determined by a subject's own type.

2.1 Introduction

Our beliefs about other people's preferences and choices are subject to a systematic bias – a bias that causes us to overestimate the percentage of people who are similar to us. Evidence shows, for example, that selfish subjects overestimate the proportion of people who are selfish (Iriberri and Rey-Biel, 2013), trustworthy subjects are more likely to think that others are trustworthy (Butler, Giuliano and Guiso, 2015), heavy people overestimate the proportion of others who are heavy (Proto and Sgroi, 2017), and people who get lucky themselves tend to attribute luck to other people's achievements (Cartwright and Wooders, 2020). The psychology literature provides even more examples – innumerable in fact – of people overestimating the commonness of their own choices.[1]

The purpose of this paper is to investigate if this overestimation is limited to our beliefs about a single statistic (such as the median or mean) of a distribution, or if it affects our beliefs about an entire distribution in a systematic way. This will help us better understand this bias and its underlying causes, and provide additional empirical evidence that supports the existence of biased beliefs. In addition, this knowledge can improve individual-level decisions as well as those made by policy makers, and can help theoretical economists in deciding when it is inappropriate to assume that distribution priors are unbiased.

This paper is closely related to the literature on *social projection* (Katz, Allport and Jenness, 1931), *false consensus effect* (Ross, Greene and House, 1977), *self-generated comparison information* (Sanders and Mullen, 1983), *truly false consensus effect* (Krueger and Clement, 1994), *interdependent preferences* (Iriberri and Rey-Biel, 2013), *self-centered beliefs* (Proto and Sgroi, 2017), *own experience bias* (Cartwright and Wooders, 2020),

[1] Krueger (1998) provides an excellent overview of this literature.

and *correlated beliefs* (Cason, Sharma and Vadovič, 2020), all of which, broadly speaking, show that people systematically overestimate the commonness of their own choices or characteristics (but differ in their more precise definitions). Since this paper hypothesizes that our beliefs about the distribution of types are systematically skewed, it is also related to the literature on *overconfidence* (Moore and Healy, 2008), which shows that individuals systematically overestimate their own skills relative to other people's skills.

Since psychologists started studying this bias long before economists, the literature contains a lot more psychology experiments relative to economics experiments. The main goal of all experiments is eliciting subjects' beliefs about other people's preferences or abilities.[2] However, psychology experiments do not provide incentives for reporting true beliefs, something that is considered crucial in economics experiments. Another important methodological difference between economics and psychology experiments is the use of deception, which seems to be prevalent in psychology experiments that study this bias.[3] This means that from the perspective of experimental economics, most of these experiments are unsatisfactory. Therefore, from this point forward I will only discuss economics experiments.

Economics experiments elicit beliefs in an incentivized manner, but contain some other problems. For example, some of them ask their subjects to guess the "average" of the choice made by other subjects, but fail to specify what "average" really means, leaving subjects free to interpret it as the median, mean, or mode, or possibly even some other statistic.[4]

[2] For example, in *false consensus* experiments, subjects are typically given two tasks: (1) answer a dichotomous question, and (2) guess how many other subjects will choose the same response as them. In *overconfidence* experiments, by contrast, subjects are typically given only one task: to rank their skills about something (e.g., driving skills) relative to other people's skills.

[3] Engelmann and Strobel (2000, p. 243), in reviewing the false consensus literature, write, "Furthermore, in many of the social psychological experiments where information about other people's decisions is provided, this information is rigged and the subjects are clearly deceived." For a more general discussion on methodological differences between economics and psychology experiments, see Camerer (1997) and Hertwig and Ortmann (2003).

[4] Benoît and Dubra (2011, p. 1595) critique an experiment in which subjects rank their driving skills

Another problem is that they elicit beliefs about only one statistic of the distribution, shedding no light on the mechanism of this bias (e.g., if it there is can cause subjects to systematically skew the entire distribution.

This paper presents an experiment that elicits subjects' beliefs about an entire distribution of responses (as opposed to a single statistic of the distribution), investigating if subjects' beliefs about the distributions are systematically skewed in relation to their own types. The experiment consists of two main tasks. In the first task, subjects choose how much money to donate from an endowment of 10 tokens, revealing how altruistic they are on a scale from 0 to 10. In the second task, subjects are asked to guess, and rewarded for accuracy, the proportion of subjects who will donate 0 tokens, 1 token, and so on – for a total of 11 guesses.

This is most closely related to the experiment by Iriberri and Rey-Biel (2013), which consists of three stages. In the first stage, they ask each subject to play several modified dictator games. In each dictator game, a dictator needs to choose one of three possible outcomes.[5] Based on the subjects' choices, they assign them "types". In the second stage, they ask subjects to guess, for each modified dictator game, how many of 10 other subjects, selected randomly, they think chose each option. After the second stage, they divide subjects into two groups. One group is shown the actual choices made by a (different)

relative to others' by saying, "Svenson ignores this issue and, in effect, asks each subject for a summary statistic of her beliefs without specifying what this statistic should be. There is no way to know if subjects responded using the medians of their beliefs, the means, the modes, or some other statistic. As a result, it is unclear what to make of Svenson's data. Svenson's experiment is hardly unique in this respect: much of the overconfidence literature, and other literatures as well, share this feature that the meaning of responses is not clear." Even the experiments that specify the statistic are limited by the fact that they only elicit one statistic of the distribution. As another example, Dominitz (1998, p. 375) writes, "In particular, what feature of the subjective probability distribution determines the category selected by respondents? Is it the mean? Or perhaps it is the median or some other quantile. Or perhaps it is the category that contains the most probability mass."

[5]For example, in one of those games, the dictator's options are (7, 10), (7, 24), and (8,17). There are 16 such dictator games that each subject plays.

group of 10 subjects and the other group is not. Lastly, the first stage is repeated, i.e., subjects play the same modified dictator games that they played earlier. Iriberri and Rey-Biel (2013) find that (i) all subject types overestimate the commonness of their own types, and (ii) subjects play the dictator game more selfishly the second time around, irrespective of their treatment groups. However, my experiment is different from theirs in at least three ways. First, subjects in my experiment are motivated only by altruism, whereas their subjects may be motivated by altruism, fairness, and/or welfare concerns. This makes it easy for me to interpret what subject type represents. Second, I assign subjects types solely based on their actions, whereas Iriberri and Rey-Biel define subject types based on subjects' motivations, determining which requires them to make some subjective judgments. Third, subject types in my study can be represented by a single, continuous random variable that is larger for higher types and smaller for lower types. This is because I define a subject's type as a measure of how altruistic she is, measured by how much money she chooses to donate on a scale from 0 to 10. By contrast, Iriberri and Rey-Biel have four different types of subjects, namely Selfish, Welfare Maximizing, Inequity Averse, and Competitive, which do not have any numerical relation between them.

The main testable hypothesis of this paper is that individuals' beliefs about the distribution from which their own type is drawn are systematically skewed in relation to their own types. If, for example, both low and high types of subjects are equally likely to inaccurately predict the distribution of types, then beliefs are not systematically skewed. If, however, low types overestimate the distribution mass around low types and high types overestimate the distribution mass near high types, then beliefs are systematically skewed. That is, if subjects' beliefs are inaccurate in a particular direction that depends on subjects' own types, then there is a systematic bias in how subjects form beliefs about the distribution of other people's types.

The results of this study confirm this hypothesis, that subjects' beliefs about distributions are systematically skewed. In particular, the more altruistic subjects are, the more they overestimate the density around their own type (which, naturally, results in an overestimation of the mean and median altruism levels). This is consistent with the hypothesis that beliefs about distributions can be presented as a transformation of the actual distribution, and specifically a transformation that only affects the distribution's skewness. I further observe that selfish subjects make the most accurate predictions about the distribution, and propose that given that selfish subjects constitute the majority of the population, their accuracy is likely *because* of this systematic bias, not due to a lack of it.

This finding has several important implications. First, it helps us better understand the nature of this bias by clarifying that this is not a systematic overestimation of the mean, but rather an overestimation of the *skewness*. Second, this result suggests that this is not a misperception about one's *own* location in the distribution, but a misperception about the *distribution* itself. Third, because altruistic subjects overestimate (and not underestimate) the proportion of altruistic subjects, such a misperception cannot be motivated by self-image concerns. If subjects wanted a self-image of someone who is altruistic, then even altruistic subjects would have an incentive to underestimate the proportion of altruistic subjects. Therefore, this systematic skewness is likely a genuine cognitive bias, not a self-serving bias. Lastly, this understanding enables us to make more accurate predictions about beliefs held by people about other people. The next two sections present a theoretical framework and the experiment design, respectively. These are followed by a presentation and discussion of the results.

2.2 Theoretical Framework and Hypotheses

Let $\theta \in [\underline{\theta}, \overline{\theta}]$ be a (continuous) random variable that represents an individual's *altruism level,* with higher values of θ indicating greater altruism. That is, values of θ are determined by a continuous probability distribution function $f(\theta)$ and a cumulative distribution function $F(\theta) = \int_0^a f(x)dx$, with $F(\underline{\theta}) = 0$ and $F(\overline{\theta}) = 1$. For each $i \in \{1, 2, ..., N\}$, let θ_i be the realized value of θ that is assigned to individual i and suppose individual i is asked to estimate the distribution $f(\theta)$ without being given any other information.

This brings us to the main testable hypothesis of this study, that each individual will overestimate the mass around their own altruism level, but otherwise estimate the distribution very accurately. In other words, each individual's belief about $f(\theta)$ is biased by her own altruism level, and the extent of the bias can be predicted by her own position in the distribution.

Hypothesis 1. Denote the true distribution by $f(\theta, \alpha)$ where α is the distribution's skewness parameter, and denote an individual i's belief about the distribution by $\hat{f}_i(\theta, \alpha)$. Then $\hat{f}_i(\theta, \alpha) = f(\theta, \hat{\alpha}(\theta_i))$, that is, there exists a function $\hat{\alpha}(\theta_i)$ that can predict each individual's belief about the distribution.

To add some context regarding how individuals are asked to estimate a distribution, let us think about a game in which there are N players, each of whom is endowed with M dollars and given an opportunity to donate some of it, say $x \in [0, M]$, to a charity. By definition (of altruism), more altruistic individuals will prefer to donate more money, which makes x a monotonically increasing function of θ. Therefore, as individuals estimate the distribution of x, they are indirectly estimating the distribution of θ.

Hypothesis 1 can be tested by using a non-parametric measure of skewness, such as the

Pearson measure of skewness (Pearson, 1895), $(\mu - M)/\sigma$.[6] The benefit of this method is that it does not require making any assumptions about the underlying distribution and involves only a simple calculation. However, a limitation of this method is that it allows predictions about the distribution to be very different from the actual distribution, not just in terms of skewness but in terms of any feature. For example, if altruism levels are distributed according to an exponential distribution but an individual believes that they are distributed according to a multi-modal normal distribution, then the skewness value would not be enough to tell us all the ways in which individuals' perceptions are different from reality. Therefore, in order to test if it is indeed only skewness that an individual misconstrues, I supplement this analysis with a parametric approach in which I hold certain distribution characteristics constant across perceptions and reality. As Section 2.4.2 will show, the non-parametric and parametric methods conform with Hypothesis 1.

To test Hypothesis 1 using a parametric approach, suppose that individuals know that θ is distributed according to a Gamma distribution, but does not know the distribution's (true) shape and scale parameters, α and β.[7] That is, the (true) probability density function is

$$f(\theta|\alpha, \beta) = \frac{\theta^{\alpha-1} e^{-\theta/\beta}}{\Gamma(\alpha)\beta^{\alpha}}$$

where $\Gamma(\alpha)$ is the gamma function. An individual's belief about the distribution depends on her own altruism level, θ_o. Since a gamma distribution is characterized by the parameters α and β, an individual's belief about the distribution can be written in terms of her beliefs about α and β. Denoting an individual's beliefs about the parameters by $\hat{\alpha}(\theta_o)$ and $\hat{\beta}(\theta_o)$, the individual's prediction about the PDF of the distribution is $\hat{f}_i = f\left(\theta|\hat{\alpha}(\theta_i), \hat{\beta}(\theta_i)\right)$. A

[6]Letting M be median would result in the Pearson Median Skewness whereas letting M be the mode would result in the Pearson Mode Skewness.

[7]I use a Gamma distribution because of its generality and the fact that its skewness is a function of a single parameter (called the shape parameter).

convenient feature about the gamma distribution is that its skewness is $\frac{2}{\sqrt{\alpha}}$, a function of the distribution's shape parameter. Given that this paper hypothesizes that individual incorrectly perceive the distribution's skewness, but not other features, this means that the hypothesis is that $\hat{\alpha} \neq \alpha$ but $\hat{\beta} = \beta$. Moreover, since the extent of deviation of $\hat{\alpha}$ from α depends on an individual's own altruism level, and the extent of that deviation depends on θ_i, we can represent $\hat{\alpha}(\theta_i, \alpha)$ as a function of an individual's altruism level and the value of the true distribution's shape parameter.

Proposition 1. The function $\hat{\alpha}(\theta_i, \alpha)$ is monotonically increasing in θ_i.

Proof. An individual whose altruism level is $\theta_i = \mu$ will not, by assumption, have an inaccurate belief about the distribution. Suppose an individual's altruism level is greater than the average altruism level, i.e., $\mu < \theta_i$. Given that she overestimates the mass around her own position in the distribution, her prediction about the mean, $\hat{\mu}$, will be closer to θ_i than μ, i.e., $\mu < \hat{\mu}$. Given that $\mu = \alpha\beta$ and $\hat{\mu} = \hat{\alpha}\beta$ (since this is a gamma distribution), this implies $\alpha < \hat{\alpha}$.

Hypothesis 2. The function $\hat{\alpha}(\theta_i, \alpha)$ increases at a decreasing rate, i.e., $\frac{\partial \hat{\alpha}}{\partial \theta_i} > 0$ and $\frac{\partial^2 \hat{\alpha}}{\partial \theta_i^2} < 0$.

This hypothesis means that the farther an individual is from the mean, the more she skews the distribution, but at a decreasing rate. This captures the idea that individuals who belong to the tail of the distribution are better aware of the fact that they are far from the mean than those who are closer to the middle of the distribution, and this awareness reduces their bias. For example, someone who is very altruistic is more likely (than someone who is slightly altruistic) to know that she is more altruistic than the average person.

2.3 Experiment Design

To test whether subjects's beliefs are systematically skewed according to their types, I conduct a laboratory experiment that consists of two main tasks, namely donation task and guessing task, followed by a demographic questionnaire. In the donation task, subjects are endowed with some money and first choose a charity from a list of 11 popular charities, each of which supports a different cause.[8] After choosing a charity, subjects decide how much of their endowment to donate to that charity. After submitting these responses, they are taken to a guessing task where they guess how many experiment participants will donate 0 tokens, how many will donate 1 token, and so on. Lastly, subjects complete a short questionnaire that collects basic demographic information, such as gender, age race, religiousness, education, and income, and contains some qualitative questions that ask how different they think they are from most other people.

2.3.1 Implementation

To conduct this experiment, I recruited 130 subjects using Amazon Mechanical Turk (MTurk), which is a micro-employment platform that allows employers to hire workers to perform small tasks for modest compensation. Anyone who is at least 18 years old can sign up as a *worker* on this platform and see available tasks posted by *requesters*. I posted this experiment in the form of a computerized survey that was visible to all workers in the United States. Amazon acts as an intermediary between workers and requesters. This guarantees subjects payments promised to them and also ensures their anonymity.

[8] I let subjects choose a charity so that they are not able to use the pre-specification of charity as an excuse for not donating. That is, a subject might prefer to donate towards, say, cancer research instead of poverty relief. Such a preference would provide a subject with an excuse to not donate, allowing her to think of herself as generous without making a donation. Therefore, allowing subjects to choose their own charity reduces the number of possible excuses available to them.

Requesters do not know workers' identities and, instead, only see workers as a 14-20 characters long alphanumeric ID that is assigned to each worker by Amazon. Amazon takes several measures to prevent fraudulent activity (e.g., one person cannot create multiple worker accounts; workers need to continuously exhibit traits of human intelligence, etc.).[9]

To ensure that subjects comprehended the experiment instructions, they were required to correctly answer some qualifying questions in order to participate in the experiment. Those who passed the qualifying questions were guaranteed a participation payment of $3. In addition to this, they were given an endowment of 10 tokens (equivalent to $1.30) in the donation task, and a performance-based reward in the guessing task, of 1 token for each accurate guess. The average subject earnings were $5 and the average time taken to complete he survey was 20 minutes. The actual survey questions, as seen by subjects, are shown in Section B.1 of the Appendix, and summary statistics of subject characteristics are shown in Table 1.1 in the Appendix.

2.3.2 Donation Task

Immediately after passing the qualifying questions, subjects are taken to the first task, titled *donation task*, and are given 10 tokens (equivalent to $1.30). Subjects are first asked to choose a charity from a list of 11 popular charities, each of which supports a different cause. I let subjects choose a charity to prevent a selection bias where the donors are limited to subjects who care for a particular cause. If there is a pre-specified charity (say, a cancer research organization) then subjects who would have given to another cause (e.g., poverty relief) would get discouraged to donate. In this case, subjects' donations would not be indicative of their altruism but of a combination of altruism *and* a preference for

[9]For more details on these measures, see https://blog.mturk.com/important-updates-on-mturk-marketplace-integrity-worker-identity-and-requester-tools-to-manage-206e4e90da0c. Amazon also charges requester a considerable fee (of 40% of what they want to pay workers) for providing this service.

donating to a particular cause. After choosing a charity, subjects are asked to decide how many tokens from their endowment of 10 tokens they would like to donate to that charity. They get to keep the remainder of their endowment. On the same screen, subjects may also provide an email address if they would like a donation receipt, as proof, once the donations have been made.

2.3.3 Guessing Task

After submitting their choices in the donation task, subjects are taken to a *guessing task*, in which they are asked to make a total of 11 guesses. They are reminded that many MTurk workers are currently participating in this experiment and will be making the same donation decision they just made. In this task, subjects need to guess how many subjects, including themselves, will donate 0 tokens, how many will donate 1 token, and so on, until the last guess of how many will donate 11 tokens. Subjects make this guess in terms of percentage, because the exact number of subjects cannot be known until all data is obtained. However, they know that this number will be more than 100.

To encourage subjects to report their true beliefs and to think carefully about each guess, subjects are rewarded for accuracy. For each guess that is within 1 percentage-point of the actual percentage, subjects will be given 1 token. This means that can earn up to 11 tokens from this task.[10] Because the actual distribution of donation amounts can only be calculated after all responses are received, these accuracy rewards are paid to them separately (approximately one week later) as a *worker bonus* on MTurk.

[10]Subjects are told, "For each accurate guess, you will earn one token (for a maximum of 11 tokens). We will consider a guess to be accurate if it is within one percentage-point of the actual percentage. For example, if we find that 5.67% of participants donated 0 tokens, then anyone who makes a guess of either 5% or 6% will receive one token for that guess."

2.4 Results

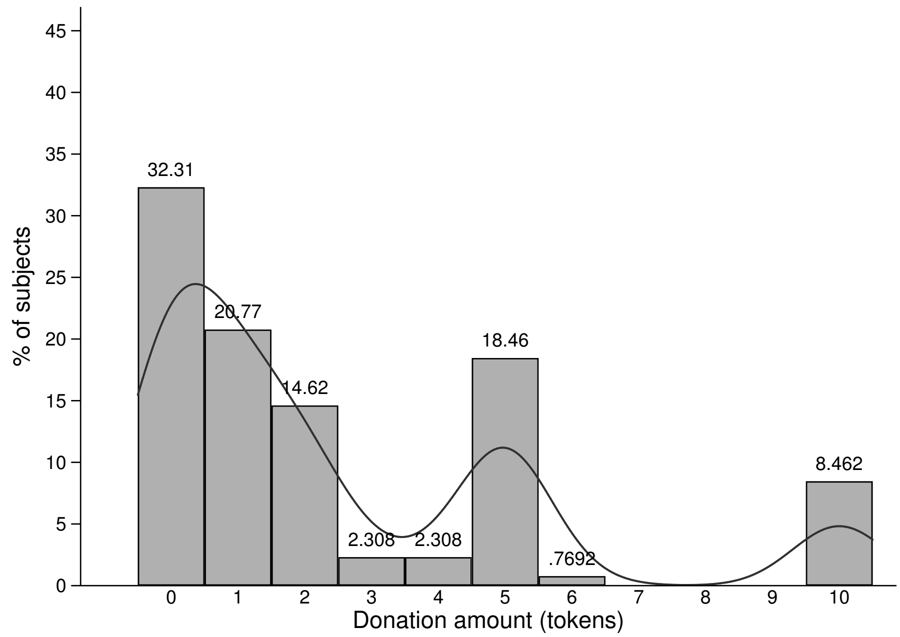
2.4.1 Donation Choices and Subject Types

Commonness of donation amounts Subjects' first task is to decide how many tokens to donate to a charity. The vast majority (about 68%) of subjects donate 2 or less tokens from their endowment of 10 tokens. Figure 2.1a shows the popularity of each choice. The most popular choices are 0 tokens, 1 token, and 5 tokens, with 32% of subjects (42 out of 130) donating 0 tokens, 21% (27 out of 130) donating 1 token, and 18% (19 out of 130) donating 5 tokens.[11] The spikes at 0, 5, and 10 are consistent with a robust finding in dictator game experiments that subjects have a preference for behaving fairly, completely altruistically, or completely selfishly (e.g., Andreoni and Bernheim, 2009; Fehr and Schmidt, 1999).

Assigning Subjects Types Based on subjects' choices, I assign each subject a *type* of either low, medium, or high. I do this by assigning a *low* type to subjects who donate 0, 1, or 2 tokens, a *medium* type to subjects who donate 3, 4, 5, 6, or 7 tokens, and a *high* type to subjects who donate 8, 9, or 10 tokens. Since subjects are assigned types solely based on their actions and assuming that actions reveal their exogenous preferences for donating money, subjects' types can be considered as exogenously determined. About 68% of subjects (88 out of 130) reveal themselves as low types, 24% (31 out of 130) reveal themselves as medium types, whereas the remaining 8% (11 out of 130) reveal themselves as high types.

[11] A significant proportion of subjects (35%) choose either 1 or 2 tokens, which is a non-zero amount, and at the same time a very small proportion of their endowment. It is likely that subjects donating 1 or 2 tokens are motivated solely by warm-glow, and specifically self-image concerns, because they are able to convince themselves that they are altruistic even by donating such a small amount.

Figure 2.1: Actual Distribution



This histogram presents the proportion of subjects who donate 0, 1, ... 10 tokens. For example, the first bar shows that 32.3% of subjects (i.e., 42 out of 130 subjects) donate 0 tokens. The fitted curve is a kernel density function estimated using subject responses.

2.4.2 Beliefs about Other Subjects

After submitting responses in the donation task, subjects are taken to a guessing task where they predict the commonness (or popularity/frequency) of each donation amount. Subjects are told that over 100 other subjects are participating in this experiment, and are asked to predict what proportion of all subjects will donate each donation amount. That is, each subject predicts what proportion of subjects will donate 0 tokens, what proportion will donate 1 token, and so on, for a total of 11 predictions.[12] In other words, each subject is asked to predict the distribution shown in Figure 2.1. To incentivize subjects to report

[12]Subjects are asked to make predictions in terms of percentages and are told that the sum of their predictions should end up being 100%. Although subjects may proceed even if their predictions do not sum to 100%, for 130 (out of an initial sample of 131) subjects the sum was in fact 100%. There was one subject whose predictions added up to 550%, and that subject was dropped from the data due to this error.

what they truly believe the distribution to be like, each subject is given a reward of 1 token for making an accurate prediction (where accuracy is defined as being no more than 1 percentage point away from the true value). Thus, subjects can potentially earn up to 11 tokens from this task.

Non-Parametric Skew The skewness of any distribution can be measured by taking the difference between the mean and median and dividing it by the standard deviation, i.e., $S = \frac{\mu - M}{\sigma}$. Based on subjects' own donation choices ($\mu = 2.477, M = 1, \sigma = 2.94$), this turns to be $S_{\text{Actual}} = 0.502$, a positive skew. Based on the guesses made by low type subjects, the average (or representative) low type subject estimates this skewness to be $S_L = 0.490$, which is remarkably accurate. Medium type and high type subjects, by contrast, guess the skewness to be -0.250 and0.132, respectively, significantly lower than the actual skewness.

Figure 2.2 shows the average predictions made by low types, medium types, and high types.[13] As the first panel of the figure shows, low types predict that, on average, 39% of subjects will donate 0 tokens. Comparing this prediction with actual proportion of subjects who donate 0 tokens – which is about 32% (see Figure 2.1) – shows that low types overestimate the commonness of completely selfish behavior by about 7 percentage points[14] ($p = 0.018$, two-tailed t-test). Likewise, medium types overestimate the commonness of donating 5 tokens by about 7 percentage points (or 39%). Lastly, high types overestimate the commonness of completely altruistic behavior (i.e., donating 10 tokens) by about 11.5 percentage points (or 136%).

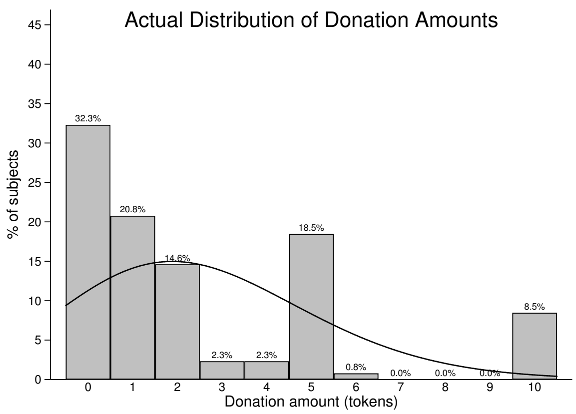
Parametric Approach To explicitly check how accurate or inaccurate each type's belief about the distribution is relative to the true distribution, Figure 2.3 presents each type's

[13]Figure 2.6 in the Appendix shows that beliefs held by subjects who donate 0, 1, or 2 tokens are very similar. This provides further justification for classifying subjects into low, medium, and high types.

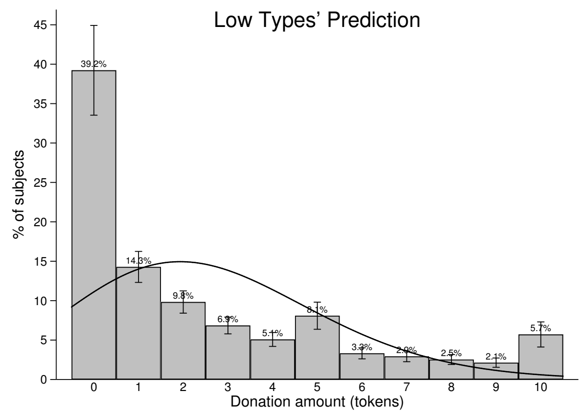
[14]In terms of *percentage*, 39 points is *21%* greater than 32 points, which is a significant overestimation.

Figure 2.2: Predictions made by Low, Medium, High types

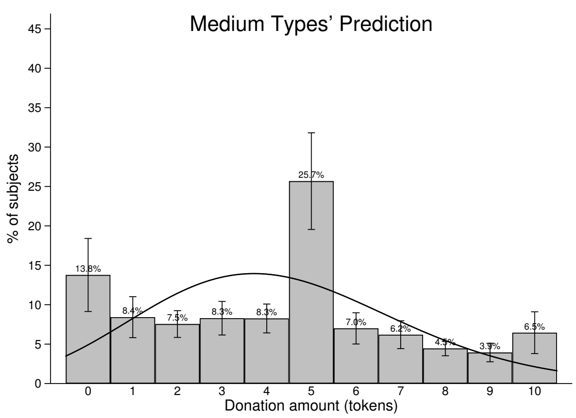
(a) Actual Distribution



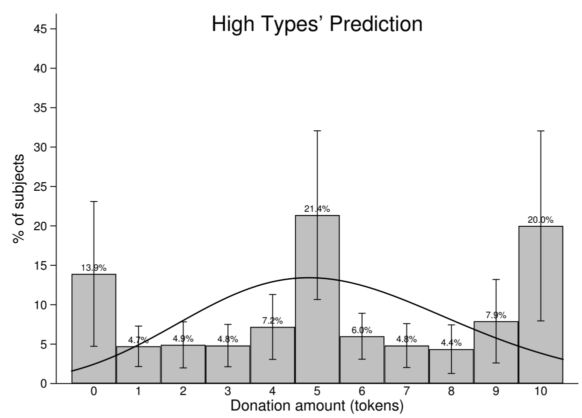
(b) Predicted Distribution



(c) Predicted Distribution



(d) Predicted Distribution



Notes. All histograms are displayed with a fitted gamma distribution. The histogram in panel (a) shows the actual distribution of donation amounts and the only difference between this and Figure 2.1 is that it is fitted distribution with a gamma distribution. Panels (b), (c), and (d) show, respectively, low, medium, and high types' beliefs about the distribution in panel (a). As an example, the first bar in panel (a) shows that low types think that 39% of subjects will donate 0 tokens, i.e., overestimating the true proportion of 32% by about 7 percentage points. The height of each bar represents the average (mean) value of guesses, and error bars represent 95% confidence intervals.

predicted distribution and the actual distribution on the same graph. A convenient feature about the Gamma distribution is that there is a simple formula to measure its skewness: $^2/_{\sqrt{\alpha}}$. The parameters of the fitted actual distribution (shown in black in each panel of Figure 2.3) turn out to be ($\alpha = 21.076, \beta = .592$), resulting in a skewness coefficient of 0.436.

The red probability distribution functions represent fitted gamma distributions based on average predictions made by subjects of each type. Alternatively, it may help to think of these average predictions as predictions made by average, or *representative*, subjects of each type, where a representative subject of a particular type is someone whose choices are equal to the average of the choices made by all subjects of that type. So, for example, a *representative low type* subject is someone whose donation choice is equal to the average of the donations made by all low type subjects, and whose belief about the percentage of subjects who will donate $x \in \{0, ..., 10\}$ tokens is equal to the average of what all low type subjects' beliefs about that percentage. Each of the predicted distribution functions are estimated under a *scale-preserving transformation* assumption, i.e., subjects make predictions by transforming the actual distribution function while maintaining the original scale of the distribution. A gamma distribution is characterized by a *shape parameter* α, and a *scale parameter* β, and a PDF of $f(x|\alpha,\beta) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}$. A scale-preserving transformation means that subjects' beliefs about β are accurate, and only their beliefs about α are inaccurate.[15] Overall, I estimate the function $\hat{\alpha}(\theta_i, \alpha)$ as: $\ln \hat{\alpha} = \ln \alpha + 0.0983 \ln \theta_i$.

Panel (a) of Figure 2.3 shows that low types make incredibly accurate predictions about

[15] In the Appendix, I show what predicted distributions would look like if we do not make this assumption and instead, allow subjects to hold incorrect beliefs about both α and β. Moreover, since a gamma distribution must start at zero but in this case there is a positive probability of donating 0 tokens, I shift the origin of the distribution by 10 units to $x_0 = -10$. This helps in obtaining much better fitted values without affecting the data – as long as I keep the same x_0 in all fitted distributions.

the distribution of donation amounts. This result is not just consistent with an egocentric bias, it is most likely *because* of an egocentric bias. Given that low types constitute the vast majority (68%) of all subjects, having an egocentric bias would actually help them make more accurate predictions. That is, assuming that others are like them is not such a bad thing because most other people *are* like them. The predictions made by low types correspond to fitted values of ($\alpha = 21.154, \beta = .592$), resulting in a skewness coefficient of 0.435, nearly identical to the actual skewness coefficient.[16]

Panels (b) and (c) show the predictions made by medium and high types respectively. As hypothesized, even medium and high types exhibit an egocentric bias. Medium types predict the mean of the distribution to be around 4.3 tokens, which is considerably closer to their own type of 5 than the actual mean of 2.48. The skewness coefficient of their fitted distribution turns out to be 0.407, which is smaller than the actual skewness coefficient of 0.436. High types, whose actual type is higher than that of medium types, overestimate the mean even more than medium types do. Their estimated skewness coefficient is 0.392, suggesting that they skew the distribution even more than medium types do.

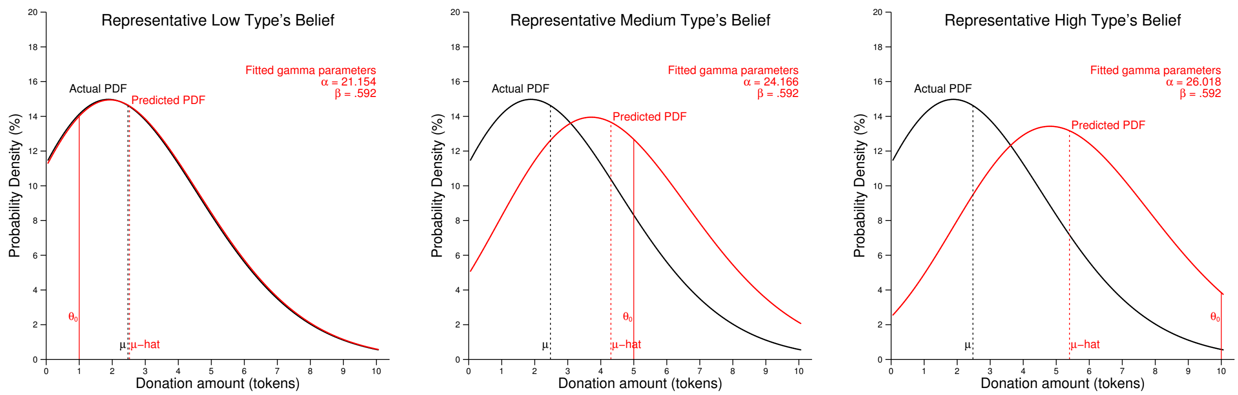
Beliefs about subject types Since subject are categorized into types based on their donation choices, as subjects estimate the commonness of each donation amount, they indirectly estimate the proportion of low, medium, and high types of subjects. For example, if a subject estimates that about 10% of all subjects will donate 0 tokens, another 10% will donate 1 token, and another 10% will donate 2 tokens, she indirectly estimates that 30% of all subjects will be low types.

Figure 2.4 presents the actual distribution of subject types, and estimates of this

[16]This is consistent with a robust finding in the literature, that "subjects holding majority positions generated more accurate estimates of consensus than did subjects holding minority positions" (Sanders and Mullen, 1983, p. 65)

Figure 2.3: Accuracy of Predictions made by Low, Medium, High types

(a) Low Types' Beliefs (b) Medium Types' Beliefs (c) High Types' Beliefs

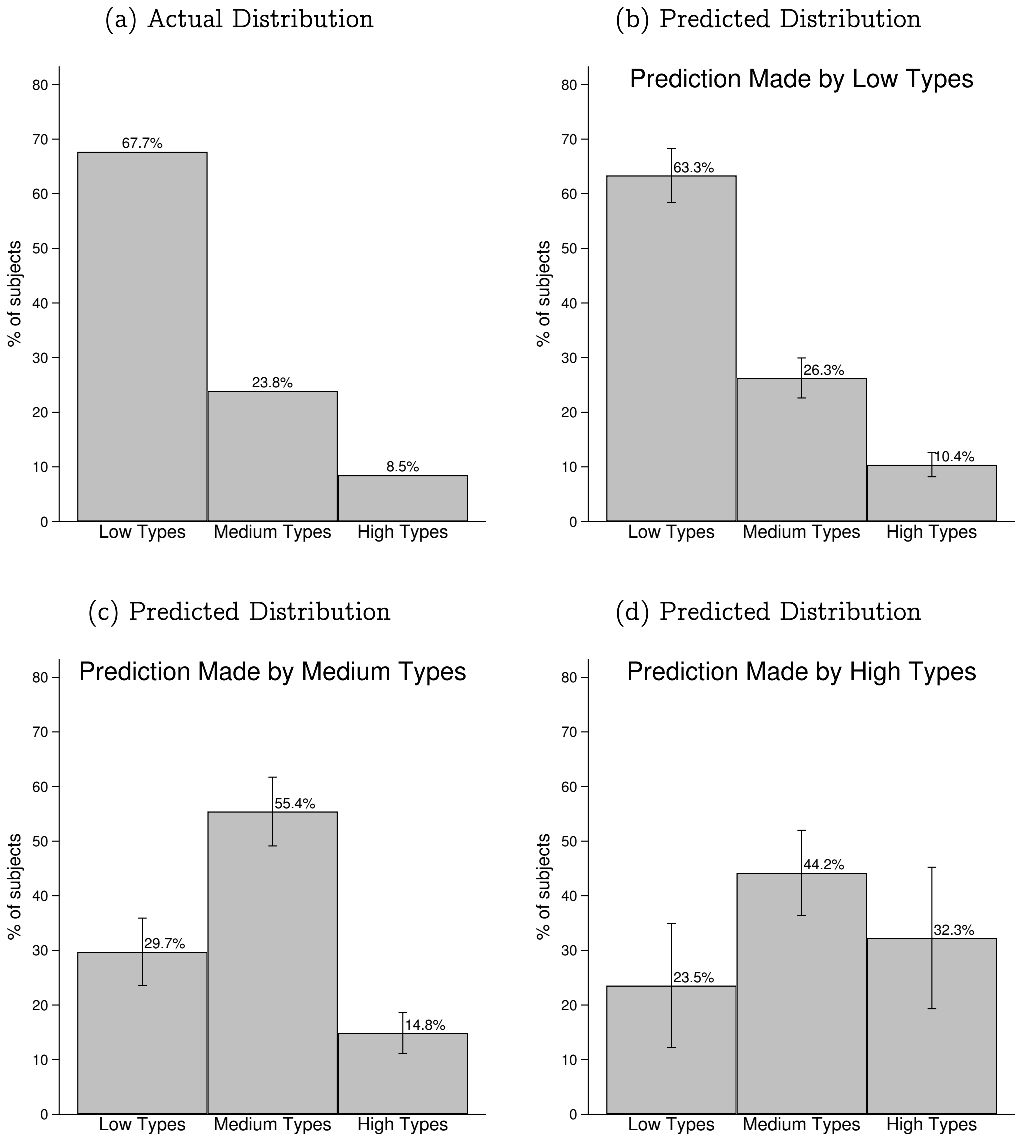


distribution made by each subject type. Panel (a) of Figure 2.4 presents the actual distribution of types, and panels (b), (c), and (d) present the beliefs of low, medium, and high types, respectively, about this distribution.

It turns out that low types' estimates are considerably more accurate than those of medium and high types. For example, they estimate that the proportion of low types will be 63.3%, which is not significantly different from the actual proportion of low types (67.7%) at a 5% significance level ($p = 0.084$ for a two-tailed t-test); the same is also true about their estimates of medium and high types. By contrast, medium and high types greatly underestimate the proportion of low types while greatly overestimating the proportion of their own types. Table 2.1 presents resulting p-values of t-tests that check the accuracy of each types' average estimate. Consistent with the bar graphs in Figure 2.4, medium and low types' estimates are considerably inaccurate, and demonstrate a systematic bias.

While low types overestimate the proportion of subjects who donate 0 tokens, interestingly, they do not overestimate the proportion of low types (i.e., those who donate 0, 1, or 2 tokens).

Figure 2.4: Distribution of Subject Types



Notes. Panel (a) shows the actual distribution of low, medium, and high types in the sample of 130 subjects. Panels (b), (c), and (d) show the beliefs about this distribution held by low, medium, and high types, respectively.

Table 2.1: T-tests Results for Accuracy of Predictions

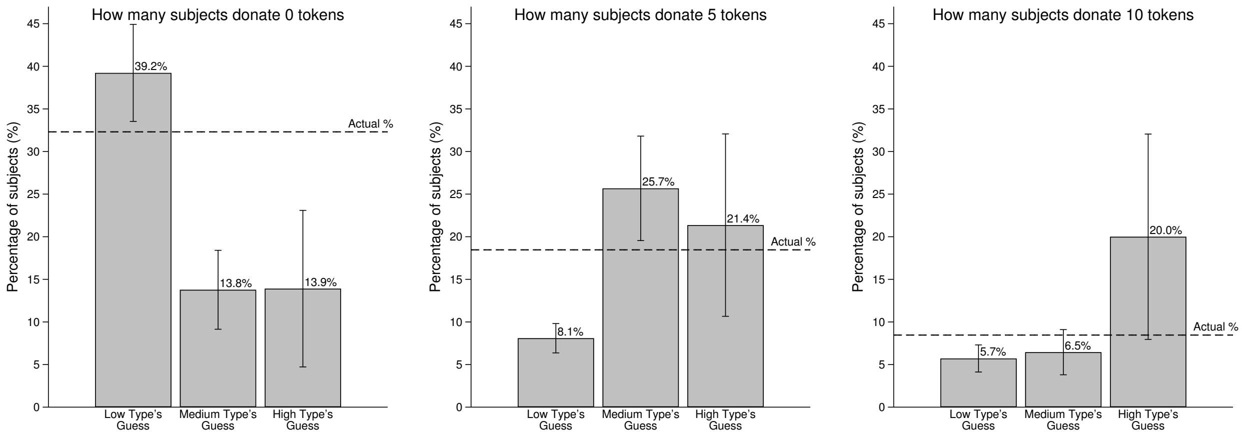
		Low Type's Guess		Medium Type's Guess		High Type's Guess	
Subject Type	Actual %	(%)	*p*-value	(%)	*p*-value	(%)	*p*-value
Low types	67.69	63.34	0.084	29.74	0.000	23.55	0.000
Medium types	23.85	26.27	0.191	55.42	0.000	44.18	0.000
High types	8.46	10.39	0.086	14.84	0.002	32.27	0.002

Focal points Another interesting aspect of the predicted distributions in Figure 2.2 is that each histogram contains spikes at 0, 5, and 10 tokens, suggesting that all subject types consider 0, 5, and 10 tokens to be some sort of focal points. However, subjects overestimate the spikes that represent their own types. That is, all subjects (correctly) predict that there will be spikes at 0, 5, and 10 tokens, but exhibit a systematic bias when it comes to estimating the *size* of each spike. This is most consistent with the hypothesis that it is easier for the mind to distort beliefs about a distribution's skewness than beliefs about a distribution's shape. Therefore, one of the first lines of attack for any subconscious bias is to distort beliefs about the skewness of a distribution.

Figure 2.5 presents a closer look at each focal point, i.e., 0, 5, and 10 tokens. Panels (a), (b), and (c) show each type's average estimate about the proportion of subjects who will donate 0, 5, and 10 tokens respectively. Panel (a) shows that the actual proportion of subjects who donate 0 tokens is 32%. On average, low types *overestimate* this proportion by 7 percentage points, medium types *underestimate* it by 18.4 percentage points, and high types *underestimate* it by 18.3 percentage points. Similarly, panel (b) shows that low types underestimate the proportion of subjects who donate 5 tokens (by 10.4 percentage points), whereas medium and high types overestimate it (7.2 and 2.9 percentage points, respectively). Lastly, panel (c) shows that low and medium types both underestimate the proportion of subjects who donate 10 tokens whereas high types overestimate this proportion. Among the three types of subjects, low types make the best estimates about the proportion of subjects who donate 0 tokens, high types make the best estimates about the proportion who donate 5 tokens, and medium types make the best estimates about the proportion who donate 10 tokens.

Figure 2.5: Predictions made by Low, Medium, High types about 0, 5, 10 tokens

(a) 0 tokens (b) 5 tokens (c) 10 tokens

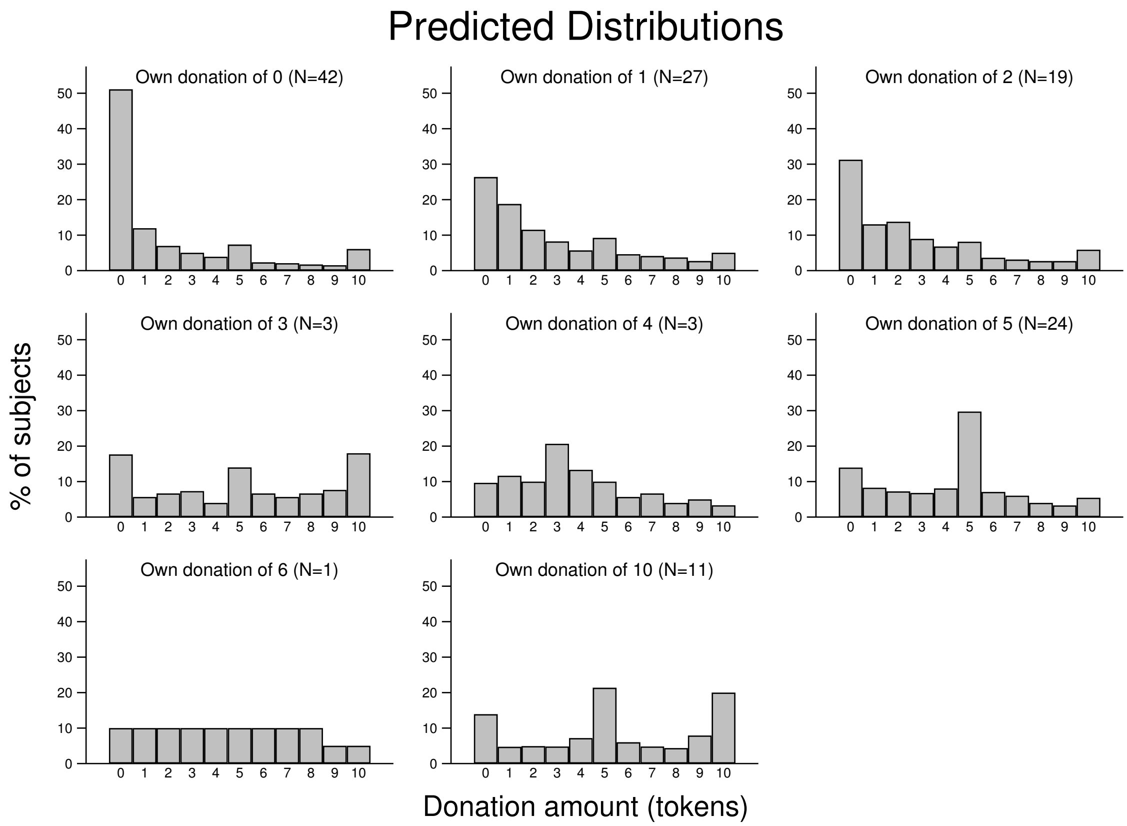


Notes. The bar graphs in panels (a), (b), and (c) represent each type's average prediction about the percentage of subjects who donate 0, 5, and 10 tokens respectively. Error bars represent 95% confidence intervals. In each panel, the error bars are smallest for low types and largest for high types. This is because the vast majority of subjects (68% to be more exact) are low types while only about 8% of subjects are high types. This is especially important to consider in panel (c), where relative bar heights indicate that high types' estimates are considerably overestimated; however, errors bars indicate that this overestimation is not statistically significant (because the actual proportion lies within the error bar).

Robustness Check for Assignment into Types The fact that multiple actions are grouped into the same type (recall that low, medium, and high types are respectively defined as those who donate 0-2, 3-7, and 8-10 tokens) raises a concern about this being an arbitrary assignment into types. However, a closer look at subjects within a single type suggests that this subjects within each type are similar after all – i.e., this assignment is not arbitrary.

Figure 2.6 shows the average predictions made by those who donated 0 tokens, those who donated 1 token, and so on. Using this method to categorize subjects can result in up to 11 different subject groups. For the sake of distinction, I call each subject's own donation amount her *sub-type*; so there are three different types of subjects and up to 11 different sub-types of subjects. Figure 2.6 also shows that this method ends up further separating the three types into 8 sub-types. High types have no variation at all (because nobody chose a donation amount of 8 or 9 tokens), so there is only one sub-type within high types. Medium types have very little variation because only three subjects donated 3 tokens, three

Figure 2.6: Predictions made by each sub-type



donated 4 tokens, and only one donated 6 tokens. However, with 42 sub-type-0 subjects, 27 sub-type-1 subjects, and 19 sub-type-2 subjects, at least low types contain enough data to allow checking for variation within these three sub-types. The first three charts in Figure 2.6 show that each of these sub-types predict that the mode of the distribution will be 0 tokens, and there will be spikes (albeit small) at 5 and 10 tokens, suggesting that each of these sub-types holds similar beliefs about the distribution. The only notable difference is that sub-type-0 subjects significantly overestimate the proportion of subjects who will donate 0 tokens in particular.

2.5 Conclusion

The main finding of this paper is that altruistic subjects systematically overestimate the proportion who are altruistic and also the average altruism level. This overestimation is consistent with a transformation that systematically skews a distribution according to an individual's own altruism level. The foremost policy implication is that people's beliefs about others, even collectively, should not be considered as correct. Second, since people systematically bias their responses depending on their own preferences, it would make sense to always include people's own characteristics and preferences when collecting their opinions about others. This would allow surveyors to correct for this bias.

In the experiment presented in this paper, subjects did not have much information that could have helped them guess how altruistic other people are and how many people of each altruism level there are. The only knew their own altruism levels and perhaps of people they know well enough. It appears that these are exactly the pieces of information they used. In other words, subjects formed beliefs about other people's altruism by extrapolating from the sample of people whose altruism levels they already knew (and this sample contains at least one person: themselves). If a subject was aware of the fact that, for example, that her circle of friends is more altruistic than average, and she also knew that people's social circles create biases in their beliefs, then she could correct for this bias. For example, if an individual is aware of the fact that her network consists mostly of liberal-minded individuals, then she can use this result to realize that her belief about the population distribution of liberals is likely to be biased. More generally, this would help her realize which of her beliefs are more likely to be inaccurate, and ultimately help her make better decisions.

The systematic bias identified in this paper also has the potential to contribute to advancements in economic theory. A standard assumption in economics, and particularly in theoretical models that involve random variables and probability distributions, is that individuals know the distribution function of a random variable but not its realized value. The finding of this paper suggests that this assumption is questionable but, given its systematic nature, also improvable. For example, in auction theory, a realistic framework might be one in which a bidder believes that other bidders value an item similarly to how to she values it, even though in reality everyone's valuation may be independent. As another example, this paper's finding would support a signaling model a receiver receives a signal through an inference function that depends not only on the sender's action but also on the receiver's type.

CHAPTER 3

Risk, Trust, and Altruism in Genetic Data Sharing

Abstract[1]

This study investigates individual attitudes toward privacy risks associated with the sharing of genetic data. How does concern about genetic data privacy compare with other concerns? We conduct behavioral experiments to elicit attitudes with respect to several conditions. First, we consider two distinct scenarios to explore how types of information provided shape behavior. We examine two types of information: 1) genetic data shared with a healthcare provider; and 2) financial data shared with a money manager. In the former case, uncertain benefit is stated in terms of health outcomes, whereas in the latter, uncertain benefit is stated in terms of financial benefit. Both scenarios involve identical decisions and monetary stakes, permitting us to focus on how the framing of data sharing influences attitudes. Second, we design experiments to investigate the motivations behind decisions in terms of altruism and trust in data sharers. Third, we consider whether data recipients protect shared data when protection is costly and benefits data sharers only. Our findings (with 162 subjects) indicate that individuals are more willing to risk a loss to privacy of genetic data (for an anticipated return in health benefits) than they are to risk loss of financial data (for an anticipated return in financial benefits). We further observe that trust has a significant impact on the investment frame, but not on the genetic frame. Finally, we find that 50 to 60 percent of data recipients choose to protect another person's data, with no significant differences between frames.

[1]Chapter 3 was written in collaboration with Wooders, Myrna; Malin, Bradley; and Vorobeychik, Yevgeniy. The dissertation author was the primary author of this material.

3.1 Introduction

Cheap DNA sequencing has enabled broad collection, processing and sharing of genetic information (van Dijk, Jaszczyszyn, Naquin and Thermes, 2018). This information is disclosed by patients to physicians, by research participants to scientists, and by general consumers to direct-to-genetic testing companies (e.g., 23andme.com) for various purposes. Genetic information sharing provides an individual with the opportunity to improve health and contribute to societal endeavors. At the same time, while some people share genetic information openly (Haeusermann et al., 2017) (e.g., posting to websites such as OpenSNP (Greshake, Bayer, Rausch and Reda, 2014) or GEDMatch (Greytak, Kaye, Budowle, Moore and Armentrout, 2018)), many others consider such information to be highly personal and potentially sensitive (Wade, 2018). As a result, many people prefer that their information be managed and used in a manner that preserves their expectation of privacy (Erlich and Narayanan, 2014). Sharing data with another party is not without risks. For instance, a data recipient may fail to protect a data sharer's privacy through weak anonymization practices or misuse the data due to lack of access control and institutional oversight.

Although various investigations have examined the extent to which people are concerned about the privacy and security of their genetic data, there are several major limitations to existing studies. First, they typically elicit information about privacy concerns through surveys (Botkin et al., 2012; Clayton et al., 2018; Condit et al., 2016; Duquette et al., 2012; Edwards et al., 2012; Kaufman et al., 2009; Lemke et al., 2010; Rauscher et al., 2015; Sanderson et al., 2017), despite the fact that survey instruments often fall prey to the *privacy paradox*, whereby reported attitudes about privacy end up being inconsistent with measures people actually take to protect their data (Barth and De Jong, 2017). Second, existing research often neglects the fact that data sharing requires an interaction between two parties (e.g., a patient and a physician), which can influence behavior.

A natural way to address these limitations is to adopt an economic perspective, whereby a genetic data sharer's decision is modeled as a risky investment. In this representation, sharing genetic data may result in long-term benefit (e.g., treatment for an undiagnosed condition (Bastarache et al., 2018)), but may expose an individual to privacy risks. Similarly, a data recipient's decision about how and when to protect genetic data has a salient economic aspect, as data protection can be costly and data recipients often have a limited budget. In both cases, a decision has ramifications for the other party.

In this work, we investigate whether decisions about sharing genetic data are different from decisions about sharing financial data, and if this information can be used to better understand how individuals value privacy of their genetic data relative to the privacy of their financial data. Better understanding is crucial to determining the conditions under which people share genetic data, and how such sharing can be encouraged.

We conduct this investigation of the implications of economic decisions involved in genetic data sharing and protection through controlled experiments with human subjects, focusing particularly on the impact of *framing*, which has been shown to affect decisions in a number of other behavioral studies.[2] Specifically, we represent the problem in two ways. First, we design a scenario where a *patient* can undergo a *genetic test*, and thereby share *genetic* data with a *physician*. Second, we design a scenario in which an *investor* can *invest in a risky asset*, thereby sharing *financial* data with a *money manager* purely

[2]Variable framing has been extensively used in economics and psychology to investigate a variety of issues. See, for example: Andreoni (1995) on positive versus negative framing; Van Dijk and Wilke (2000) on give versus take frames; Bacharach and Bernasconi (1997) on framing effects and focal points; Rege and Telle (2004) on the impact of social approval on cooperation; Dufwenberg, Gächter and Hennig-Schmidt (2011) on the impact of frames on beliefs. To our knowledge, ours is one of the first or the first to use framing effects in the study of health issues and in particular, of genetic privacy.

for financial gain. We refer to these scenarios as the *genetic frame* and the *investment frame*, respectively. Crucially, actual monetary gain and loss potentials are identical in both frames.

We expand on this experimental design through several variations that allow us to isolate the effects on decisions of economic and altruistic motivations, as well as trust and reciprocity. To do so, at the beginning of each experiment, subjects, acting as patients or investors, are provided with a fixed amount of money as an endowment. Each subject makes a decision on whether to share data, that is, whether or not to undergo a genetic test (in the genetic frame) or whether or not to make an investment (in the investment frame). If a subject does not share data, she keeps her endowment as her final payout. If she does share data, the subject faces two uncertain possibilities, one positive and the other negative. First, with a small probability, she may receive a substantial sum of money (i.e., a return on investment), representing treatment in the genetic frame or financial gain in the investment frame. Second, with a larger probability, the subject loses *all* of her endowment, representing the risk associated with a potential privacy compromise of shared data as a result of her decision.

The design of the aforementioned experimental setting effectively removes the role of the data recipient from the equation, thus casting the problem purely as a risky decision. While the presence of the data recipient does not change the economic nature of this choice, one may expect that people will not act out of pure self-interest. If data sharing can benefit another party, it is possible that this will serve as an incentive to share. This motivates an *altruism setting*, in which data sharing also benefits a data recipient.

Finally, we establish a *trust setting* in which data sharing benefits both the data sharer and the data recipient but, in addition, the data recipient can, at a cost, reduce the data sharer's risk of privacy loss. This enables us to measure the extent to which trust in the recipient's reciprocity motivates a greater degree of data sharing. Thus, all data sharers (i.e., patients or investors) decide whether or not to share data in three different settings, namely base, altruism, and trust.

Turning our attention to data recipients, we also study their motivations to reduce the privacy risk faced by data sharers. Data recipients decide whether or not to bear a monetary cost that would reduce the data sharer's risk. They make this decision in two settings, a *reciprocity setting* and a *recipient-altruism* setting. In the reciprocity setting, the data sharer's decision to share data has a direct financial benefit on the recipient, independent of the latter's decision to reduce data privacy exposure risk. In the recipient-altruism setting, the data recipient starts with data and can spend money to reduce the risk of loss to the data sharer, to whom the data belongs according to the experiment narrative. Comparing these two settings allows us to isolate the effects of altruism and reciprocity on motivating the data recipient to reduce the sharer's risk.

Our findings show that data sharers are significantly more likely to tolerate risk when sharing genetic data than when sharing financial data. We further find that trust and reciprocity appear to encourage sharing more than altruism does.

3.2 Methods

We study data sharing through human subject experiments involving monetary stakes. Our experiments involve two classes of settings: a decision under uncertainty involving a single subject (e.g., the base setting), and an interaction between two subjects (e.g., the trust & reciprocity setting). In the latter class of settings, we first divide subjects into data sharer and data recipient pairs, randomly and anonymously. Data sharers can share their data and data recipients can expend resources to protect that data and reduce data sharers' risks.

Each {data sharer, data recipient} pair is randomly assigned to either the genetic frame or the investment frame. In the genetic frame, data sharers are assigned to a role of a patient and data recipients are assigned to the role of a physician. In the investment frame, data sharers are assigned to a role of an investor and data recipients are assigned to a role of a money manager. Thus, we pair patients with physicians in the genetic frame, and investors with money managers in the investment frame. All subjects maintain their randomly assigned roles in all experimental settings.

The underlying decision problem - in terms of monetary payoffs, risks, and choices - is the same in both frames. The only difference is what the numbers represent. In the genetic frame, the decision problem simulates a situation where a patient needs to decide whether or not to undergo a genetic test. In the investment frame, the decision problem simulates a situation where an investor needs to decide whether or not to make a risky investment. The risk in either case stems from potential privacy breaches associated with data sharing by the patient or the investor.

Within each frame, we conduct four versions, i.e., settings, of the experiment that allow us to deconstruct the various motivations that individuals may have to share their data, similar to a triadic design approach pioneered by Cox (Cox, 2004; Cox et al., 2008, 2015; Di Bartolomeo and Papa, 2016) for decomposing motivations. We hypothesize that data sharers have three reasons to share their data: (i) potential health/financial benefits resulting from genetic testing/investment; (ii) benefits to data recipients, such as money managers or healthcare institutions/physicians who are engaged in research; and (iii) trust in the ultimate data custodian to protect their data from being stolen or misused. The four settings of our experiment are called : 1) base, 2) altruism, 3) trust and reciprocity, and 4) recipient-altruism. Data sharers participate in the first three settings, i.e., 1) base, 2) altruism, and 3) trust and reciprocity. Data recipients participate in the last three settings,

Table 3.1: Summary of differences among settings

Setting name	Sharers' motivation	Recipients' motivation
Base	1. Personal benefit	N/A
Altruism	1. Personal benefit 2. Altruism	N/A
Trust & Reciprocity	1. Personal benefit 2. Altruism 3. Trust	1. Altruism 2. Reciprocity
Recipient- Altruism	1. Personal benefit 2. Altruism 3. Trust	1. Altruism

In the base setting, data recipients do not even exist. In the *altruism* setting, data recipients participate only as passive recipients of their data sharer's altruism. Thus, data recipients do not have any actions in the *base* and *altruism* settings. The *trust and reciprocity* setting is the full version of the game with the highest number of confounding motivations. In the *recipient-altruism* setting, a data recipient is given money from the experimenter, regardless of the data sharer's action. Note that in both the *trust and reciprocity* setting and the *recipient-altruism* setting, a data recipient has an action only if the data sharer chooses to share data.

i.e., 2) altruism, 3) trust and reciprocity, and 4) recipient-altruism. Below, we describe each setting in detail, while Table 3.1 provides a summary of the settings and the associated motivations that are tested.

3.2.1 Base setting

In this setting, a data sharer is endowed with \$6 and must choose whether or not to share her data, which costs \$2 out of the \$6 endowment. If she chooses to share data, she can win \$60 with a 5% probability, but also lose her remaining \$4 with a 25% probability. Note that there is no mention of a data recipient in this setting, so other-regarding preferences cannot affect a data sharer's decision.

Patients are told that the \$60 represents potential health benefits from genetic testing and the loss of \$4 represents the loss of privacy of genetic data. Investors are told that

$60 represents a financial gain from the investment while $4 represents loss of privacy of financial data. If a person's genetic data gets hacked and/or misused, she could potentially lose her health insurance, and in the worst case, her job and family ties, depending on the information revealed through the genetic test. To capture this extreme case, we represent the cost of exposure of sensitive data as the loss of all wealth. In order to obtain meaningful results, we deliberately give participants a higher risk of their data being exposed than what it is in the real world. The fact that we do not use real-world numbers does not create a problem because our goal is not to estimate the proportion of population that is willing to perform a genetic test given awareness of actual risks. Rather, our objective is to compare the impact of privacy concerns when they pertain to genetic versus financial data. Financial data, in turn, are natural baselines for comparison, as this is the most common frame used in behavioral economic studies.

3.2.2 Altruism setting

Typically genetic testing and investment settings are not single-agent decisions, but involve another party such as physicians or money managers. In such encounters, a number of factors may influence people's decisions in addition to egocentric motivations, including altruism, trust, and reciprocity. In order to tease apart these factors, we introduce an *altruism setting*. In this setting, a data sharer's decision problem is the same as in the base setting with one exception. Specifically, if a data sharer chooses to share data and give up $2, a data recipient (who is said to be a physician or a money manager, depending on the frame) receives $4. This $4 represents either the physician's benefit from advancing his research or the money manager's benefit. Note that while a data recipient receives money, he does not have a decision to make. This setting includes two motivations for a data sharer to share data: (i) her potential benefit (which is also present in the base setting) and (ii) her desire to provide benefit to a data recipient, or altruism. We hypothesize that if altruism is a motivation for sharing data, then a data sharer will be more likely to share data in the altruism setting than in the base setting.

3.2.3 Trust and Reciprocity setting

For data sharers, this setting is called the trust setting and for data recipients, it is called the reciprocity setting. In the experiment, these separate settings are implemented as a single, interactive game in which both parties make decisions that affect their own as well as each other's payoffs. This setting resembles the extensively-studied trust game (Berg et al., 1995) in which a player, say Player 1, is endowed with some money and chooses whether to send a portion of it to another player, say Player 2. Player 2 receives three times the amount sent by Player 1, and then has the opportunity to send some of his money back to Player 1. For example, if Player 1 sends \$5 from her \$10 endowment to Player 2, then Player 2 receives \$20. Player 2 can then send some of his \$20 back to Player 1. There are robust findings in this game: (a) Player 1 often sends a non-zero amount to Player 2, and (b) Player 2 commonly reciprocates and sends a portion of her funds back to Player 1.

The trust and reciprocity setting adds one more feature to the altruism setting. Specifically, a data recipient may now spend some of his \$4 to reduce the data sharer's risk of loss. Specifically, a data recipient has three options: spend nothing (i.e., \$0), spend half of his earnings (i.e., \$2), or spend all of his earnings (i.e., \$4). If a data recipient spends \$0, then the data sharer faces a 25% probability of losing her remaining wealth of \$4 (which is the same probability as in the other settings); if a data recipient spends \$2, then the data sharer faces a 15% probability of losing her \$4 wealth; and lastly, if a data recipient spends all of his \$4, the data sharer faces no risk (i.e., a 0% probability) of losing her \$4 wealth.

In the base and altruism settings, a data sharer's expected payoff is \$6 from sharing data and \$6 from not sharing data (thus, a risk neutral person would be ambivalent between the two choices). Now a data sharer's expected payoff depends on a data recipient's decision. If a data sharer believes that the data recipient will spend his entire \$4, then her expected payoff is \$7. If she believes that a data recipient will spend \$2, then her expected payoff is \$6.40. Finally, if she does not expect the data recipient to spend anything, then her expected payoff is \$6, the same as in the other settings. Therefore, a data sharer now has three motivations to share data: (i) her own benefit (this motivation is present in all three settings), (ii) her altruism, or desire to benefit a data recipient (this is also present in the altruism setting but is not present in the base setting), and (iii) whether she trusts her data recipient to protect her wealth (which is not present in either of the previous two settings). If a data sharer is more likely to share data in the trust setting than in the altruism setting, then trust is what accounts for the difference.

Let us now turn shift our focus to the motivations of a data recipient. If a data recipient spends a positive amount to reduce a data sharer's probability of loss, then there are two possible factors that could be motivating this choice: (i) altruism: he wants to protect the data sharer from a loss; and (ii) reciprocity: he feels beholden from knowing that his income is a result of the data sharer's decision, and thus reciprocates the favor by spending some money to help the data sharer. To disentangle these two factors, we construct a *recipient-altruism* setting that is described below.

3.2.4 Recipient-altruism setting

The purpose of this setting is to deconstruct the different underlying motivations of data recipients (i.e., physicians and money managers) to spend money to protect a data sharer's data and, consequently, her endowment. The only difference between this recipient-altruism setting and the reciprocity setting is that the data recipient's earning of

$4 is not a consequence of the data sharer's decision to share data. Instead, data recipients are now endowed with $4 regardless of data sharers' choice. Each data recipient is told that a data sharer's personal data, and hence wealth, is at risk and data recipients need to decide whether to spend some of their $4 to protect it. Thus, a data recipient only has one motivation to spend money to protect a data sharer's wealth: altruism. The increase in the tendency of data recipients to protect data sharers' data in the reciprocity setting compared to this altruism setting would then be motivated by reciprocity. Note that in both reciprocity and recipient-altruism settings, data recipients have no purely economic motivation to spend money to protect a data sharer's endowment.

3.2.5 Implementation

We conducted several experiment sessions at Vanderbilt University in March 2018 and October 2018, recruiting a total of 162 undergraduate students to participate in the experiment. All sessions took place on weekdays in a computer lab that had 30 computer stations, and subjects submitted all responses using computers. Subjects were paid in cash for all settings at the end of the experiment session. While payments were being calculated and put into envelopes, subjects were afforded the opportunity to complete a demographic questionnaire.

We conducted a total of 8 sessions, 4 for the genetic frame and 4 for the investment frame, with 78 subjects in the genetic frame (39 patient and physician pairs) and 84 subjects in the investment frame (42 investor and money manager pairs). Each data sharer participated in the first three settings, i.e., 1) base, 2) altruism, and 3) trust and reciprocity, while each data recipient participated in the last three settings, i.e., 2) altruism, 3) trust and reciprocity, and 4) recipient-altruism. Subjects were not informed about their outcomes, and therefore also earnings, from each setting until the end of the experiment session.

At the beginning of the experiment, subjects were informed that they would be participating in multiple decision-making tasks, but were not informed about the exact number or nature of the tasks. They only learned about each task at the beginning of each setting. This helped us ensure that each subject's first decision remained pure from any undesirable behavioral effects, such as priming or portfolio effects.

To ensure that responses were free from such effects, we gave these settings to subjects in various sequences (i.e., some subjects were given the base setting first, some the altruism setting, and others the trust setting). Any undesirable effects would have contaminated the results in the second (or later) setting in which a subject participated. In the event of such contamination, we would have discarded subjects' responses in the second and third settings, keeping only the responses recorded from the first setting, which would be guaranteed to be free from such effects.

We found that the order in which a subject received a setting did not make a difference in the proportion of subjects who shared their data. For example, the proportion of patients who chose to get a genetic test in, say, the base setting remained the same regardless of whether they participated in the base setting as their first, second, or third setting. Therefore, we conclude that subjects made their decision in each setting independently of their decision in another setting. This implies that our results are free from any undesirable effects (e.g., priming or portfolio effects) and we can keep responses from all settings.

To ensure anonymity while paying everyone their correct earnings, we provided each experiment participant a unique numerical code, which was used as their pseudonym throughout the experiment. All subjects were guaranteed minimum earnings of $5 for participating, but could earn substantially more, depending on both their choices and the choices of the participants with whom they were paired. Average subject earnings were approximately $20, and average time spent was approximately 45 minutes (including

Table 3.2: Summary Statistics

Variables	Investment	Genetic	Difference
Age	19.71	20.03	-0.326
Victim of ID theft	0.076	0.029	0.033
Real-life Choice	0.519	0.116	0.409***
% Female	0.506	0.406	0.133
% White	0.544	0.580	-0.003
% Hispanic	0.101	0.130	-0.030
Religiousness (0-3)	1.329	1.304	-0.060

*p< 0.10, **p< 0.05, ***p< 0.01

This table reports means of demographic characteristics, separately for the genetic and investment groups. The last column shows the difference in means. A statistically significant difference indicates that subjects in the genetic and investment groups are different in terms of that characteristic. The variable *Real-life Choice* is different across the two groups, but that is not a concern because this variable represents an entirely different question for each group. For the investment group, it represents whether subjects have made personal investments in real life while for the genetic frame it represents whether subjects have ever undergone genetic testing in real life. Thus, overall the assignment into groups was perfectly randomized. The p-values reported in this table are based on two-tailed t-test results (we also conducted χ^2 and Mann-Whitney tests, both of which provide the same conclusions).

administrative tasks such as introduction, questions, and payments).

Table 3.2 provides summary statistics for the demographic characteristics of subjects. The fact that none of the subjects' characteristics are different across the two frames indicates that assignment to frames was sufficiently random. . There is no statistically significant difference between the two groups in terms of age, gender, race, religiosity, or being a victim of identity theft. Only the variable real-life investment is statistically different across the two groups, but this is because this question is actually different for each group. For the genetic group, the question asks subjects whether they have undergone genetic testing in real life, whereas for the investment group, the question asks whether they have made any investments in real life.

3.2.6 Online Replication

We also replicate the laboratory experiment using the online platform Amazon Mechanical Turk. The results of the online experiment are highly similar to the in-person laboratory experiments, but we chose not to pool data from the two experiments because the subject populations and experimental environments are different. Rather, the online settings demonstrate robustness of our results to different settings and subject populations. We refer the reader to the Supplementary Information for details of this experiment.

3.3 Results

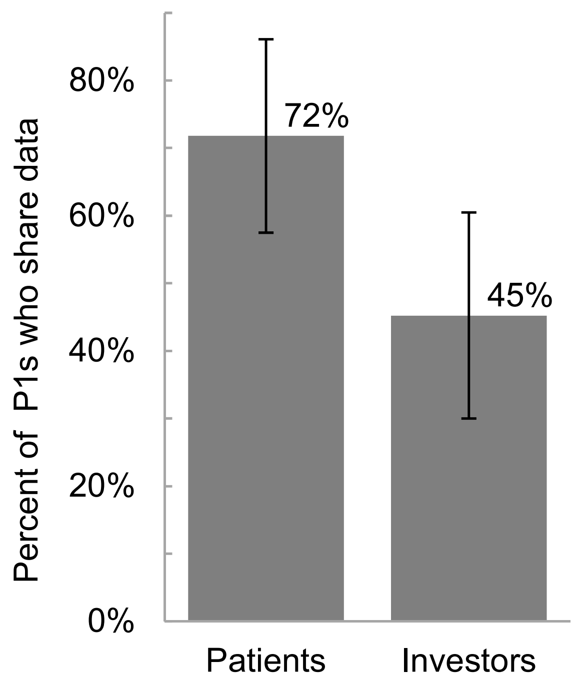
We first study the decision of an individual with respect to whether or not to share data with another party. In the genetic frame, a subject who is assigned the role of a patient needs to decide whether to take a genetic test and share her genetic data with a physician. In the investment frame, a subject assigned the role of an investor needs to decide whether to invest in an asset and share financial data with a money manager. In both frames, i.e., regardless of the narrative presented to the subject, experiment outcomes are monetary, determined as follows. Subjects are endowed with \$6 at the beginning of an experiment. If they choose not to share data (either by getting a genetic test or by making an investment, depending on the frame), they receive the full endowment at the end of the experiment. If instead they choose to share their data, they give up \$2 from their endowment to face an uncertain outcome that is defined as follows. On the one hand, data sharers have a 5% probability of receiving \$60 (representing either a health or financial benefit, depending on the frame), for a total final payout of \$64 at the end of the experiment. On the other hand, they have a 25% probability of losing their remaining endowment, representing the consequence of the loss of privacy of shared data, leaving them with nothing at the end of the experiment. Note that from a purely decision-theoretic viewpoint, the two decisions have identical expected earnings of $6; consequently, a risk-neutral individual would be indifferent between either sharing her data or not sharing her data, whereas a risk averse person would not share data.

More People Choose to Share Data in the Genetic Frame In the base setting, we remove the data recipient from the equation. Instead, this setting involves only a data sharer who decides whether or not to put her data (genetic or financial) at risk, without explicit reference to a data recipient. Our main result in the base setting, shown in Figure 3.1, is that there is a substantial framing effect. Specifically, 72% of subjects in the genetic frame choose to share data, strongly suggesting that the uncertain outcome is seen as favorable. By contrast, only 45% of subjects in the investment frame choose to share data, suggesting that the uncertain outcome is seen as somewhat unfavorable.

In addition to comparing the means using a t-test, we compare the two frames through a regression analysis, where we control for various demographic characteristics such as age, race, gender, degree of religious beliefs, and whether subjects have ever been victims of data theft. Table 3.3 presents these regression results. The first column of the table shows results produced by an OLS regression and the second column shows results produced by a Probit regression. In both regressions, the dependent variable is the percentage of subjects who choose to share data (framed as choosing either to undergo a genetic test or to make an investment).

The key independent variable is *Genetic Frame*, which is a binary variable indicating whether a subject is given the genetic or investment frame. In the OLS regression, the variable Genetic Frame has a coefficient of 0.254, which suggests that subjects in the genetic

Figure 3.1: Framing Effects



This bar chart shows framing effects in the base setting (in which patients/investors are asked to make decisions without any mention of physicians/money managers). We find that 72% of subjects in the genetic frame share their data, compared to only 45% of subjects in the investment frame. This difference is statistically significant (p= 0.015 for a two-tailed t-test). Error bars represent 95% confidence interval.

Table 3.3: Regression results

Variables	OLS	Probit
Genetic Frame	0.254**	0.675**
Failed control questions	0.0243	0.0768
Age	-0.0389	-0.111
Victim of ID theft	0.174	0.493
% Female	0.0522	0.123
% White	0.169	0.466
% Hispanic	-0.216	-0.584
Religiousness (0-3)	-0.0101	-0.0302
Constant	1.089	1.716

* p<0.10, ** p<0.05, *** p<0.01

This table shows results of OLS and Probit regressions. In both regressions, the dependent variable is the percentage of subjects who share data. The variable *Genetic Frame* is statistically significant. This is a binary variable indicating whether a subject is assigned to the genetic frame or investment frame. The results show that patients (i.e., subjects in the genetic frame) are about 25 percentage points more likely to get a genetic test than investors are likely to make an investment. Other variables do not significantly affect the decision to share data.

frame (i.e., patients) are about 25.4 percentage points more likely to get a genetic test than investors are likely to make an investment. In the Probit regression, a marginal analysis shows that patients choose to undergo a genetic test with a probability of 0.686 while investors choose to make an investment with a probability of 0.425, which is a difference of 26.1 percentage points.[3] Thus, the results of Figure 3.1 are broadly consistent with the OLS and Probit results. The regression results also show that none of the demographic variables has a significant influence on the decision to share data, suggesting that people's willingness to put data (whether genetic or financial) at risk is similar across different age groups, gender groups, and racial groups.

Trust Increases Data Sharing Next, we investigate whether the involvement of a data recipient influences the decision to get a genetic test or make an investment. The presence of a data recipient/custodian (who may be an individual or an organization) introduces

[3] The probit coefficient shown in the table, 0.675, is the difference between the z-scores of the genetic frame and of the investment frame. We use this coefficient to perform a marginal analysis whose interpretation is more meaningful for our purposes.

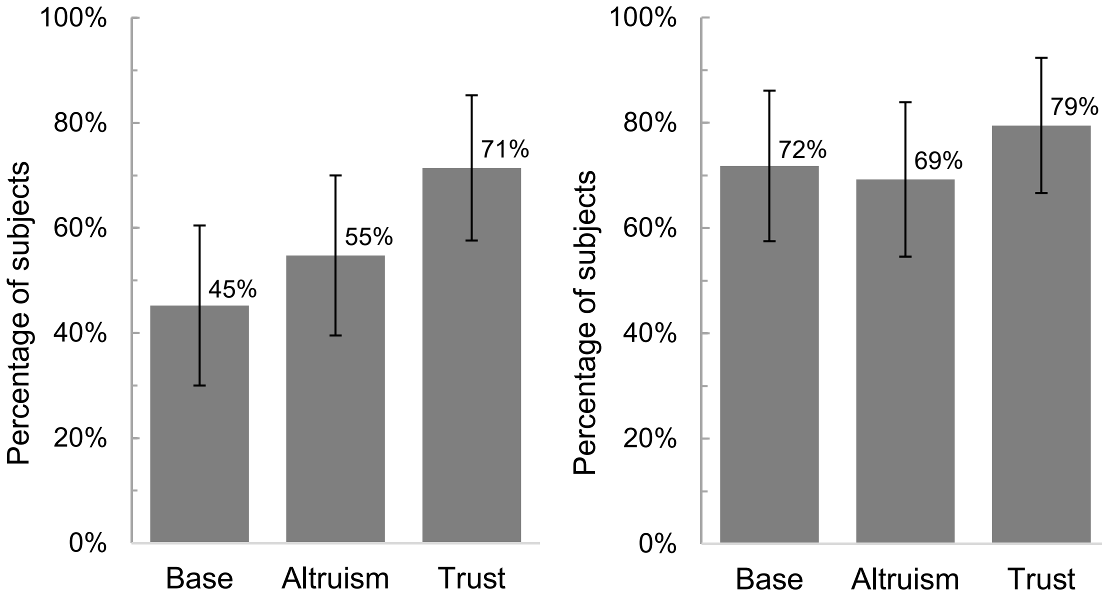
considerations into the decision about whether data is shared. These may include other-regarding preferences as well as a degree of trust in the data custodian's willingness to protect data. Thus, the following two settings consider the impact of these considerations, in terms of *altruism* and *trust*, on a data sharer's decision as to share data.

In the altruism setting, data sharers (i.e., patients and investors) are randomly and anonymously paired with data recipients (i.e., physicians and money managers). By choosing to share data, the data sharer *also benefits the data recipient*, who receives $4 if the data sharer decides to share data. This gain represents investment gain to the money manager in the investment frame and helps the physician with biomedical research in the genetic frame.

The trust setting also features a data recipient (i.e., a physician or a money manager) who receives $4 if the data recipient shares data. However, now the recipient is viewed as a data custodian who may choose to protect the sharer's data. In particular, the data recipient now has three options: (i) spend $0 to protect the data, which keeps the data sharer's risk of loss at 25%; (ii) spend $2 to protect the data, which reduces the data sharer's risk of loss to 15%; and (iii) spend all $4, reducing the sharer's risk to 0%. Thus, if the data sharer trusts the recipient to spend at least some of the money received for better data protection (concretely, reducing the risk of loss to the sharer), the data sharing option becomes far more appealing.

Figure 3.2 compares the results of all three settings in which a data sharer chooses to share data. While there is no statistically significant difference between any pair of settings in the genetic frame, there are statistically significant differences between settings in the investment frame. Specifically, in the trust setting, a significantly greater number of investors choose to share data, compared to those in the altruism and base settings. This implies that investors are more likely to invest when their money manager has the ability

Figure 3.2: Choices made by Data Sharers in each setting



This chart shows the proportion of subjects who share data in each setting and each frame. In the genetic frame (left panel), there is no statistically significant difference between the heights of any two bars. In the investment frame (right panel), two pairwise differences are statistically significant: (i) between the trust and altruism settings (p=0.018) and (ii) between the trust and base settings (p=0.003). That is, when money managers are able to protect investors' data, investors are much more likely to invest. This suggests that investors trust money managers to protect their financial data and that trust is an important factor in their decision to invest. By contrast, when physicians are able to protect patients' data, it does not increase patients' likelihood to get a genetic test. This suggests that trust is not an important factor for patients in their decisions to undergo genetic testing. Rather, personal benefit is the only important motivator for getting a genetic test.

to reduce their risk, indicating that investors trust their money managers to have their best interests at heart.

Nevertheless, altruism increases the frequency of data sharing in the investment frame by 10 percentage points (from 45% to 55%). This is in stark contrast to the genetic frame, where altruism decreases the frequency of data sharing by 3 percentage points (from 72% to 69%). Indeed, our results in the genetic frame contrast with typical results in related *trust games* (Berg et al., 1995), where a robust finding is that altruism accounts for a large portion of investment (Cox, 2004). Our results suggest a potential displacement effect (Gneezy and Rustichini, 2000) because fewer subjects share data in the altruism setting than in the base setting, even though their purely economic motivations are identical.

Approximately 50-60% of Data Recipients Choose to Protect Data One of the unique features of our study design is the interaction between a patient and a physician in the genetic frame and an investor and a money manager in the investment frame. Data recipients (physicians and money managers) may spend some of the money they receive to reduce the data sharer's risk of loss. This is modeled in our experiment as a reduction of risk of loss for the data sharer from 25% to either 15% or 0%. Based solely on economic self-interest, data recipients would not spend any of the money they receive on data protection. Nevertheless, they do.

Two additional factors, however, can motivate data recipients to protect the sharer's data: altruism and reciprocity. As these two motivations for a non-zero contribution by the data custodian cannot be directly disentangled in the trust setting, we introduce a *recipient-altruism* setting. In this setting, the data recipient receives $4 from the experimenter rather than the data sharer, and may still use it to reduce the data sharer's risk of loss. Thus, any money spent by the data recipient in the altruism setting is motivated solely by altruism.

The previously termed trust setting then becomes the *reciprocity* setting when viewed from the perspective of the data recipient, as now any *additional* spending beyond what we observe in the altruism setting must be due to reciprocity. The distinction between altruism and reciprocity is significant in practice as well: data custodians are often different from people who directly interact with the individuals sharing data (particularly as individuals in charge of data protection may change over time), and the nature of their decisions is best captured by the recipient-altruism setting.

Since spending the full $4 to protect the sharer's data is exceedingly rare, we pool all non-zero contributions by the data recipient into a single *decision to protect*, with the complement corresponding to spending nothing for the sharer's protection. Figure

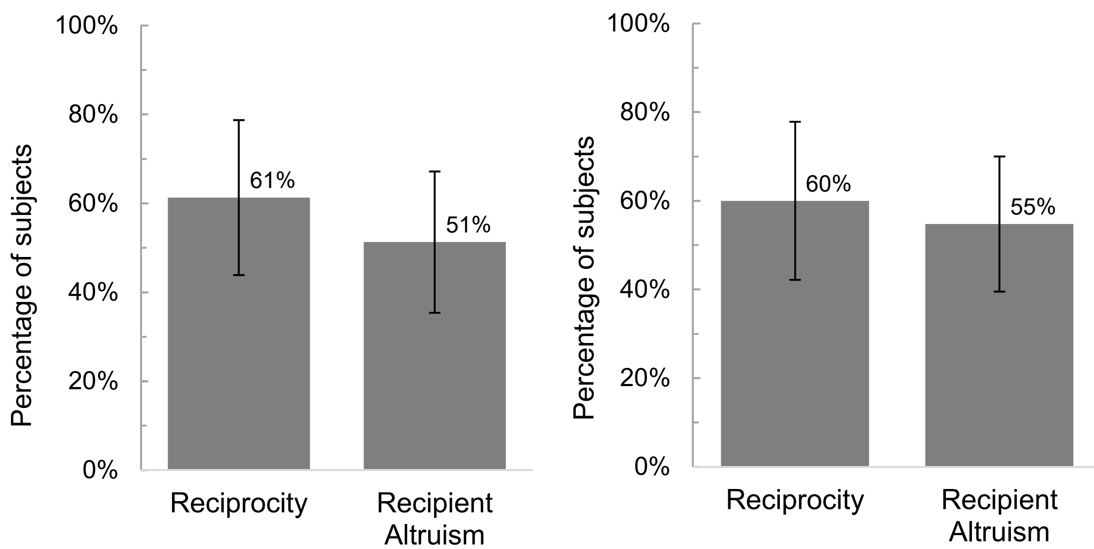
3.3 presents the fraction of physicians and money managers who decide to spend some of their received funds on data protection, which benefits the data sharer. It can be seen that differences between the two frames are not dramatic, although the stronger result is encountered in the genetic frame. In the genetic frame, only 51% of data recipients choose to spend *anything* into reducing the risk to the data sharer when the sole motivation is altruism. In the investment frame, 55% of data recipients choose to spend anything to reduce the sharer's risk. The difference between these two numbers is not statistically significant. Reciprocity boosts the data recipient's spending from 51% to 61% in the genetic frame, and from 55% to 60% in the investment frame. The former comparison is marginally significant (one-tailed test; $p = 0.0516$). Nevertheless, even reciprocity is insufficient motivation because over one-third of data recipients still do not spend money to protect the sharer's data.

3.4 Discussion

There has been much public discussion about tension between data sharing and privacy in general (Acquisti et al., 2015), as well as about how privacy affects the sharing of genetic data in particular (Sanderson et al., 2017). On the one hand, the ability to share fine-grained data is crucial to both research and policy (of Directors, 2017). On the other, privacy concerns must necessarily put a check on what, how, and with whom data is shared (Joly et al., 2012). An important aspect of this discussion is the perceived risk to privacy of shared data, and how perceptions of risk affect trust. Perceived degrees of risk influence individual decisions to opt in or opt out of participating in clinical studies or research programs that involve shared datasets (Hull et al., 2008).

In this paper, we use methods from behavioral economics to understand how people make decisions about sharing data as well as the extent to which data custodians are

Figure 3.3: Choices made by Data Recipients in each setting



Recall that a data recipient assumes the role of a physician in the genetic frame and the role of a money manager in the investment frame. When reciprocity is a motivator, it increases the proportion of physicians who spend some money to protect their patients' data by 10 percentage points (61% - 51%) and the proportion of money managers who spend money to protect their clients' financial data by 5 percentage points (60% - 55%). In either frame, the difference is not statistically significant in a two-tailed comparison. However, a one-tailed t-test results in a marginally statistically significant difference in the genetic frame ($p = 0.05$). This suggests that reciprocity motivates physicians to spend some money to protect their patients' genetic data, but it does not motivate money managers to spend money to protect their clients' financial data.

motivated to protect it. Two especially important factors we explore are: 1) the effect of framing a decision in terms of risky investment; and 2) the effect of viewing data sharing explicitly as an encounter between two parties, a data sharer and a data recipient/custodian.

Our investigation yields several notable findings. First, despite identical monetary stakes, far more people choose to share their data when the decision is narrated as involving genetic testing, compared to when it is cast as financial in nature. As the distinction is solely one of framing, our finding suggests that people are more willing to take on risk when sharing genetic data than when sharing financial data. Second, it is noteworthy that both altruism and trust are significantly stronger motivators for sharing financial data than for sharing genetic data. Third, our finding that 50-60% of data custodians bear a cost to reduce a sharer's risk is, on the one hand, remarkable considering that such decisions only benefit data sharers and not the custodians themselves. On the other hand, social considerations suggest that without additional incentives or explicit regulation, shared data may be inadequately protected.

There are several limitations to our investigation that we wish to highlight, which stem from the fact that our experiments were naturally highly stylized, and thus lack several aspects of actual data sharing settings. The first of these is that the economic nature of the experiment means that ultimately the design revolves around monetary payouts, whatever the narrative. In reality, differences between genetic and financial data sharing situations also involve differences in stakes and in the nature of outcomes. Second, data sharing is often an encounter between an individual and an organization (e.g., individuals may opt into or out of data sharing) or among organizations, whereas our experiment focuses either on an individual's decision or on an encounter between two individuals. Third, data sharing encounters are typically not anonymous, whereas our experiment pairs subjects anonymously. Fourth, data sharing is largely implicitly captured through associated risks, whereas the actual decision is presented as making a financial investment or undergoing genetic testing. Despite these limitations, our observations provide insights into the nature of data sharing encounters and into the variations in behavior that, depending on the nature of data being shared, we should expect.

References

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for truth-telling," *Econometrica*, 87, 1115–1153.

Acquisti, A., L. Brandimarte, and G. Loewenstein (2015): "Privacy and human behavior in the age of information," *Science*, 347, 509–514, 10.1126/science.aaa1465.

Ahmed, A. M., and O. Salas (2011): "Implicit influences of Christian religious representations on dictator and prisoner's dilemma game decisions," *The Journal of Socio-Economics*, 40, 242–246.

Akerlof, G. A., and W. T. Dickens (1982): "The economic consequences of cognitive dissonance," *The American economic review*, 72, 307–319.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2006): "Elicitation using multiple price list formats," *Experimental Economics*, 9, 383–405.

Andreoni, J. (1990): "Impure altruism and donations to public goods: A theory of warm-glow giving," *The economic journal*, 100, 464–477.

——— (1995): "Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments," *The Quarterly Journal of Economics*, 110, 1–21.

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

Andreoni, J., J. M. Rao, and H. Trachtman (2017): "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving," *Journal of Political Economy*, 125, 625–653.

Andreoni, J., and A. Sanchez (2019): "FOOLING MYSELF OR FOOLING OBSERVERS? AVOIDING SOCIAL PRESSURES BY MANIPULATING PERCEPTIONS OF DESERVINGNESS OF OTHERS," *Economic Inquiry*.

Andreoni, J., and C. Sprenger (2011): "Uncertainty equivalents: Testing the limits of the independence axiom,"Technical report.

Ariely, D., A. Bracha, and S. Meier (2009): "Doing good or doing well? Image motivation and monetary incentives in behaving prosocially," *American Economic Review*, 99, 544–55.

Bacharach, M., and M. Bernasconi (1997): "The variable frame theory of focal points: An experimental study," *Games and Economic Behavior*, 19, 1–45.

Barth, S., and M. D. De Jong (2017): "The privacy paradox - Investigating discrepancies between expressed privacy concerns and actual online behavior - A systematic literature review," *Telematics and Informatics*, 34, 1038–1058.

Bastarache, L., J. J. Hughey, S. Hebbring, J. Marlo, W. Zhao, W. T. Ho, S. L. Van Driest, T. L. McGregor, J. D. Mosley, Q. S. Wells et al. (2018): "Phenotype risk scores identify patients with unrecognized Mendelian disease patterns," *Science*, 359, 1233–1239.

Battigalli, P., G. Charness, and M. Dufwenberg (2013): "Deception: The role of guilt," *Journal of Economic Behavior & Organization*, 93, 227–232.

Benabou, R., and J. Tirole (2002): "Self-confidence and personal motivation," *The Quarterly Journal of Economics*, 117, 871–915.

Bénabou, R., and J. Tirole (2006): "Incentives and Prosocial Behavior," *The American Economic Review*, 96, 1652–1678, http://www.jstor.org/stable/30034989.

Benoît, J.-P., and J. Dubra (2011): "Apparent overconfidence," *Econometrica*, 79, 1591–1625.

Bentley, J. W., R. J. Bloomfield, S. Davidai, and M. Ferguson (2019): "Identifying Bias Through Post-Report Interactions," *Available at SSRN 2423131*.

Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, reciprocity, and social history," *Games and economic behavior*, 10, 122–142.

Botkin, J. R., E. Rothwell, R. Anderson, L. Stark, A. Goldenberg, M. Lewis, M. Burbank, and B. Wong (2012): "Public attitudes regarding the use of residual newborn screening specimens for research," *Pediatrics*, 129, 231.

Broberg, T., T. Ellingsen, and M. Johannesson (2007): "Is generosity involuntary?" *Economics Letters*, 94, 32–37.

Butler, J. V., P. Giuliano, and L. Guiso (2015): "Trust, values, and false consensus," *International Economic Review*, 56, 889–915.

Camerer, C. (1997): "Rules for experimenting in psychology and economics, and why they differ," in *Understanding Strategic Interaction*: Springer, 313–327.

Cartwright, E., and M. Wooders (2020): "Own experience bias in evaluating the efforts of others," *Journal of Economic Behavior & Organization*, 173, 164–178.

Cason, T. N., T. Sharma, and R. Vadovič (2020): "Correlated beliefs: Predicting outcomes in 2× 2 games," *Games and Economic Behavior*.

Clayton, E. W., C. M. Halverson, N. A. Sathe, and B. A. Malin (2018): "A systematic literature review of individuals' perspectives on privacy and genetic information in the United States," *PloS one*, 13, e0204417.

Condit, C. M., L. Shen, K. L. Edwards, D. J. Bowen, D. M. Korngiebel, and C. O. Johnson (2016): "Participants' role expectations in genetics research and re-consent: Revising the theory and methods of mental models research relating to roles," *Journal of health communication*, 21, 16–24.

Cox, J. C. (2004): "How to identify trust and reciprocity," *Games and economic behavior*, 46, 260–281.

Cox, J. C., K. Sadiraj, and V. Sadiraj (2008): "Implications of trust, fear, and reciprocity for modeling economic behavior," *Experimental Economics*, 11, 1–24.

Cox, J. C., V. Sadiraj, and U. Schmidt (2015): "Paradoxes and mechanisms for choice under risk," *Experimental Economics*, 18, 215–250.

Dana, J., D. M. Cain, and R. M. Dawes (2006): "What you don't know won't hurt me: Costly (but quiet) exit in dictator games," *Organizational Behavior and Human Decision Processes*, 100, 193–201.

Dana, J., R. A. Weber, and J. X. Kuang (2007): "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," *Economic Theory*, 33, 67–80.

Di Bartolomeo, G., and S. Papa (2016): "Trust and reciprocity: extensions and robustness of triadic design," *Experimental Economics*, 19, 100–115.

Di Tella, R., R. Perez-Truglia, A. Babino, and M. Sigman (2015): "Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism," *American Economic Review*, 105, 3416–42.

van Dijk, E. L., Y. Jaszczyszyn, D. Naquin, and C. Thermes (2018): "The third revolution in sequencing technology," *Trends in Genetics*, 34, 666–681.

of Directors, A. B. (2017): "Laboratory and clinical genomic data sharing is crucial to improving genetic health care: a position statement of the American College of Medical Genetics and Genomics," *Genetics in Medicine*, 19, 721–722.

Dominitz, J. (1998): "Earnings expectations, revisions, and realizations," *Review of Economics and statistics*, 80, 374–388.

Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt (2011): "The framing of games and the psychology of play," *Games and Economic Behavior*, 73, 459–478.

Duquette, D., C. Langbo, J. Bach, and M. Kleyn (2012): "Michigan BioTrust for Health: public support for using residual dried blood spot samples for health research," *Public Health Genomics*, 15, 146–155.

Edwards, K. L., A. A. Lemke, S. B. Trinidad, S. M. Lewis, H. Starks, K. W. Snapinn, M. Q. Griffin, G. L. Wiesner, and W. Burke (2012): "Genetics researchers' and IRB professionals' attitudes toward genetic research review: a comparative analysis," *Genetics in medicine*, 14, 236.

Engelmann, D., and M. Strobel (2000): "The false consensus effect disappears if representative information and monetary incentives are given," *Experimental Economics*, 3, 241–260.

Erlich, Y., and A. Narayanan (2014): "Routes for breaching and protecting genetic privacy," *Nature Reviews Genetics*, 15, 409–421.

Exley, C., and J. B. Kessler (2019): "Motivated Errors," *Working Paper*.

Exley, C. L. (2016): "Excusing selfishness in charitable giving: The role of risk," *The Review of Economic Studies*, 83, 587–628.

Fehr, E., and K. M. Schmidt (1999): "A theory of fairness, competition, and cooperation," *The quarterly journal of economics*, 114, 817–868.

Feiler, L. (2014): "Testing models of information avoidance with binary choice dictator games," *Journal of Economic Psychology*, 45, 253–267.

Festinger, L. (1957): *A theory of cognitive dissonance* Volume 2: Stanford university press.

Fox, C. R., and A. Tversky (1995): "Ambiguity aversion and comparative ignorance," *The quarterly journal of economics*, 110, 585–603.

Gneezy, U., and A. Rustichini (2000): "A fine is a price," *The Journal of Legal Studies*, 29, 1–17.

Gneezy, U., S. Saccardo, M. Serra-Garcia, and R. van Veldhuizen (2016): "Motivated self-deception, identity and unethical behavior," *Working Paper*.

Greshake, B., P. E. Bayer, H. Rausch, and J. Reda (2014): "OpenSNP–A Crowdsourced Web Resource for Personal Genomics," *PLoS ONE*, 9, 10.1371/journal.pone.0089204.

Greytak, E., D. H. Kaye, B. Budowle, C. Moore, and S. Armentrout (2018): "Privacy and genetic genealogy data," *Science*, 361, 857, 10.1126/science.aav0330.

HAEUSERMANN, T., B. GRESHAKE, A. BLASIMME, D. IRDAM, M. RICHARDS, AND E. VAYENA (2017): "Open sharing of genomic data: Who does it and why?" *PLoS One*, 12.

HAISLEY, E. C., AND R. A. WEBER (2010): "Self-serving interpretations of ambiguity in other-regarding behavior," *Games and economic behavior*, 68, 614–625.

HARBAUGH, W. T. (1998): "The prestige motive for making charitable transfers," *The American Economic Review*, 88, 277–282.

HERTWIG, R., AND I. EREV (2009): "The description–experience gap in risky choice," *Trends in cognitive sciences*, 13, 517–523.

HERTWIG, R., AND A. ORTMANN (2003): "Economists' and psychologists' experimental practices: How they differ, why they differ, and how they could converge," in *The psychology of economic decisions*: Oxford University Press, 253–272.

HOLT, C. A., AND S. K. LAURY (2002): "Risk aversion and incentive effects," *American economic review*, 92, 1644–1655.

HULL, S. C., R. R. SHARP, J. R. BOTKIN, M. BROWN, M. HUGHES, J. SUGARMAN, D. SCHWINN, P. SANKAR, D. BOLCIC-JANKOVIC, B. R. CLARRIDGE ET AL. (2008): "Patients' views on identifiability of samples and informed consent for genetic research," *The American Journal of Bioethics*, 8, 62–70.

IRIBERRI, N., AND P. REY-BIEL (2013): "Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do?" *Quantitative Economics*, 4, 515–547.

Jacobson, S., and R. Petrie (2009): "Learning from mistakes: What do inconsistent choices over risk tell us?" *Journal of risk and uncertainty*, 38, 143–158.

Joly, Y., E. S. Dove, B. M. Knoppers, M. Bobrow, and D. Chalmers (2012): "Data sharing in the post-genomic world: the experience of the International Cancer Genome Consortium (ICGC) Data Access Compliance Office (DACO)," *PLoS computational biology*, 8.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–292.

Katz, D., F. H. Allport, and M. B. Jenness (1931): "Students' attitudes; a report of the Syracuse University reaction study.."

Kaufman, D., J. Murphy, L. Erby, K. Hudson, and J. Scott (2009): "Veterans' attitudes regarding a database for genomic research," *Genetics in Medicine*, 11, 329.

Krueger, J. (1998): "On the perception of social consensus," in *Advances in experimental social psychology* Volume 30: Elsevier, 163–240.

Krueger, J., and R. W. Clement (1994): "The truly false consensus effect: an ineradicable and egocentric bias in social perception.," *Journal of personality and social psychology*, 67, 596.

Lambarraa, F., and G. Riener (2015): "On the norms of charitable giving in Islam: two field experiments in Morocco," *Journal of Economic Behavior & Organization*, 118, 69–84.

Larson, T., and C. M. Capra (2009): "Exploiting moral wiggle room: Illusory preference for fairness? A comment," *Judgment and decision Making*, 4, 467.

Lazear, E. P., U. Malmendier, and R. A. Weber (2012): "Sorting in experiments with application to social preferences," *American Economic Journal: Applied Economics*, 4, 136–63.

Lemke, A. A., W. A. Wolf, J. Hebert-Beirne, and M. E. Smith (2010): "Public and biobank participant attitudes toward genetic research participation and data sharing," *Public Health Genomics*, 13, 368–377.

Levine, D. K. (1998): "Modeling altruism and spitefulness in experiments," *Review of economic dynamics*, 1, 593–622.

Loewenstein, G. F., E. U. Weber, C. K. Hsee, and N. Welch (2001): "Risk as feelings.," *Psychological bulletin*, 127, 267.

Mazar, N., O. Amir, and D. Ariely (2008): "The dishonesty of honest people: A theory of self-concept maintenance," *Journal of marketing research*, 45, 633–644.

Meier, S., and C. Sprenger (2010): "Present-biased preferences and credit card borrowing," *American Economic Journal: Applied Economics*, 2, 193–210.

Mijović-Prelec, D., and D. Prelec (2010): "Self-deception as self-signalling: a model and experimental evidence," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 365, 227–240.

Moore, D. A., and P. J. Healy (2008): "The trouble with overconfidence.," *Psychological review*, 115, 502.

Olson, K. R. (2006): "A literature review of social mood," *The Journal of Behavioral Finance*, 7, 193–203.

Pearson, K. (1895): "Contributions to the mathematical theory of evolution, II. Skew variation in homogeneous material," *Philosophical Transactions of the Royal Society of London.(A.)*, 343–414.

Prietzel, T. T. (2019): "The effect of emotion on risky decision making in the context of prospect theory: a comprehensive literature review," *Management Review Quarterly*, 1–41.

Proto, E., and D. Sgroi (2017): "Biased beliefs and imperfect information," *Journal of Economic Behavior & Organization*, 136, 186–202.

Quattrone, G. A., and A. Tversky (1984): "Causal versus diagnostic contingencies: On self-deception and on the voter's illusion.," *Journal of personality and social psychology*, 46, 237.

Rabin, M. (1994): "Cognitive dissonance and social change," *Journal of Economic Behavior & Organization*, 23, 177–194.

Rauscher, E. A., C. Hesse, S. Miller, W. Ford, and E. L. Youngs (2015): "Privacy and family communication about genetic cancer risk: Investigating factors promoting women's disclosure decisions," *Journal of Family Communication*, 15, 368–386.

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

Regner, T., and A. Matthey (2017): "Actions and the self: I give, therefore I am," *Jena Economic Research Papers*, 2017, 018.

Ross, L., D. Greene, and P. House (1977): "The false consensus effect: An egocentric bias in social perception and attribution processes," *Journal of experimental social psychology*, 13, 279–301.

Sanders, G. S., and B. Mullen (1983): "Accuracy in perceptions of consensus: Differential tendencies of people with majority and minority positions," *European journal of social psychology*, 13, 57–70.

Sanderson, S. C., K. B. Brothers, N. D. Mercaldo, E. W. Clayton, A. H. M. Antommaria, S. A. Aufox, M. H. Brilliant, D. Campos, D. S. Carrell, J. Connolly et al. (2017): "Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the US," *The American Journal of Human Genetics*, 100, 414–427.

Schlag, K. H., J. Tremewan, and J. J. Van der Weele (2015): "A penny for your thoughts: a survey of methods for eliciting beliefs," *Experimental Economics*, 18, 457–490.

Schotter, A., and I. Trevino (2014): "Belief elicitation in the laboratory," *Annu. Rev. Econ.*, 6, 103–128.

Schwardmann, P., and J. Van der Weele (2019): "Deception and self-deception," *Nature human behaviour*, 1–7.

Shariff, A. F., and A. Norenzayan (2007): "God is watching you: Priming God concepts increases prosocial behavior in an anonymous economic game," *Psychological science*, 18, 803–809.

Smith, M. K., R. Trivers, and W. von Hippel (2017): "Self-deception facilitates interpersonal persuasion," *Journal of Economic Psychology*, 63, 93–101.

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

Van Dijk, E., and H. Wilke (2000): "Decision-induced focusing in social dilemmas: Give-some, keep-some, take-some, and leave-some dilemmas.," *Journal of Personality and Social Psychology*, 78, 92.

Wade, L. (2018): "To overcome decades of mistrust, a workshop aims to train indigenous researchers to be their own genome experts," *Science*, 9.

Wu, G., and R. Gonzalez (1996): "Curvature of the probability weighting function," *Management science*, 42, 1676–1690.

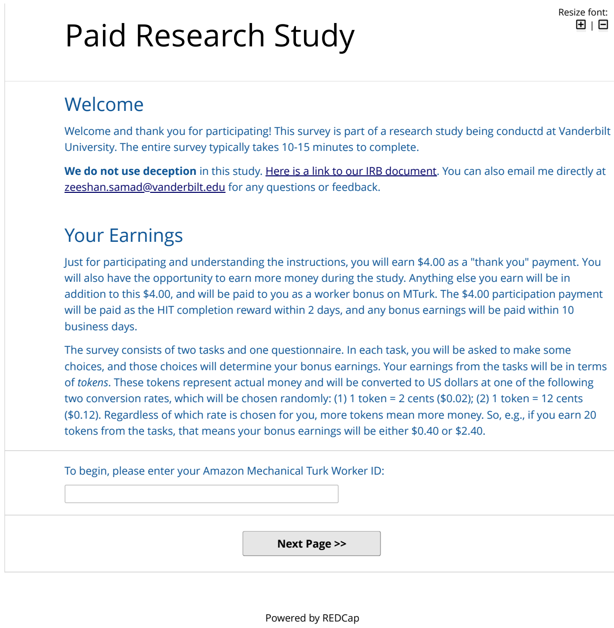
APPENDIX A

Supplementary Material for Chapter 1

A.1 Experiment Instructions

Some screenshots of the computerized survey, as seen by experiment participants, are shown below in Figures A.1 through A.11

Figure A.1: Experiment Instructions - Screen 1



Resize font:

Paid Research Study

Welcome

Welcome and thank you for participating! This survey is part of a research study being conductd at Vanderbilt University. The entire survey typically takes 10-15 minutes to complete.

We do not use deception in this study. Here is a link to our IRB document. You can also email me directly at zeeshan.samad@vanderbilt.edu for any questions or feedback.

Your Earnings

Just for participating and understanding the instructions, you will earn $4.00 as a "thank you" payment. You will also have the opportunity to earn more money during the study. Anything else you earn will be in addition to this $4.00, and will be paid to you as a worker bonus on MTurk. The $4.00 participation payment will be paid as the HIT completion reward within 2 days, and any bonus earnings will be paid within 10 business days.

The survey consists of two tasks and one questionnaire. In each task, you will be asked to make some choices, and those choices will determine your bonus earnings. Your earnings from the tasks will be in terms of *tokens*. These tokens represent actual money and will be converted to US dollars at one of the following two conversion rates, which will be chosen randomly: (1) 1 token = 2 cents ($0.02); (2) 1 token = 12 cents ($0.12). Regardless of which rate is chosen for you, more tokens mean more money. So, e.g., if you earn 20 tokens from the tasks, that means your bonus earnings will be either $0.40 or $2.40.

To begin, please enter your Amazon Mechanical Turk Worker ID:

Next Page >>

Powered by REDCap

Figure A.2: Experiment Instructions - Screen 2

Resize font:

Paid Research Study

Instructions

We are starting you off with 100 tokens and giving you the opportunity to donate some of these to a charity, keeping the remaining tokens for yourself. You will have two options:

Donate 0 tokens OR Donate 50 tokens

If you donate 50 tokens, then with some probability, which is unknown right now, the charity will get 120 tokens; and with the remaining probability, the charity will get 0 tokens. That probability will be a randomly generated number between 0 and 100, and rounded off to the nearest 10. This means that probability will be one of these:

0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%

Although you will be told what it is at the end of the survey, for now, you need to choose a donation amount without knowing that probability.

As an example, suppose the randomly generated probability is "70%". If you donate 50 tokens, then a randomization based on a 70-30 chance will determine if the charity will get 120 tokens or 0 tokens.

Understanding Questions

To ensure that you understand this, please answer the questions below. Failure to answer all questions correctly will disqualify you from participating in this study.

1 If you send 0 Tokens to the charity, then **You** will get:

0 Tokens	50 or 0 Tokens, depending on chance
50 Tokens	100 or 0 Tokens, depending on chance
100 Tokens	120 Tokens or 0, depending on chance
120 Tokens	

reset

2 If you send 0 Tokens to the charity, then the **Charity** will get:

0 Tokens	50 or 0 Tokens, depending on chance
50 Tokens	100 or 0 Tokens, depending on chance
100 Tokens	120 Tokens or 0, depending on chance
120 Tokens	

reset

3 If you send 50 Tokens to the charity, then **You** will get:

0 Tokens	50 or 0 Tokens, depending on chance
50 Tokens	100 or 0 Tokens, depending on chance
100 Tokens	120 Tokens or 0, depending on chance
120 Tokens	

reset

4 If you send 50 tokens to the charity, then the **Charity** will get:

0 Tokens	50 or 0 Tokens, depending on chance
50 Tokens	100 or 0 Tokens, depending on chance
100 Tokens	120 Tokens or 0, depending on chance
120 Tokens	

reset

5 Suppose you send 50 tokens to the charity. Also suppose that we use the conversion rate B (1 token = $0.12) to determine your earnings.

Then your bonus earnings will be:

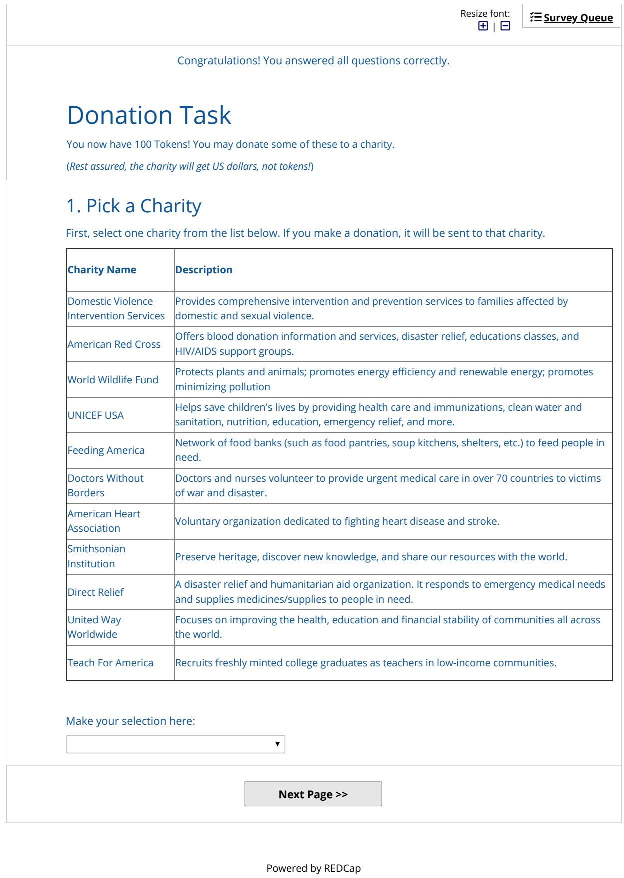
$0.00	$2.00
$4.00	$6.00
$8.00	$10.00

reset

Submit

Powered by REDCap

Figure A.3: Experiment Instructions - Screen 3



Resize font:

Survey Queue

Congratulations! You answered all questions correctly.

Donation Task

You now have 100 Tokens! You may donate some of these to a charity.

(*Rest assured, the charity will get US dollars, not tokens!*)

1. Pick a Charity

First, select one charity from the list below. If you make a donation, it will be sent to that charity.

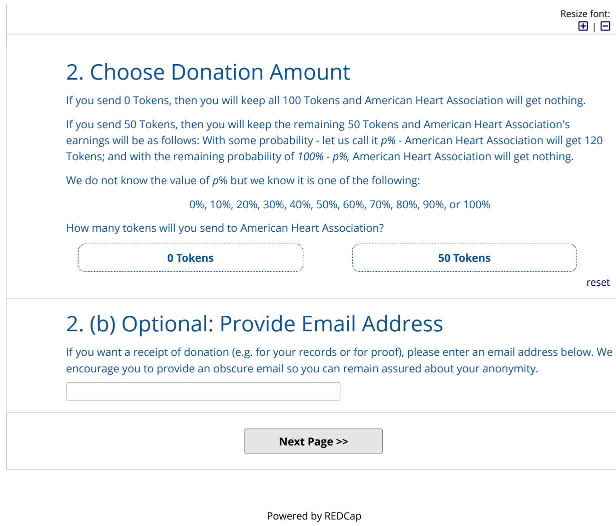
Charity Name	**Description**
Domestic Violence Intervention Services	Provides comprehensive intervention and prevention services to families affected by domestic and sexual violence.
American Red Cross	Offers blood donation information and services, disaster relief, educations classes, and HIV/AIDS support groups.
World Wildlife Fund	Protects plants and animals; promotes energy efficiency and renewable energy; promotes minimizing pollution
UNICEF USA	Helps save children's lives by providing health care and immunizations, clean water and sanitation, nutrition, education, emergency relief, and more.
Feeding America	Network of food banks (such as food pantries, soup kitchens, shelters, etc.) to feed people in need.
Doctors Without Borders	Doctors and nurses volunteer to provide urgent medical care in over 70 countries to victims of war and disaster.
American Heart Association	Voluntary organization dedicated to fighting heart disease and stroke.
Smithsonian Institution	Preserve heritage, discover new knowledge, and share our resources with the world.
Direct Relief	A disaster relief and humanitarian aid organization. It responds to emergency medical needs and supplies medicines/supplies to people in need.
United Way Worldwide	Focuses on improving the health, education and financial stability of communities all across the world.
Teach For America	Recruits freshly minted college graduates as teachers in low-income communities.

Make your selection here:

Next Page >>

Powered by REDCap

Figure A.4: Experiment Instructions - Screen 4


Resize font:

2. Choose Donation Amount

If you send 0 Tokens, then you will keep all 100 Tokens and American Heart Association will get nothing.

If you send 50 Tokens, then you will keep the remaining 50 Tokens and American Heart Association's earnings will be as follows: With some probability - let us call it *p%* - American Heart Association will get 120 Tokens; and with the remaining probability of *100% - p%*, American Heart Association will get nothing.

We do not know the value of *p%* but we know it is one of the following:

0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%

How many tokens will you send to American Heart Association?

0 Tokens **50 Tokens**

reset

2. (b) Optional: Provide Email Address

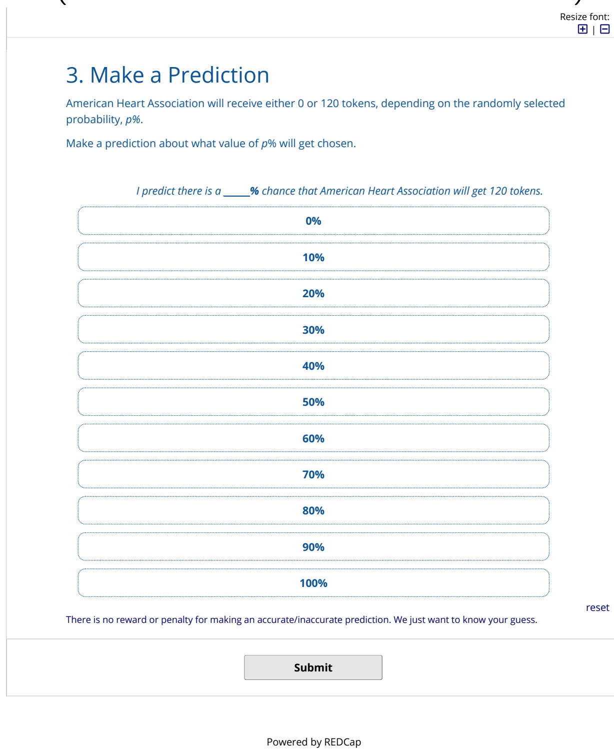
If you want a receipt of donation (e.g. for your records or for proof), please enter an email address below. We encourage you to provide an obscure email so you can remain assured about your anonymity.

Next Page >>

Powered by REDCap

Figure A.5: Experiment Instructions - Screen 5a

(show to those who donate 50 tokens)


Resize font:

3. Make a Prediction

American Heart Association will receive either 0 or 120 tokens, depending on the randomly selected probability, *p*%.

Make a prediction about what value of *p*% will get chosen.

I predict there is a ____% chance that American Heart Association will get 120 tokens.

- 0%
- 10%
- 20%
- 30%
- 40%
- 50%
- 60%
- 70%
- 80%
- 90%
- 100%

reset

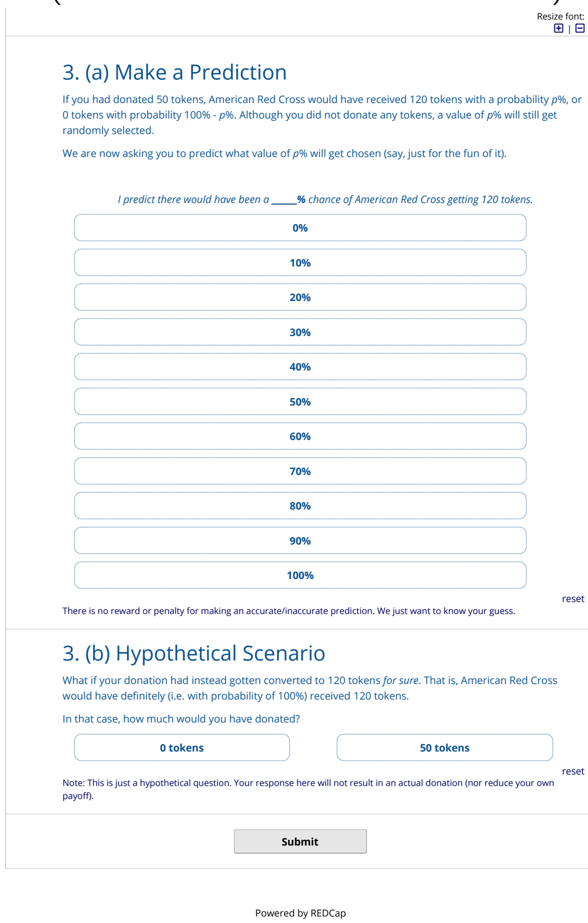
There is no reward or penalty for making an accurate/inaccurate prediction. We just want to know your guess.

Submit

Powered by REDCap

Figure A.6: Experiment Instructions - Screen 5b

(show to those who donate 0 tokens)



Resize font:

3. (a) Make a Prediction

If you had donated 50 tokens, American Red Cross would have received 120 tokens with a probability *p*%, or 0 tokens with probability 100% - *p*%. Although you did not donate any tokens, a value of *p*% will still get randomly selected.

We are now asking you to predict what value of *p*% will get chosen (say, just for the fun of it).

I predict there would have been a ______% chance of American Red Cross getting 120 tokens.

0%

10%

20%

30%

40%

50%

60%

70%

80%

90%

100%

reset

There is no reward or penalty for making an accurate/inaccurate prediction. We just want to know your guess.

3. (b) Hypothetical Scenario

What if your donation had instead gotten converted to 120 tokens *for sure*. That is, American Red Cross would have definitely (i.e. with probability of 100%) received 120 tokens.

In that case, how much would you have donated?

0 tokens | **50 tokens**

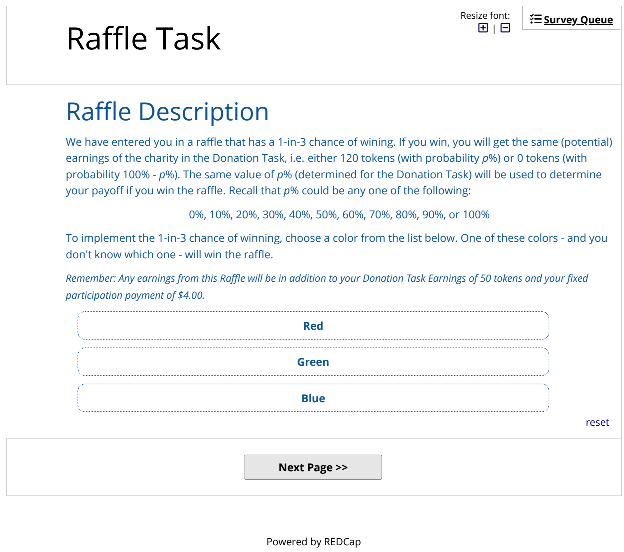
reset

Note: This is just a hypothetical question. Your response here will not result in an actual donation (nor reduce your own payoff).

Submit

Powered by REDCap

Figure A.7: Experiment Instructions - Screen 6


Resize font:

Survey Queue

Raffle Task

Raffle Description

We have entered you in a raffle that has a 1-in-3 chance of wining. If you win, you will get the same (potential) earnings of the charity in the Donation Task, i.e. either 120 tokens (with probability *p*%) or 0 tokens (with probability 100% - *p*%). The same value of *p*% (determined for the Donation Task) will be used to determine your payoff if you win the raffle. Recall that *p*% could be any one of the following:

0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%

To implement the 1-in-3 chance of winning, choose a color from the list below. One of these colors - and you don't know which one - will win the raffle.

Remember: Any earnings from this Raffle will be in addition to your Donation Task Earnings of 50 tokens and your fixed participation payment of $4.00.

Red

Green

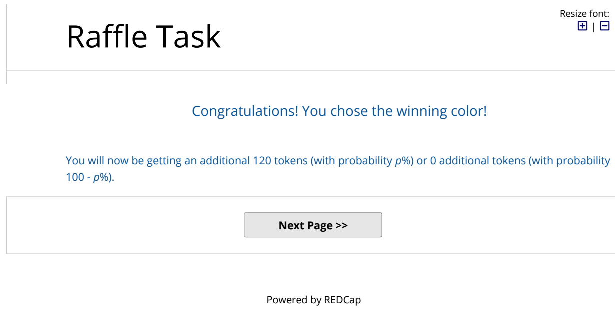
Blue

reset

Next Page >>

Powered by REDCap

Figure A.8: Experiment Instructions - Screen 7



Resize font:

Raffle Task

Congratulations! You chose the winning color!

You will now be getting an additional 120 tokens (with probability *p*%) or 0 additional tokens (with probability 100 - *p*%).

Next Page >>

Powered by REDCap

Figure A.9: Experiment Instructions - Screen 8

Resize font:

Raffle Task

Payment Options

On this page, tell us how you would like to receive your earnings from the Raffle Task. In each decision row below, choose whether you would like to be paid according to the first option or the second option.
(*These payments options are only for your Raffle Task Earnings and do not affect your Participation Fee or Donation Task Earnings.*)

In each row, the first option is your current raffle earnings: "120 tokens with prob *p*%; 0 tokens with prob 100 - *p*%". The second option is to get 120 tokens with a known probability and 0 tokens with the remaining probability stated in that row. One row will be randomly selected for payment, so it makes good sense to answer each row as if it is the one that determines your payment.

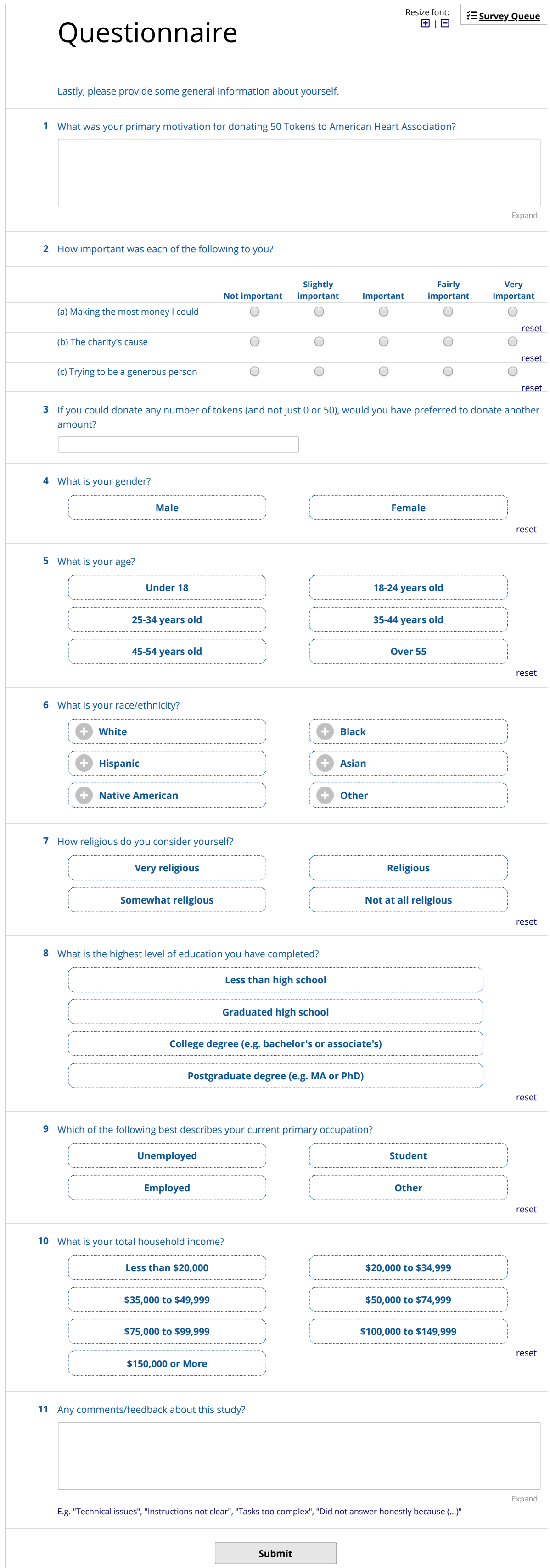
For example, suppose "Decision Row 3" is selected for payment and in this row you chose the second option. Then you will get 120 tokens (with a probability of 80%) or 0 tokens (with the remaining probability of 20%). Another randomization based on an 80-20 chance will determine the actual payoff. You will see your final earnings at the end of the survey.
(*If you would like to see for yourself that the chances are indeed what we say they are, please reach out to us. We will try to explain our randomization formula to you*)

Hint: The decisions in the first and last rows should be rather easy. In the first row, the second option is to receive 120 tokens for sure, which is clearly the better choice. In the last row, the second option is to receive 0 tokens for sure, which is clearly the *worse* choice. We have highlighted the options that are clearly better.

Decision Row	First option	Second option	
Decision Row 1	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 100% chance 0 tokens: 0% chance	reset
Decision Row 2	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 90% chance 0 tokens: 10% chance	reset
Decision Row 3	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 80% chance 0 tokens: 20% chance	reset
Decision Row 4	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 70% chance 0 tokens: 30% chance	reset
Decision Row 5	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 60% chance 0 tokens: 40% chance	reset
Decision Row 6	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 50% chance 0 tokens: 50% chance	reset
Decision Row 7	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 40% chance 0 tokens: 60% chance	reset
Decision Row 8	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 30% chance 0 tokens: 70% chance	reset
Decision Row 9	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 20% chance 0 tokens: 80% chance	reset
Decision Row 10	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 10% chance 0 tokens: 90% chance	reset
Decision Row 11	120 tokens: *p*% chance 0 tokens: 100-*p*% chance	120 tokens: 0% chance 0 tokens: 100% chance	reset

Submit

Figure A.10: Experiment Instructions - Screen 9



Resize font:

Survey Queue

Questionnaire

Lastly, please provide some general information about yourself.

1 What was your primary motivation for donating 50 Tokens to American Heart Association?

Expand

2 How important was each of the following to you?

	Not important	Slightly important	Important	Fairly important	Very Important
(a) Making the most money I could	○	○	○	○	○
(b) The charity's cause	○	○	○	○	○
(c) Trying to be a generous person	○	○	○	○	○

reset

reset

reset

3 If you could donate any number of tokens (and not just 0 or 50), would you have preferred to donate another amount?

4 What is your gender?

Male | Female

reset

5 What is your age?

Under 18 | 18-24 years old
25-34 years old | 35-44 years old
45-54 years old | Over 55

reset

6 What is your race/ethnicity?

White | Black
Hispanic | Asian
Native American | Other

7 How religious do you consider yourself?

Very religious | Religious
Somewhat religious | Not at all religious

reset

8 What is the highest level of education you have completed?

Less than high school
Graduated high school
College degree (e.g. bachelor's or associate's)
Postgraduate degree (e.g. MA or PhD)

reset

9 Which of the following best describes your current primary occupation?

Unemployed | Student
Employed | Other

reset

10 What is your total household income?

Less than \$20,000 | \$20,000 to \$34,999
\$35,000 to \$49,999 | \$50,000 to \$74,999
\$75,000 to \$99,999 | \$100,000 to \$149,999
\$150,000 or More

reset

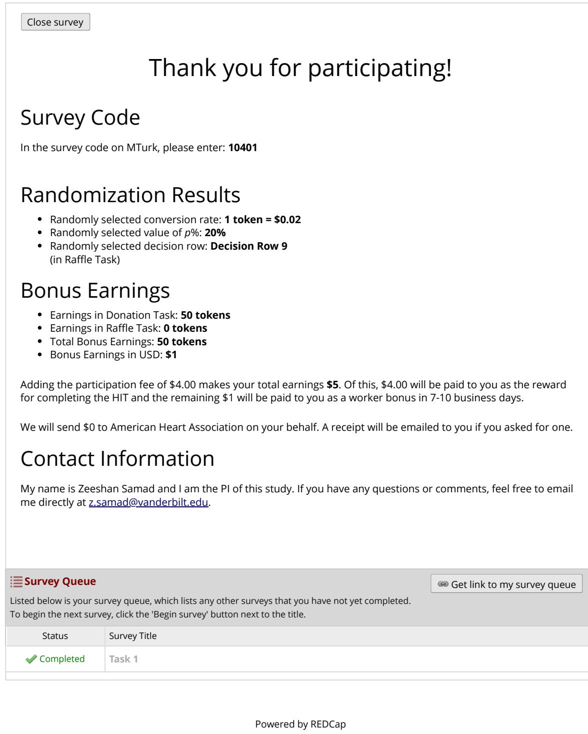
11 Any comments/feedback about this study?

Expand

E.g. "Technical issues", "Instructions not clear", "Tasks too complex", "Did not answer honestly because (...)"

Submit

Figure A.11: Experiment Instructions - Screen 10



Close survey

Thank you for participating!

Survey Code

In the survey code on MTurk, please enter: **10401**

Randomization Results

- Randomly selected conversion rate: **1 token = $0.02**
- Randomly selected value of *p*%: **20%**
- Randomly selected decision row: **Decision Row 9** (in Raffle Task)

Bonus Earnings

- Earnings in Donation Task: **50 tokens**
- Earnings in Raffle Task: **0 tokens**
- Total Bonus Earnings: **50 tokens**
- Bonus Earnings in USD: **$1**

Adding the participation fee of $4.00 makes your total earnings **$5**. Of this, $4.00 will be paid to you as the reward for completing the HIT and the remaining $1 will be paid to you as a worker bonus in 7-10 business days.

We will send $0 to American Heart Association on your behalf. A receipt will be emailed to you if you asked for one.

Contact Information

My name is Zeeshan Samad and I am the PI of this study. If you have any questions or comments, feel free to email me directly at z.samad@vanderbilt.edu.

Survey Queue

Get link to my survey queue

Listed below is your survey queue, which lists any other surveys that you have not yet completed.
To begin the next survey, click the 'Begin survey' button next to the title.

Status	Survey Title
Completed	Task 1

Powered by REDCap

A.2 Additional Hypotheses

This section presents additional hypotheses that can be tested using the data obtained from the lab experiments, but are not related to the primary research question.

Hypothesis 3: An individual can manipulate her belief only up to a certain extent.

Result 3: For the vast majority of subjects – specifically 87% – of subjects, the absolute value of the difference between adopted and true beliefs is less than or equal to 20%, suggesting that subjects do not go overboard while adopting false beliefs, because they must also find any adopted belief convincing – otherwise it would not reduce their cognitive dissonance.

Hypothesis 4: Subjects' true belief is $\mathbb{E}[p] = 0.50$.

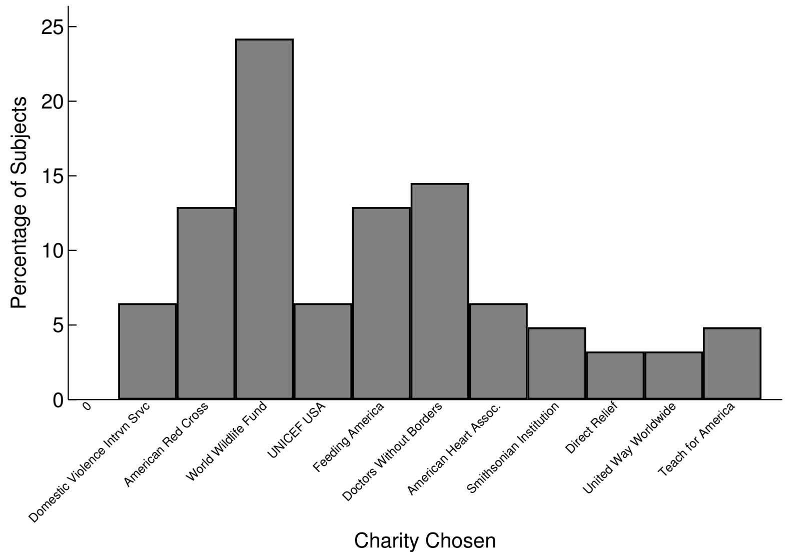
Result 4: The true belief of about 82% of subjects truly is in the range $0.40 \leq p_t \leq 0.60$.

In my experiment, subjects are only given an environment that contains ambiguity, and not one that contains unambiguous uncertainty. Therefore, my findings are consistent with a prominent finding by Fox and Tversky (1995), that individuals are ambiguity neutral in the absence of a choice between an ambiguous and an unambiguous outcome.

A.3 Additional Findings

This section contains additional figures that are not directly related to the main findings but are nonetheless insightful and contain interesting results. Figure A.12 shows the popularity of each charity. For some reason, WWF seems to be the most popular among MTurk workers. However, I do not observe any correlation between the charity chosen and donation decision, i.e., subjects who choose WWF are not any more likely to make a donation than subjects who choose other charities. Since subjects who have no intention to

Figure A.12: Popularity of Charities



donate might be more likely to choose the first listed option, simply because of convenience, each subject is given the list of charities in a randomized order.

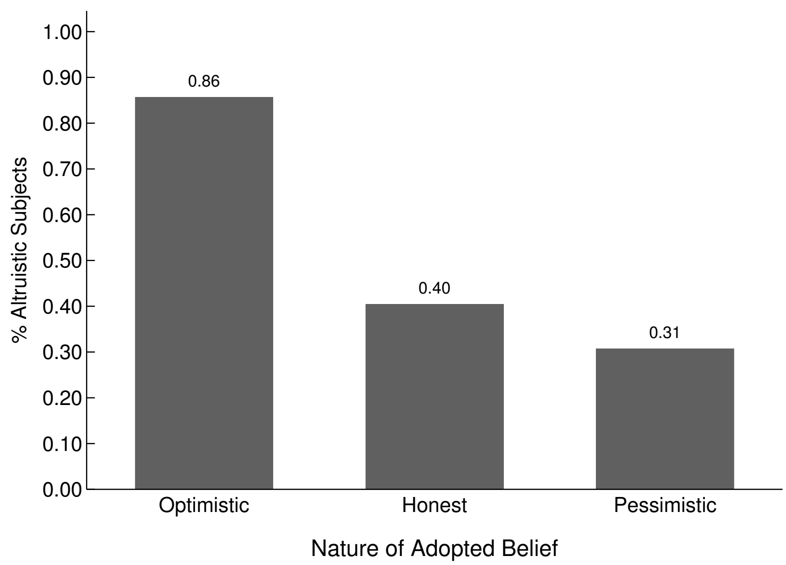
Figure A.13 shows the direction in which subjects manipulated their beliefs. In particular, whether subjects adopted more optimistic or more pessimistic beliefs. A total of 8 subjects adopted beliefs that were more optimistic than their true beliefs, 7 (86%) of whom behaved altruistically (i.e., made a donation). This is consistent with the hypothesis that only altruistic people have an incentive to adopt optimistic beliefs, probably because they want to believe that their act of donating will not go in vain.

There were 42 subjects who did not manipulate their beliefs in either direction (i.e., remained honest with themselves). Out of these, 17 subjects (40%) donated while the remaining 25 (60%) did not donate. These 42 subjects behaved according to their true beliefs. Assuming that everyone's true belief is $\mathbb{E}[p] = 0.5$, this suggests that about 40% of the population has a probability threshold (i.e., the lowest probability at which they are willing to donate) of something lower than 0.5, i.e., $p^* \leq 0.5$, and 60% of the population has a probability threshold of $p^* > 0.5$. This tells us something about the distribution of people's altruism level, $f_P(p^* = 0.5) = 0.4$.

Lastly, 13 subjects adopted beliefs that were more pessimistic than their true beliefs. And 4 of these 13 subjects (31%) also donated money along with adopting pessimistic beliefs. However, since this proportion (31%) is not significantly different (a two-tailed t-test results in $p = 0.538 > 0.05$) from the true proportion of altruistic subjects (40%), the most plausible explanation is that these 4 subjects did not think about the question carefully since it was un-incentivized, and ended up becoming part of the noise of the data. In other words, these 4 subjects acted based on their true beliefs, but responded inaccurately to the question about adopted beliefs.

It seems counter-intuitive for altruistic subjects to falsely adopt more pessimistic beliefs, so I examine this group of subjects a bit further in order to determine if they deliberately adopt more pessimistic beliefs or if it is something else (e.g., random selection or a mistake, etc). Out of a total of 10 altruistic and manipulative (AM) subjects, four subjects adopt more pessimistic beliefs than their true beliefs and the remaining six adopt more optimistic beliefs than their true beliefs. However, three of these four subjects adopt a belief that is greater than or equal to 50%, which suggests they are not deliberately adopting a pessimistic belief about p – if they were, they would adopt something lower than 50%. Instead, the more reason seems to be that their true beliefs about p are incredibly optimistic (at 70%, 90%, and 90%). Moreover, the last of these four subjects adopts a belief of 20% and holds a true belief of 40%, which does not make sense with her altruistic behavior. Moreover, this subject's qualitative responses are also consistent with selfish behavior, raising the possibility that she might have made a mistake when choosing whether or not to make a donation.

Figure A.13: Direction of Manipulation



Pairwise Comparison of Proportion of Altruistic Subjects, by Adopted Beliefs

Groups Compared	t-statistic	p-value
Optimistic vs Honest	2.292995	.0263682
Optimistic vs Pessimistic	2.611165	.0176792
Honest vs Pessimistic	.6202099	.5377773

APPENDIX B

Supplementary Material for Chapter 2

B.1 Experiment Instructions

Figures B.1 through B.11 show screenshots of the computerized version of the experiment, as seen by participants.

Figure B.1: Experiment Instructions - Screen 1

Resize font:

Paid Research Study

Welcome

Welcome, and thank you for participating! We are a team of researchers at Vanderbilt University in Nashville, Tennessee. This survey is part of a research study related to donation behavior and your participation will make an important contribution to research. This survey contains 2 decision tasks and a short questionnaire. The entire survey takes about 5 minutes to complete, so we hope you will not rush through it :)

We will not use deception in this study. This means we cannot go back on our word. So, if we say something like "We are giving you some money and want to see how you spend it", then at a later stage we canNOT say something like "We were just kidding about giving you money. We just wanted you to act as if it was actual money."

Click here to view our IRB document.

For any concerns/feedback, email zeeshan.samad@vanderbilt.edu.

Your Earnings

Just for participating and understanding the instructions, you will earn $3.00 as a "thank you" payment. This is the HIT completion reward and will be paid to you within 2 days.

You will also have the opportunity to earn more money within the survey tasks. These additional earnings will be in terms of "tokens", which represent actual money. At the end of the survey, your earnings in tokens will be converted to US dollars at the rate of **1 token = $0.13**. So, for example, if we say "we are giving you 10 tokens" it means we are giving you $1.30. These additional earnings will be paid to you as a worker bonus on MTurk within 2 weeks.

Lastly, we hope that the choices you make in this study are similar to the choices you would make with larger amounts of money.

To begin, please enter your Amazon Mechanical Turk Worker ID:

Next Page >>

Powered by REDCap

Figure B.2: Experiment Instructions - Screen 2

Resize font:
⊞ | ⊟

Paid Research Study

Qualifying Questions

This survey asks you to make decisions that require a certain level of mathematical understanding and comprehension. Please answer the following questions to ensure that you have these necessary skills.

You must answer all questions correctly in order to proceed.

Please be sure about your answers before hitting submit, because you will not be able to attempt these a second time.

Question 1

If you earn 10 tokens during the survey tasks, what will be your *bonus earnings in USD*?

Recall that the conversion rate is 1 token = $0.13

Question 2

A survey question asks respondents to choose between three options: option A, option B, or option C. We find that 20% choose option A and 25% choose option B. What percentage must have chosen option C?

Question 3

See the figure below and answer the questions that follow.

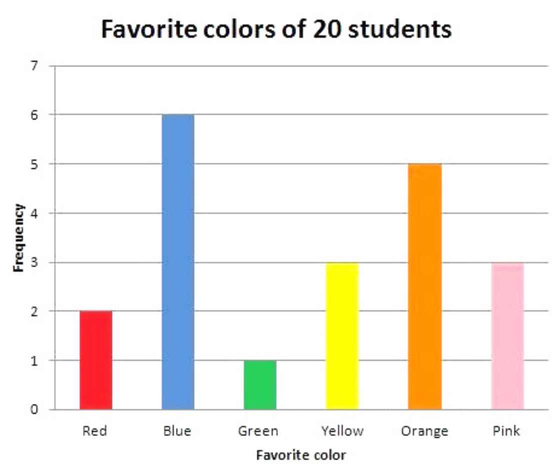


Figure B.3: Experiment Instructions - Screen 2 (contd)

3 (a) Which is the most popular favorite color?

3 (b) What % of students report their favorite color to be Blue?

Hint: the question asks about *percentage*, not *number* of students.

3 (c) What % of students report a favorite color *Other Than* Red?

Question 4

The table below describes characteristics of survey respondents. Use this table to answer the questions that follow.

	Full Sample	
Variables	*n*	%
Gender		
Female	239	69
Male	108	31
Race/Ethnicity		
Caucasian	130	37
Hispanic	72	21
African American	73	21
Asian American	39	11
Mixed race/Other	34	10
Relationship Status		
Married/In a relationship	239	70
Not in a relationship	102	30

4 (a) What is the sample size?

Hint: use the gender variable

(in terms of number of people)

4 (b) What % of respondents are either male or female?

Submit

Powered by REDCap

Figure B.4: Experiment Instructions - Screen 3

Resize font:

Survey Queue

Donation Task

Congratulations! You answered all questions correctly.

This is where the actual research study begins.

Donation Task

We are now giving you 10 tokens. You may donate some (or none, or all) of these to a charity and keep the remainder for yourself.

Remember that tokens represent actual money and will be converted to US dollars based on the conversion rate mentioned earlier.

1. Pick a Charity

First, select one charity from the list below. Any donation you make will be sent to that charity.

Charity Name	**Charity's Cause**
Domestic Violence Intervention Services	Provides comprehensive intervention and prevention services to families affected by domestic and sexual violence.
American Red Cross	Offers blood donation information and services, disaster relief, educations classes, and HIV/AIDS support groups.
World Wildlife Fund	Protects plants and animals; promotes energy efficiency and renewable energy; promotes minimizing pollution
UNICEF USA	Helps save children's lives by providing health care and immunizations, clean water and sanitation, nutrition, education, emergency relief, and more.
Feeding America	Network of food banks (such as food pantries, soup kitchens, shelters, etc.) to feed people in need.
Doctors Without Borders	Doctors and nurses volunteer to provide urgent medical care in over 70 countries to victims of war and disaster.
American Heart Association	Voluntary organization dedicated to fighting heart disease and stroke.
Smithsonian Institution	Preserve heritage, discover new knowledge, and share our resources with the world.
Direct Relief	A disaster relief and humanitarian aid organization. It responds to emergency medical needs and supplies medicines/supplies to people in need.
United Way Worldwide	Focuses on improving the health, education and financial stability of communities all across the world.
Teach For America	Recruits freshly minted college graduates as teachers in low-income communities.

Make your selection here:

Figure B.5: Experiment Instructions - Screen 3 (contd)

2. Donation Amount

How many of your 10 tokens would like to donate to ______?

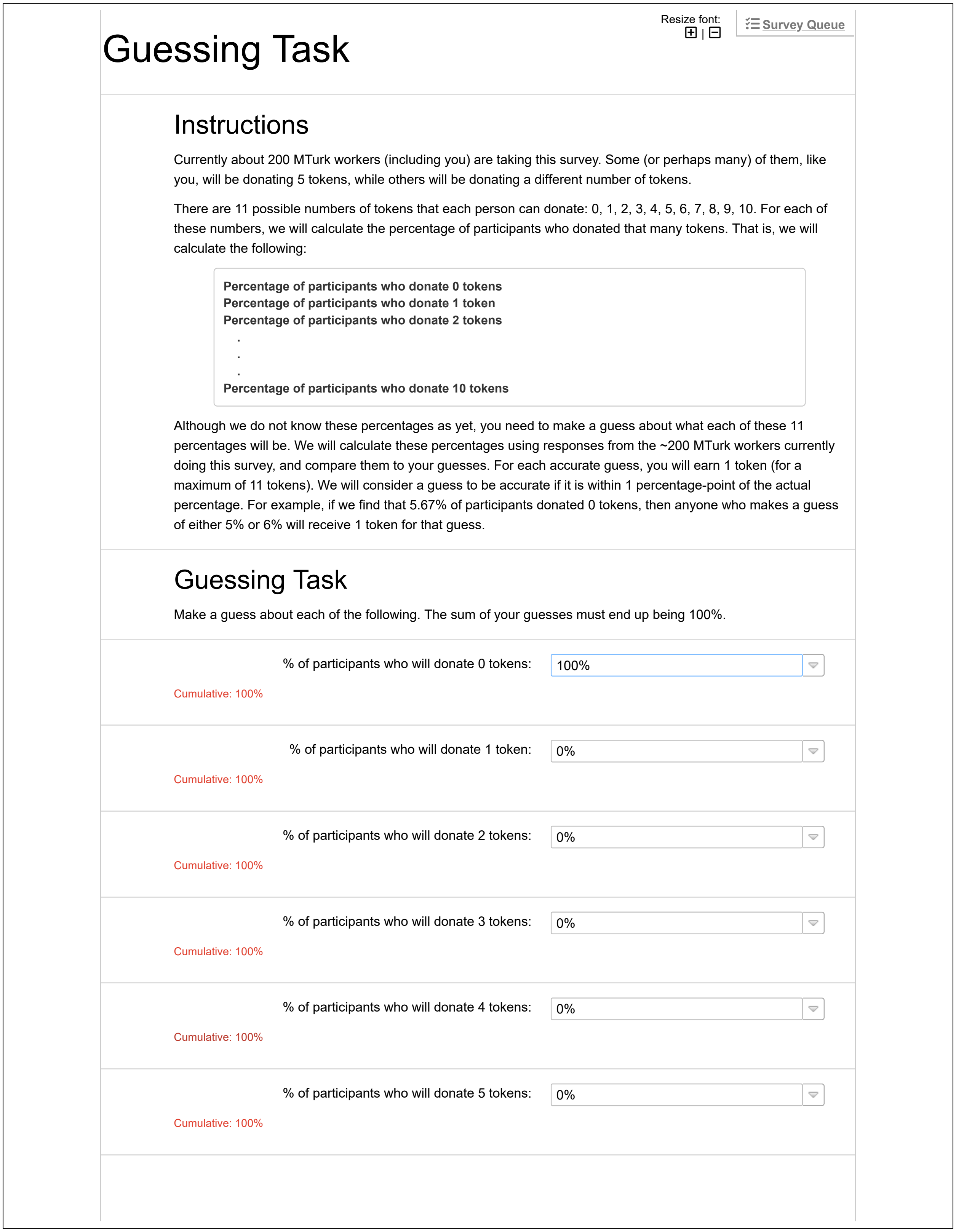
3. [Optional] Provide Email Address

If you want a receipt of donation (e.g. for proof), please enter an email address below. Feel free to provide an obscure email address if that helps in assuring you your anonymity.

Submit

Powered by REDCap

Figure B.6: Experiment Instructions - Screen 4



Resize font:

Survey Queue

Guessing Task

Instructions

Currently about 200 MTurk workers (including you) are taking this survey. Some (or perhaps many) of them, like you, will be donating 5 tokens, while others will be donating a different number of tokens.

There are 11 possible numbers of tokens that each person can donate: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. For each of these numbers, we will calculate the percentage of participants who donated that many tokens. That is, we will calculate the following:

Percentage of participants who donate 0 tokens
Percentage of participants who donate 1 token
Percentage of participants who donate 2 tokens
.
.
.
Percentage of participants who donate 10 tokens

Although we do not know these percentages as yet, you need to make a guess about what each of these 11 percentages will be. We will calculate these percentages using responses from the ~200 MTurk workers currently doing this survey, and compare them to your guesses. For each accurate guess, you will earn 1 token (for a maximum of 11 tokens). We will consider a guess to be accurate if it is within 1 percentage-point of the actual percentage. For example, if we find that 5.67% of participants donated 0 tokens, then anyone who makes a guess of either 5% or 6% will receive 1 token for that guess.

Guessing Task

Make a guess about each of the following. The sum of your guesses must end up being 100%.

% of participants who will donate 0 tokens: 100%
Cumulative: 100%

% of participants who will donate 1 token: 0%
Cumulative: 100%

% of participants who will donate 2 tokens: 0%
Cumulative: 100%

% of participants who will donate 3 tokens: 0%
Cumulative: 100%

% of participants who will donate 4 tokens: 0%
Cumulative: 100%

% of participants who will donate 5 tokens: 0%
Cumulative: 100%

Figure B.7: Experiment Instructions - Screen 4 (contd)

% of participants who will donate 6 tokens: 0%

Cumulative: 100%

% of participants who will donate 7 tokens: 0%

Cumulative: 100%

% of participants who will donate 8 tokens: 0%

Cumulative: 100%

% of participants who will donate 9 tokens: 0%

Cumulative: 100%

% of participants who will donate 10 tokens: 0%

Cumulative: 100%

Next Page >>

Powered by REDCap

Figure B.8: Experiment Instructions - Screen 5

Resize font:

Guessing Task

Review

Below are your guesses about previous participants' donation choices.

100% will donate 0 tokens.
0% will donate 1 token.
0% will donate 2 tokens.
0% will donate 3 tokens.
0% will donate 4 tokens.
0% will donate 5 tokens.
0% will donate 6 tokens.
0% will donate 7 tokens.
0% will donate 8 tokens.
0% will donate 9 tokens.
0% will donate 10 tokens.

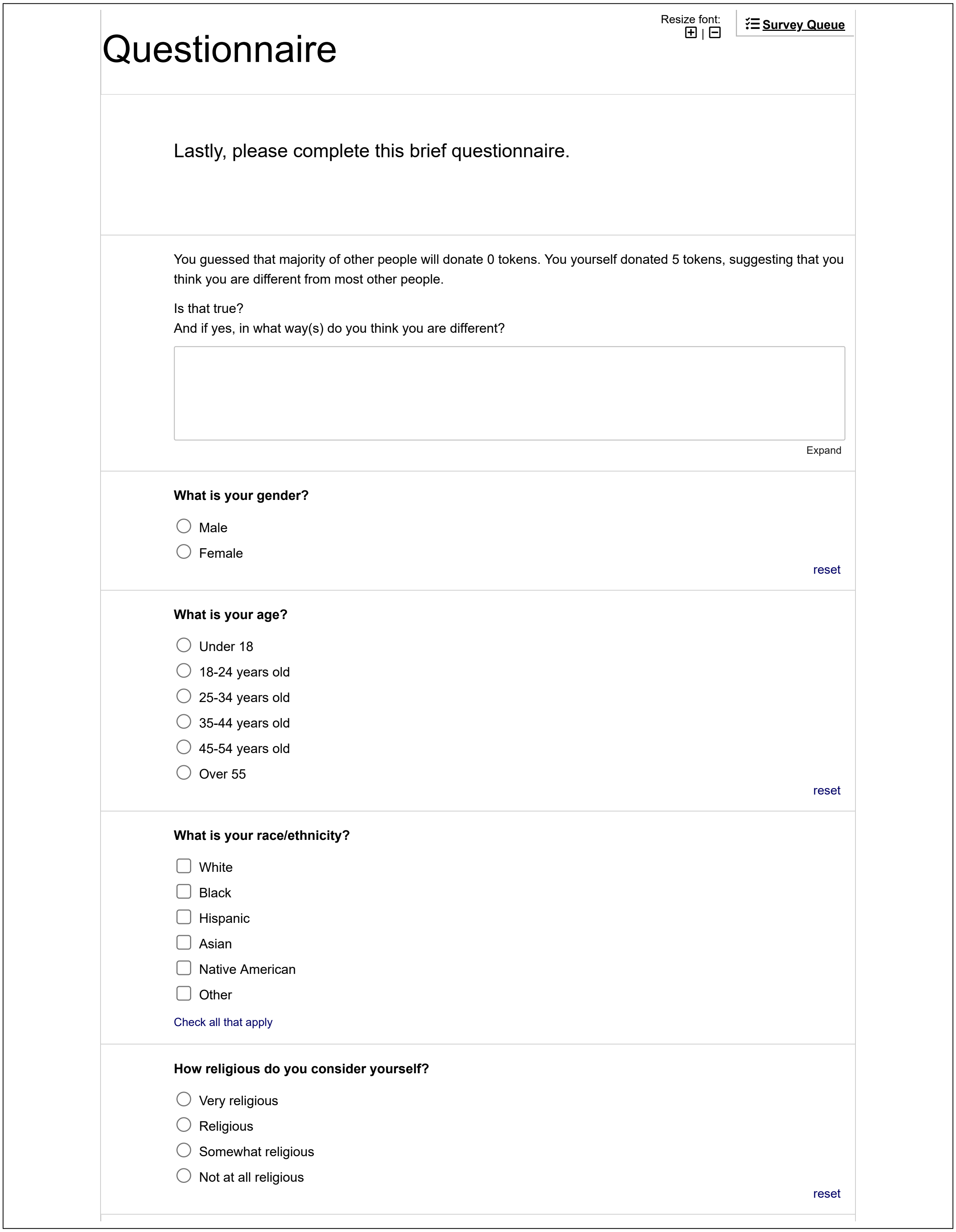
Recall that you yourself donated 5 token(s).

If you want to revise your guesses, you may go back to the Previous Page. Otherwise go ahead and Submit.

<< Previous Page **Submit**

Powered by REDCap

Figure B.9: Experiment Instructions - Screen 6



Resize font:

Survey Queue

Questionnaire

Lastly, please complete this brief questionnaire.

You guessed that majority of other people will donate 0 tokens. You yourself donated 5 tokens, suggesting that you think you are different from most other people.

Is that true?
And if yes, in what way(s) do you think you are different?

Expand

What is your gender?

- ○ Male
- ○ Female

reset

What is your age?

- ○ Under 18
- ○ 18-24 years old
- ○ 25-34 years old
- ○ 35-44 years old
- ○ 45-54 years old
- ○ Over 55

reset

What is your race/ethnicity?

- ☐ White
- ☐ Black
- ☐ Hispanic
- ☐ Asian
- ☐ Native American
- ☐ Other

Check all that apply

How religious do you consider yourself?

- ○ Very religious
- ○ Religious
- ○ Somewhat religious
- ○ Not at all religious

reset

Figure B.10: Experiment Instructions - Screen 6 (contd)

What is your highest level of education?

- ◯ Less than high school
- ◯ Graduated high school
- ◯ College degree (e.g. bachelor's or associate's)
- ◯ Postgraduate degree (e.g. MA or PhD)

reset

Which of the following best describes your primary occupation?

- ◯ Unemployed
- ◯ Student
- ◯ Employed
- ◯ Other

reset

What is your total annual household income?

- ◯ Less than $20,000
- ◯ $20,000 to $34,999
- ◯ $35,000 to $49,999
- ◯ $50,000 to $74,999
- ◯ $75,000 to $99,999
- ◯ $100,000 to $149,999
- ◯ $150,000 or More

reset

Any additional feedback/suggestions?

We read all comments carefully, and will use your suggestions to improve the way we create/conduct surveys.

Expand

Submit

Powered by REDCap

Figure B.11: Experiment Instructions - Screen 7

Close survey

Thank you for participating!

Survey Code

Your unique survey code for MTurk is: **101324**

Participation Payment

Your participation payment of $3.00 will be approved on MTurk within 2 days.

Bonus Earnings

Bonus Earnings in Donation Task = 5 tokens = $0.65

Bonus Earnings from Guessing Task = To be determined

Your bonus earnings will be paid to you as a worker bonus on MTurk within 2 weeks. Please feel free to get in touch if you do not receive any bonus within that time frame or if you feel that you received an incorrect amount.

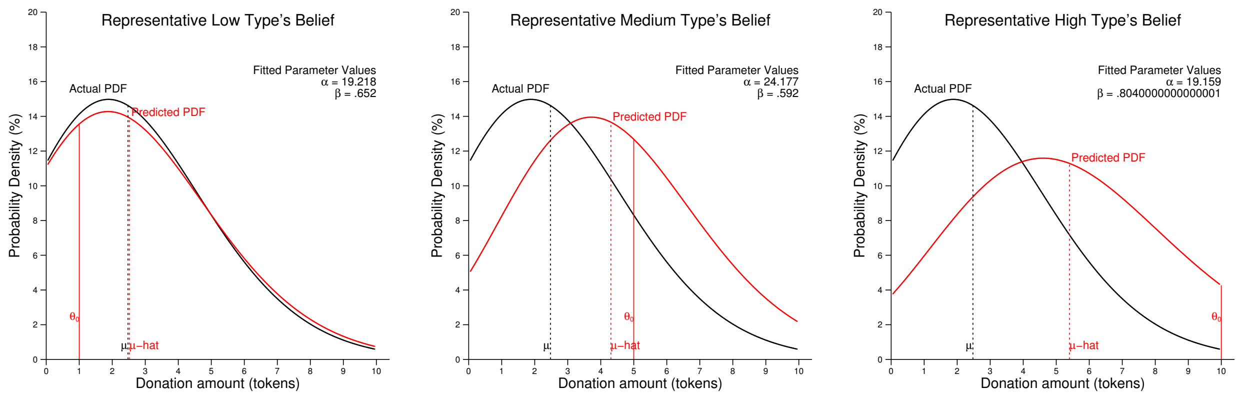
We will also send $0.65 to Direct Relief on your behalf. If you asked for a receipt, it will be emailed to you.

Contact Information

Name: Zeeshan Samad
Email: zeeshan.samad@vanderbilt.edu
Affiliation: Department of Economics, Vanderbilt University

Figure B.12: Actual vs. Predicted Distributions

(a) Low Types' Beliefs (b) Medium Types' Beliefs (c) High Types' Beliefs



B.2 Additional Results

Table B.1 shows the mean characteristics of experiment subjects, for each subject type. The variables for age, gender, race, education, and income are not significantly different across types. Interestingly, religiousness is significantly lower for low types ($t = 4.4743$ for a two-tailed t-test between low and medium types), suggesting that less religious people are less altruistic.

Table B.2 provides a tabular presentation of the actual distribution (Figure 1.4) and each type's belief about it (Figure 2.2).

Figure B.12 shows how accurately each type of subject predicts the distribution of donation amounts.

Table B.1: Summary Statistics by Subject Type

Variable	Low Types	Medium Types	High Types	Total
Own Donation Amount	0.739	4.742	10	2.477
	(0.0847)	(0.122)	(.)	(0.258)
Mode of Predicted Distribution	0.693	3.806	5.364	1.831
	(0.200)	(0.496)	(1.397)	(0.259)
% Female	0.273	0.419	0.364	0.315
	(0.0477)	(0.0901)	(0.152)	(0.0409)
Age Bracket (1-6)	3.602	3.839	4	3.692
	(0.111)	(0.203)	(0.381)	(0.0951)
% White	0.830	0.774	0.818	0.815
	(0.0403)	(0.0763)	(0.122)	(0.0342)
Religiousness (0-3)	0.670	1.581	1.182	0.931
	(0.102)	(0.184)	(0.377)	(0.0934)
Years of Education	15.18	15.35	15.45	15.25
	(0.228)	(0.352)	(0.718)	(0.184)
Income Category (1-7)	3.909	3.710	3.545	3.831
	(0.175)	(0.218)	(0.593)	(0.138)
Number of Subjects	88	31	11	130

Notes: Standard errors are in parentheses; All values are averages over subject type. The variable *own donation amount* is the donation made by subjects themselves (this is also the variables that defines type; in particular, low type is defined as subjects who donate 2 tokens or less, medium type as subjects who donate 3-6 tokens, and high type as subjects who donate 7 or more tokens); *mode of prediction* is the mode of the distribution predicted by subjects; *age bracket* is a categorical variable, with 1: <18, 2: 18-25, 3: 35-35, 4: 35-45, 5: 45-55, and 6: 55+; *income category* ranges from 1 to 7, with 1: < $20k p.a., 2: $25-35k, 3: $35-50k, 4: $50-75k, 5: $75-100k, 6: $100-150k, 7: > $150k.

Table B.2: Average Predictions Made by Low, Medium, and High Types

	Actual Frequency		Low Types' Beliefs		Medium Types' Beliefs		High Types' Beliefs	
Donation Amount	number	%	(%)	*p*-value	(%)	*p*-value	(%)	*p*-value
0	42	32.308	39.2273	0.0178	13.7742	0.0000	13.9091	0.0012
1	27	20.769	14.2841	0.0000	8.4194	0.0000	4.7273	0.0000
2	19	14.615	9.8295	0.0000	7.5484	0.0000	4.9091	0.0000
3	3	2.308	6.8523	0.0000	8.2903	0.0000	4.8182	0.0638
4	3	2.308	5.0795	0.0000	8.2581	0.0000	7.1818	0.0248
5	24	18.462	8.0909	0.0000	25.6774	0.0226	21.3636	0.5592
6	1	0.769	3.3182	0.0000	7.0000	0.0000	6.0000	0.0025
7	0	0.000	2.9318	0.0000	6.1935	0.0000	4.8182	0.0032
8	0	0.000	2.5341	0.0000	4.4516	0.0000	4.3636	0.0103
9	0	0.000	2.1364	0.0000	3.9355	0.0000	7.9091	0.0076
10	11	8.462	5.7159	0.0009	6.4516	0.1321	20.0000	0.0586

Notes. The *p*-value columns conducts a t-test for inaccuracy of predictions. Thus, a statistically significant difference ($p < 0.05$) means that the prediction is inaccurate. For example, the first row of Table B.2a shows that, on average, low types predict that 39% of all subjects will donate 0 tokens. This is different from the actual value (32%) by 7 percentage points (or 17.6 *percent*). The corresponding *p*-value of 0.0178 shows that this prediction is inaccurate at a 5% significance level (we reject the null hypothesis that the prediction is accurate). Table B.2b shows that low types think that about 68% of the sample consists of low types.

APPENDIX C

Supplementary Material for Chapter 3

C.1 Appendix Introduction

In this appendix, we provide all experiment instruments employed in the study, present more detailed results of regression results, and discuss several theoretical and methodological aspects of the study. The structure of this document is as follows. Section C.2 provides the survey instructions that were used in the experiment. Section C.3 provides detailed results of regression and statistical analyses that are not only informative but also serve as robustness checks for our results. Section C.4 further discusses the interpretation of our result, particularly that people value potential health benefits more than potential financial benefits. Section C.5 discusses how we develop our experimental methodology given our specific research question. Section C.6 provides results and details of an online replication of the laboratory experiment.

C.2 Experiment Instructions

Subjects are asked to make the same decision in three different settings (called *scenarios* in the actual instructions). Data sharers participate in the *base*, *altruism*, and *trust* settings while data recipients participate in the *recipient-altruism* and *reciprocity* settings. In order to prevent any experimenter-demand effects, we do not use these names with subjects and instead refer to the settings as Scenario 1, Scenario 2, etc. Since subjects receive settings in different orders, Scenario 1 does not always refer to any one particular setting (e.g., the base setting). Each Scenario is presented to subjects on a single computer screen.

Figures C.1 through C.12 show experiment instructions as seen by subjects. It may help to divide these 12 figures into four sets, one each for patients, investors, physicians, and money managers, where each set contains three settings. Thus, Figures C.1, C.2, and C.3 show the text seen by patients (i.e., data sharers in the genetic frame), Figures C.4, C.5, and C.6 contain the text seen by investors (i.e., data sharers in the investment frame), Figures C.7, C.8, and C.9 contain the text seen by physicians (i.e., data recipients in the genetic frame), and Figures C.10, C.11, and C.12 contain the text seen by money managers (i.e., data recipients in the investment frame)

C.3 Additional regression results

As shown in Table 1 of the main paper, we find framing effects in the base setting. However, we do not get the same result in other settings. Regression results for each setting are shown in Tables C.1 through C.5. Specifically, Tables C.1, C.2, and C.3 present regression results for data sharers in the base, altruism, and trust settings respectively; and Tables C.4 and C.5 present regression results for data recipients in the recipient-altruism and reciprocity settings respectively. In each of these regressions, the dependent variable represents the proportion of subjects who share data (in case of data sharers) or protect another person's data (in case of data recipients). The key independent variable of interest is a binary variable that indicates whether a subject got assigned to the genetic frame or investment frame. The estimated coefficient of this variable tells us the extent to which responses are different across the genetic and investment frames or, in other words, the extent to which people's concerns about genetic data privacy are different from their concerns about financial data privacy.

Tables C.6 and C.7 show, for data sharers and data recipients respectively, how much responses differ across each setting and in each frame. Table C.6a shows the proportion of patients who choose to get a genetic test in the base, altruism, and trust settings; and

Table C.6b shows the proportion of investors who choose to get a genetic test in the base, altruism, and trust settings. Table C.7a shows the proportion of physicians who protect a physician's genetic data in the recipient-altruism and reciprocity settings. Table C.7b shows the proportion of money managers who protect an investor's financial data in the recipient-altruism and reciprocity settings.

The specific findings presented in each table are described in the table's notes.

C.4 Interpretation of results

Our main finding is that individuals are more willing to put their genetic data at risk for the sake of a health benefit than they are willing to put their financial data at risk for the sake of a monetary benefit. There are two interrelated forces that appear to drive our results and their interpretation: (i) different perceptions of benefits, and/or (ii) different perceptions about risk.

The first possible reason for our result is that people consider a $60 earning as more valuable when it is presented as a health benefit than when it is presented as a financial benefit. The second possible reason is that people consider a 25% chance of losing $4 as worse when it is presented as their financial data being compromised than when it is presented as their genetic data being compromised. This second possibility is supported by a robust finding about cognitive biases that people are more prone to overstating and understating probabilities than monetary amounts (Kahneman and Tversky, 1979; Loewenstein, Weber, Hsee and Welch, 2001; Wu and Gonzalez, 1996).[1] In particular, people overstate probabilities smaller than 50% and understate probabilities greater than 50% and these perceptions are increasingly incorrect for probabilities farther from 50%. Moreover, a

[1] See Hertwig and Erev (2009); Olson (2006); Prietzel (2019) for a more comprehensive literature survey on cognitive biases.

variety of factors can influence the extent to which people distort probabilities, from the dollar value of the amount to their current mood/emotions (e.g. subjects become more risk-seeking if they are feeling more fearless/adventurous).

We observe that people are more willing to put their genetic data at risk for a health benefit than they are willing to put their financial data at risk for a financial benefit. Because both frames have the same expected monetary payoffs and because a larger percentage of subjects choose to share data in the genetic test than in the investment frame, we conclude that on average subjects value the potential health benefits (i.e., derive more utility from potential health benefits) more than the potential financial benefits.

C.5 Experiment design

Instead of using framing effects, we could have approached our research question by presenting the problem only as a decision to get a genetic test (i.e. keeping only the genetic frame), and could have compared subjects' average willingness to share data with that of a risk-neutral individual. In that case, we would have attributed risk-averse or risk-seeking behavior to the framing. However, this approach would require us to assume that individuals are typically risk-neutral, which contradicts experimental evidence about risk-aversion.

By contrast, by using two separate frames, we can measure subjects' willingness to share genetic data *relative* to their willingness to share financial data, without making any assumptions about their baseline risk preferences. That is, we investigate whether subjects are systematically *more* or *less* risk-averse with respect to sharing genetic data than with respect to sharing financial data. For example, a subject who exhibits risk aversion in sharing genetic data, but even greater risk aversion in sharing financial data, is actually revealing that she is relatively risk-seeking with respect to sharing genetic data.

C.6 Online Replication of Experiment

We replicated the experiment described in the main paper using Amazon Mechanical Turk (AMT), which is a micro-employment platform that allows employers to hire workers to perform small tasks for modest compensation. We tried to replicate the experiment as faithfully as possible but, due to technical limitations, were not able to have subject pairs (of data sharers and data recipients) play the game simultaneously in real-time, as subjects did in the lab experiment. However, we do not believe this discrepancy caused any significant bias or problems in comparing results for multiple reasons.

First, the base setting is naturally a single-person decision without even a mention of a data recipient, which means that as far as this setting is concerned, the online version is a perfect replication of the lab version. As a result, at least the base setting results are completely comparable.

Second, the altruism setting only mentions the data recipient (i.e., physician or money manager) as a passive recipient, not giving him the ability to take any action. Thus, payoffs in the altruism setting are determined solely by data sharers' (i.e., patients' or investors') own actions, not actions of data recipients.

Third, in the *trust and reciprocity* setting, which is the only setting potentially threatened by technical issues, the interaction between data sharers and data recipients is minimal, computerized, and anonymous. Therefore, although each party's payoff partially depends on the other party's choice, the act of interacting itself is unlikely to have any influence on subjects' choices. The only way this shortcoming might be problematic is if it makes subjects systematically more skeptical about whether the other party is really represented by an AMT worker and not the experimenter. For example, if physicians think that their patient (who chose to undergo genetic testing) is really the experimenter, it would blur the difference between the *reciprocity* setting and the *recipient-altruism* setting. Based on these observations, it appears that this replication shortcoming does not create a bias in the results, especially with regard to framing effects.

In the online version, we conducted computerized surveys individually. We assigned 120 subjects the role of patients (i.e. assigned them the role of data sharers and the genetic frame), another 120 subjects the role of investors, 50 subjects the role of physicians, and 50 subjects the role of money managers. Thus, we recruited a total of 340 AMT workers for the online version of the experiment. Patients/investors who opted to get a genetic test/make an investment were randomly assigned to physicians/money managers, and subjects' earnings were calculated accordingly. This was all done after all subjects had submitted their responses and their earnings were paid to them as a "worker bonus" on AMT.

Tables C.8, C.9, and C.10 show the regression results for the base, altruism, and trust settings respectively. In all three regressions, the dependent variable represents the proportion of data sharers who choose to share data (either in the form of getting a genetic test or in the form of making an investment). Comparing these results to those obtained through the lab experiment (i.e., Tables C.1, C.2, and C.3, respectively) we find the two sets of results are highly similar in several notable respects. First, in both the lab and online experiments, patients are more willing to put their genetic data privacy at risk for the sake of a health benefit than investors are willing to put their financial data privacy at risk for the sake of monetary benefits. Second, the same variables are statistically significant in both the lab and online versions of the experiment.

Figure C.13 presents results of the AMT experiment in the form of bar charts, which are comparable to the bar charts for the lab experiment (Figure 2 in the main paper).

Comparing the lab and online versions of the experiment, we observe two interesting patterns.

First, in all frames and settings, lab subjects are about 20 percentage points more likely to share their data than online subjects. This suggests that online subjects are more risk-averse than lab subjects. As for why this might be the case, one conjecture is that online subjects are somewhat skeptical about the 5% chance of winning $60, whereas lab subjects have greater faith in the experimenter's integrity of implementing this chance. Another conjecture is that online subjects have systematically different characteristics, such as income levels, which is shown to have a correlation with risk-aversion.

Second, in both online and lab versions of the experiment, we find that the number of patients who share their data is always greater by about 60% than the number of investors who share their data. For example, comparing results for the base setting, the lab experiment numbers show that 72% of patients share their data, which is 59% more than the proportion of investors who share their data (45%). In the online experiment, 45% of patients share their data, which is also 59% more than the proportion of investors who share their data, which is 28% in the online version.

Figure C.1: Instructions given to Patients in Base Setting

Scenario 1

Introduction

As a participant in this experiment, you will make a hypothetical decision about whether to get genetic testing. There is a small chance that this test will result in a large benefit – for instance, you could learn that you are at high risk for a certain health problem that can be prevented with a simple intervention. There is also a small chance that the genetic test could result in a large harm – for example, your genetic data could get hacked and publicly exposed and, as a consequence, you may have trouble finding a job and end up broke.

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with $6, the first $4 of which represents your current wealth. The other $2 represents the cost of genetic testing (either actual financial cost, or time required). You need to decide whether to get genetic testing. If you decide to do so, you must give up the $2.

You have a 5% chance of receiving $60 (gigantic relative to your wealth), which represents your chance of receiving a substantial benefit from the genetic test in the form of a life-saving treatment. That is, you have one chance out of 20 to receive a return of $60. The following table summarizes this.

Table 1: Chance of Benefit

Amount You Invest	Chance you receive $60
$0	0%
$2	5%

Imagine that you decide to get genetic testing and pay $2. As a result, your genetic data (and your $4 wealth in this experiment) is at risk. In this case, you will incur a chance of 25% of losing your remaining $4 and ending up broke.

Will you get genetic testing? Yes or No.

Figure C.2: Instructions given to Patients in Altruism Setting

Scenario 2: The following scenario is similar to the one above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will be paired with another participant who will play the role of your physician. Suppose that your physician suggests that you get genetic testing. There is a small chance that this test will result in a large benefit – for instance, you could learn that you are at high risk for a certain health problem that can be prevented with a simple intervention. The test is beneficial to the physician too because he or she can use the resulting data for research.

There is also a small chance that the genetic test could result in a large harm – for example, your genetic data could get hacked and publicly exposed and, as a consequence, you may have trouble finding a job and end up broke.

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with $6, the first $4 of which represents your current wealth. The other $2 represents the cost of genetic testing (either actual financial cost, or time required). You need to decide whether to get genetic testing. If you decide to do so, you must give the physician the $2. The physician, in turn, receives another $2 from a funding agency for a total of $4 – this captures the fact that the physician gets some additional benefit from being able to use your genetic for research purposes.

You have a 5% chance of receiving $60 (gigantic relative to your wealth), which represents your chance of receiving a substantial benefit from the genetic test in the form of a life-saving treatment. That is, you have one chance out of 20 to receive a return of $60. The following table summarizes this.

Table 1: Chance of Benefit

Amount You Invest	Chance you receive $60
$0	0%
$2	5%

Imagine that you decide to get genetic testing and pay $2. As a result, your genetic data (and your $4 wealth in this experiment) is at risk. In this case, you will incur a chance of 25% of losing your remaining $4 and ending up broke.

Will you get genetic testing? Yes or No

Figure C.3: Instructions given to Patients in Trust Setting

Scenario 3: The following scenario is similar to the one above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will be paired with another participant who will play the role of your physician. Suppose that your physician suggests that you get genetic testing. There is a small chance that this test will result in a large benefit – for instance, you could learn that you are at high risk for a certain health problem that can be prevented with a simple intervention. The test is beneficial to the physician too because he or she can use the resulting data for research.

There is also a small chance that the genetic test could result in a large harm – for example, your genetic data could get hacked and publicly exposed and, as a consequence, you may have trouble finding a job and end up broke. The physician can reduce the chance that your medical record is exposed by investing in information security (better encryption, better firewalls, and so on).

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with $6, the first $4 of which represents your current wealth. The other $2 represents the cost of genetic testing (either actual financial cost, or time required). You need to decide whether to get genetic testing. If you decide to do so, you must give the physician the $2. The physician, in turn, receives another $2 from a funding agency for a total of $4 – this captures the fact that the physician gets some additional benefit from being able to use your genetic for research purposes.

You have a 5% chance of receiving $60 (gigantic relative to your wealth), which represents your chance of receiving a substantial benefit from the genetic test in the form of a life-saving treatment. That is, you have one chance out of 20 to receive a return of $60. The following table summarizes this.

Table 1: Chance of Benefit

Amount you invest	**Chance you earn $60**
$0	0%
$2	5%

Imagine that you decide to get genetic testing and pay $2. As a result, your genetic data (and your $4 wealth in this experiment) is at risk. The extent of the risk depends on how much money the other participant (your physician) decides to spend on security to protect your genetic data. The following table shows exactly how much different amounts of security investment reduce the risk to your $4 wealth.

Table 2: Chance of Loss

Amount the physician spends on security	**Chance you lose your additional $4**
$0.00	25%
$2.00	15%
$4.00	0%

Will you get genetic testing? Yes or No

Figure C.4: Instructions given to Investors in Base Setting

Scenario 1

Introduction

As a participant in this experiment, you will play the role of an investor and decide whether you want to invest money with your money manager. This investment might result in a large return, but could also result in a large loss – for example, your financial data could get hacked and, as a consequence, you may lose all your assets and end up broke.

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with \$6, the first \$4 represents your (non-liquid) wealth and the other \$2 represents your liquid assets that you may invest. You need to decide whether to invest your liquid assets with your money manager. If you decide to do so, you must give up the \$2.

You have a 5% chance of receiving \$60 (gigantic relative to your wealth), which represents the case that the investment is successful. If the investment is unsuccessful, you will receive nothing. That is, you have one chance out of 20 to receive a return of \$60. The following table summarizes this.

Table 1: Chance of Return on Investment

Amount You Invest	**Chance you will receive \$60**
\$0	0%
\$2	5%

Imagine that you decide to invest your \$2 of liquid assets. As a result, your financial data (and your \$4 wealth in this experiment) is at risk. In this case, you will incur a risk of 25% of losing your remaining \$4.

Will you invest your \$2 of liquid assets? Yes or No.

Figure C.5: Instructions given to Investors in Altruism Setting

Scenario 2: The following scenario is similar to the one above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will play the role of an investor and will be paired with another participant who will play the role of your money manager. Suppose that your money manager suggests that you invest with him. This investment might result in a large return, and is beneficial to the money manager too, who makes money when people invest with him or her.

There is a small chance that the investment could also result in a large loss – for example, your financial data could get hacked and, as a consequence, you may lose all your assets and end up broke.

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with $6, the first $4 represents your (non-liquid) wealth and the other $2 represents your liquid assets that you may invest. You need to decide whether to invest your liquid assets with your money manager. If you decide to do so, you must give the money manager the $2. The money manager, in turn, receives an additional $2 for a total of $4 – this captures the fact that the money manager gets some additional benefit from your decision to invest.

You have a 5% chance of receiving $60 (gigantic relative to your wealth), which represents the case that the investment is successful. If the investment is unsuccessful, you will receive nothing. That is, you have one chance out of 20 to receive a return of $60. The following table summarizes this.

Table 1: Chance of Return on Investment

Amount You Invest	**Chance you will receive $60**
$0	0
$2	5%

Imagine that you decide to invest your $2 of liquid assets. As a result, your financial data (and your $4 wealth in this experiment) is at risk. In this case, you will incur a risk of 25% of losing your remaining $4.

Will you invest your $2 of liquid assets? Yes or No.

Figure C.6: Instructions given to Investors in Trust Setting

Scenario 3: The following scenario is similar to the ones above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will play the role of an investor and will be paired with another participant who will play the role of your money manager. Suppose that your money manager suggests that you invest with him. This investment might result in a large return, and is beneficial to the money manager too, who makes money when people invest with him or her.

There is a small chance that the investment could also result in a large loss – for example, your financial data could get hacked and, as a consequence, you may lose all your assets and end up broke. The money manager can reduce the chance that your financial information is exposed by investing in information security (better encryption, better firewalls, and so on).

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with $6, the first $4 represents your (non-liquid) wealth and the other $2 represents your liquid assets that you may invest. You need to decide whether to invest your liquid assets with your money manager. If you decide to do so, you must give the money manager the $2. The money manager, in turn, receives an additional $2 for a total of $4 – this captures the fact that the money manager gets some additional benefit from your decision to invest.

You have a 5% chance of receiving $60 (gigantic relative to your wealth), which represents the case that the investment is successful. If the investment is unsuccessful, you will receive nothing. That is, you have one chance out of 20 to receive a return of $60. The following table summarizes this.

Table 1: Chance of Return on Investment

Amount You Invest	Chance you will receive $60
$0	0
$2	5%

Imagine that you decide to invest your $2 of liquid assets. As a result, your financial data (and your $4 wealth in this experiment) is at risk. The extent of the risk depends on how much money the other participant (your money manager) decides to spend on security to protect your financial data. The following table shows exactly how much different amounts of security investment reduce the risk to your $4 wealth.

Table 2: Chance of Loss

Amount money manager spends on security	Chance you lose your wealth
$0.00	25%
$2.00	15%
$4.00	0%

Will you invest your $2 of liquid assets? Yes or No.

Figure C.7: Instructions given to Physicians in Recipient-Altruism Setting

Scenario 1

Introduction

As a participant in this experiment, you will play the role of a physician and will be paired with another participant who will play the role of your patient. In this experiment, we simulate a situation where a patient potentially faces a large loss. You have records of this patient's genetic data and now there is a small chance that the patient's genetic data could get hacked and publicly exposed, and, as a consequence, the patient may have trouble finding a job and end up broke. As a physician, you can reduce the chance that the patient's genetic data is exposed by spending on information security (better encryption, better firewalls, and so on).

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with $4. We also give a patient of yours $4, which represents the patient's wealth. Due to a potential data breach at your facility, the patient's genetic data is at risk. As a result, the patient incurs some risk of losing his or her wealth of $4. The extent of the risk depends on how much money you decide to spend on security to protect his or her data. The following table shows exactly how different amounts of security investment reduce the risk to your patient's $4 wealth.

Effect of your spending on security

Amount You Spend	Chance Patient Loses $4 Wealth
$0.00	25%
$2.00	15%
$4.00	0%

The patient's wealth might face some risk even after your spending, unless you spend the entire $4. You get to keep anything you do not spend on security.

How much do you choose to spend on security?

Figure C.8: Instructions given to Physicians in Reciprocity Setting

Scenario 2: The following scenario is similar to the one above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will play the role of a physician and will be paired with another participant who will play the role of your patient. In this experiment, we simulate a situation where a patient decides whether to get a genetic test from you.

There is a small chance that the genetic test will result in a large benefit to the patient – for instance, the patient may learn that he or she has a high risk for a certain health problem that can be prevented with a simple intervention. This test is also beneficial to you because you can use the resulting data for research.

There is also a small chance that the genetic test could potentially result in a large loss to the patient – the patient's genetic data could get hacked and publicly exposed, and, as a consequence, the patient may have trouble finding a job and end up broke. As a physician, you can reduce the chance that the patient's genetic data is exposed by spending on information security (better encryption, better firewalls, and so on).

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We gave your patient $6, the first $4 of which represents his or her wealth. The other $2 represents the cost of a genetic test (either actual financial cost, or time required). Your patient decided to get genetic testing and gave you $2. In turn, we are giving you an additional $2 for a total of $4 – this captures the fact that you get some additional benefit from being able to use the patient's genetic data for research purposes.

The patient has a 5% chance of receiving $60 (gigantic relative to her wealth), which represents the patient's chance of receiving a substantial benefit from the genetic test in the form of a life-saving treatment. As a result of the genetic test, the patient's genetic data (and her money in this experiment) is now at risk, and the patient incurs some risk of losing his or her wealth of $4. The extent of the risk depends on how much money you choose to spend on security to protect the patient's genetic data. The following table shows exactly how different amounts of security investment reduce the risk to your patient's $4 wealth.

Effect of your spending on security

Amount You Spend	Chance Patient Loses His or Her Wealth
$0.00	25%
$2.00	15%
$4.00	0%

The patient's wealth might face some risk even after your spending, unless you spend the entire $4. You get to keep anything you do not spend on security.

How much do you choose to spend on security?

Figure C.9: Instructions given to Physicians in Altruism Setting

Scenario 3: The following scenario is similar to the ones above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will play the role of a physician and will be paired with another participant who will play the role of your patient. In this experiment, we simulate a situation where a patient decides whether to get a genetic test from you. This test is beneficial to you because you can use the resulting data for research.

Experiment

We gave your patient $6, the first $4 of which represents the patient's wealth. The other $2 represents the cost of a genetic test (either actual financial cost, or time required). If your patient decides to get genetic testing, he or she will give you this $2. In turn, we will give you an additional $2 for a total of $4 – this captures the fact that you get some additional benefit from being able to use the patient's genetic data for research purposes.

You do not have a decision to make in this experiment. You will see the decision at the end of this experiment session.

Figure C.10: Instructions given to Money Managers in Recipient-Altruism Setting

Scenario 1

Introduction

As a participant in this experiment, you will play the role of a money manager and will be paired with another participant who will play the role of an investor. In this experiment, we simulate a situation where an investor's wealth is at risk. This investor has invested money with you and now there is a small chance that the investor's financial data could get hacked and, as a consequence, the investor may lose all of his or her assets and end up broke. As a money manager, you can reduce the chance that the investor's financial data is exposed by spending on information security (better encryption, better firewalls, and so on).

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We are starting you off with \$4. We also gave an investor \$4, which represents his or her wealth. Due to a potential data breach at your facility, the investor's financial data is at risk. As a result, the investor incurs some risk of losing his or her \$4 wealth. The extent of the risk depends on how much money you decide to spend on security to protect his or her financial data. The following table shows exactly how much different amounts of security investment reduce the risk.

Effect of your spending in security

Amount You Spend	**Chance Investor Loses \$4 Wealth**
\$0.00	25%
\$2.00	15%
\$4.00	0%

The investor's wealth might face some risk even after your spending, unless you spend the entire \$4. You get to keep anything you do not spend on security.

How much do you choose to spend for the security?

Figure C.11: Instructions given to Money Managers in Reciprocity Setting

Scenario 2: The following scenario is similar to the one above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will play the role of a money manager and will be paired with another participant who will play the role of an investor. In this experiment, we simulate a situation where an investor decides whether to invest his or her money with you.

There is a small chance that the investment will result in a large return to the investor. The investment is also very beneficial to you because you earn money when people invest with you.

There is also a small chance that the investment could potentially result in a large loss to the investor – the investor's financial data could get hacked and, as a consequence, the investor may lose all of his or her assets and end up broke. As a money manager, you can reduce the chance that the investor's financial data is exposed by spending on information security (better encryption, better firewalls, and so on).

In the experiment below, we simulate the benefits and losses associated with this scenario with monetary payments. The figures in the experiment are small, but they are meant to represent significant possibilities of gain or loss.

Experiment

We have given an investor \$6, the first \$4 of which represents the investor's wealth. The other \$2 represents the investor's liquid assets that he or she may invest. This investor chose to invest with you and gave you their \$2. In turn, we are giving you an additional \$2 for a total of \$4 – this captures the fact that you make some additional money because of the investor's decision to invest with you.

The investor has a 5% chance of receiving \$60, which represents the event that the investment is successful. As a result of the investment, the investor's financial data (and his or her money in this experiment) is now at risk, and the investor incurs some risk of losing his or her remaining \$4. The extent of the risk depends on how much money you choose to spend on security to protect the investor's financial data. The following table shows exactly how much different amounts of security spending reduces the risk to the investor's \$4 wealth.

Effect of your spending on security

Amount You Spend	**Chance Investor Loses \$4 Wealth**
\$0.00	25%
\$2.00	15%
\$4.00	0%

The investor's wealth might face some risk even after your spending, unless you spend the entire \$4. You get to keep anything you do not spend on security.

How much do you choose to spend towards security?

Figure C.12: Instructions given to Money Managers in Altruism Setting

Scenario 3: The following scenario is similar to the one above. The changes between this scenario and the previous one are highlighted in red. Please answer the question at the end after accounting for these changes. We encourage you to read the entire passage.

Introduction

As a participant in this experiment, you will play the role of a money manager and will be paired with another participant who will play the role of an investor. In this experiment, we simulate a situation where an investor decides whether to invest his or her money with you. The investment might result in a large return to the investor, and is also very beneficial to you because you earn money when people invest with you.

Experiment

We have given an investor \$6, the first \$4 of which represents the investor's wealth. The other \$2 represents the investor's liquid assets that he or she may invest. If the investor chooses to invest with you, you will receive their \$2. In turn, we will give you an additional \$2 for a total of \$4 – this captures the fact that you make some additional money because of the investor's decision to invest with you. If the investor does not invest with you, you will receive \$0.

You do not have a decision to make in this experiment. You will see the investor's decision at the end of this experiment session.

Table C.1 below is an augmented version of Table 1 of the main paper. The coefficient of *Genetic Frame* is statistically significant in all specifications, indicating that there is a framing effect in the base setting. Specifically, coefficient of 0.254 in the second column indicates that the proportion of subjects who share data in the Genetic Frame is about 25 percentage points greater than the proportion of subjects who share data in the Investment Frame. The Probit coefficient of 0.675 tells us the difference between the coefficient z-score obtained in the Genetic Frame and the z-score in the investment frame. Although this does not have a straightforward interpretation, we can use this value to perform a marginal analysis (results of which are not shown but can be provided upon request). The marginal analysis results tell us that patients choose to undergo a genetic test with a probability of 0.686 while investors choose to make an investment with a probability of 0.425, holding all variables at their means. These two values result in a difference of 26.1 percentage points, which suggests that patients are about 26.1 percentage points more likely to get a genetic test than investors are likely to make an investment. In the last column, the results present the odds ratios from a Logit specification. The coefficient of 2.997 suggests that the odds of a patient getting a genetic test (which are 28/11) are about 3 times greater than the odds of an investor making an investment (which are 19/23).

Table C.1: Framing effects for data sharers in the base setting

Variable	**OLS**	**OLS**	**Probit**	**Logit O/R**
Genetic	0.266**	0.254**	0.675**	2.997**
	(0.107)	(0.125)	(0.325)	(1.607)
failed control questions		0.0243	0.0768	1.119
		(0.124)	(0.327)	(0.600)
Age		-0.0389	-0.111	0.836
		(0.0403)	(0.108)	(0.148)
Victim of ID theft		0.174	0.493	2.134
		(0.298)	(0.828)	(2.796)
% Female		0.0522	0.123	1.256
		(0.123)	(0.321)	(0.664)
% White		0.169	0.466	2.126
		(0.127)	(0.335)	(1.163)
% Hispanic		-0.216	-0.584	0.380
		(0.280)	(0.749)	(0.484)
Religiousness (0-3 scale)		-0.0101	-0.0302	0.957
		(0.0546)	(0.144)	(0.222)
Constant	0.452***	1.089	1.716	15.65

Table C.2 presents regression results for decisions made by data sharers in the altruism setting. The coefficient of *Genetic Frame* is not statistically significant, which means we cannot say there is a framing effect for this setting, that is, the proportion of subjects who share data in the Genetic Frame is not significantly different from the proportion of subjects who share data in the Investment Frame.

Table C.2: Framing effects for data sharers in the altruism setting

Variable	**OLS**	**OLS**	**Probit**	**Logit O/R**
Genetic Frame	0.145	0.120	0.325	1.695
	(0.108)	(0.127)	(0.322)	(0.889)
failed control questions		0.0142	0.0283	1.064
		(0.127)	(0.325)	(0.560)
Age		-0.0171	-0.0432	0.928
		(0.0410)	(0.104)	(0.158)
Victim of ID theft		0.0840	0.238	1.421
		(0.303)	(0.800)	(1.818)
% Female		-0.138	-0.372	0.548
		(0.125)	(0.316)	(0.282)
% White		0.0866	0.241	1.484
		(0.129)	(0.331)	(0.800)
% Hispanic		-0.0900	-0.245	0.673
		(0.285)	(0.708)	(0.763)
Religiousness (0-3 scale)		0.0287	0.0781	1.137
		(0.0556)	(0.141)	(0.262)
Constant	0.548***	0.858	0.901	4.697
	(0.0751)	(0.815)	(2.076)	(15.88)
N	81	73	73	73
R^2	0.072	0.120		

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

Table C.3 below presents regression results for decisions made by data sharers in the trust setting. The coefficient of *Genetic Frame* is not statistically significant, which means we cannot say there is a framing effect for this setting. That is, in the trust setting, the proportion of subjects who share data in the Genetic Frame is not significantly different from the the proportion of subjects who share data in the Investment Frame.

Table C.3: Framing effects for data sharers in the trust setting

Variable	**OLS**	**OLS**	**Probit**	**Logit O/R**
Genetic Frame	0.0806	0.0979	0.327	1.696
	(0.0967)	(0.116)	(0.356)	(1.006)
failed control questions		-0.0482	-0.167	0.745
		(0.115)	(0.345)	(0.433)
Age		0.0405	0.154	1.287
		(0.0373)	(0.128)	(0.281)
Victim of ID theft		-0.0369	-0.154	0.814
		(0.276)	(0.752)	(1.048)
% Female		0.118	0.404	1.917
		(0.114)	(0.350)	(1.125)
% White		-0.0367	-0.120	0.818
		(0.118)	(0.356)	(0.487)
% Hispanic		-0.186	-0.525	0.423
		(0.260)	(0.742)	(0.501)
Religiousness (0-3 scale)		0.0108	0.0301	1.059
		(0.0506)	(0.151)	(0.267)
Constant	0.714***	-0.134	-2.597	0.0143
	(0.0671)	(0.742)	(2.521)	(0.0615)
N	81	73	73	73
R^2	0.072	0.120		

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

Table C.4 presents regression results for decisions made by data recipients in the recipient-altruism setting. The coefficient of *Genetic Frame* is not statistically significant, which means we cannot say there is a framing effect for this setting. That is, the proportion of physicians who protect a patient's genetic data is not significantly different from the the proportion of money managers who protect an investor's financial data.

Table C.4: Framing effects for data recipients in the recipient-altruism setting

Variable	**OLS**	**OLS**	**Probit**
Genetic Frame	-0.0348	0.00952	0.0340
	(0.112)	(0.126)	(0.321)
failed control questions		-0.119	-0.310
		(0.187)	(0.478)
Age		0.00713	0.0171
		(0.0272)	(0.0708)
Victim of ID theft		0.103	0.312
		(0.246)	(0.642)
% Female		0.225*	0.591*
		(0.120)	(0.302)
% White		-0.0836	-0.236
		(0.126)	(0.320)
% Hispanic		-0.108	-0.286
		(0.177)	(0.443)
Religiousness (0-3 scale)		0.0474	0.128
		(0.0626)	(0.158)
Constant	0.548***	0.295	-0.504
	(0.0779)	(0.565)	(1.460)
N	81	75	75
R^2	0.001	0.080	

Standard errors in parentheses

* $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$

Table C.5 presents regression results for decisions made by data recipients in the reciprocity setting. The coefficient of *Genetic Frame* is not statistically significant, which means we cannot say there is a framing effect for this setting. That is, the proportion of physicians who protect a patient's genetic data is not significantly different from the the proportion of money managers who protect an investor's financial data.

Table C.5: Framing effects for data recipients in the reciprocity setting

Variable	**OLS**	**OLS**	**Probit**
Genetic Frame	0.0129	0.0594	0.238
	(0.127)	(0.137)	(0.393)
failed control questions		-0.116	-0.435
		(0.204)	(0.547)
Age		-0.0613	-0.204*
		(0.0401)	(0.118)
Victim of ID theft		-0.149	-0.448
		(0.238)	(0.639)
% Female		0.214	0.688*
		(0.130)	(0.383)
% White		-0.0784	-0.287
		(0.141)	(0.407)
% Hispanic		-0.0323	-0.0690
		(0.188)	(0.523)
Religiousness (0-3 scale)		-0.0966	-0.312
		(0.0668)	(0.195)
Constant	0.600***	1.920**	4.684*
	(0.0907)	(0.804)	(2.403)
N	61	56	56
R^2	0.000	0.173	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tables C.6a and C.6b show the contribution of different motivations for data sharers to share data in each frame. The first table (panel a) shows the contribution of each motivation in the Genetic Frame and the second table (panel b) shows the contribution of each motivation in the Investment Frame. The *Motivation* column shows the motivation being tested in that setting. In the base setting, the only reason/motivation for subjects to share data is personal benefit. In the altruism setting, there are two motivations for subjects to share data: personal benefit and altruism. Any *additional* data sharing, relative to the base setting, must be due to altruism. In the trust setting, there are three motivations to share data: personal benefit, altruism, and trust. Any *additional* data sharing - relative to the altruism setting - must be due to trust, because that is the only difference between the trust setting and the altruism setting. The column *Pct who share data* shows the proportion of subjects who choose to share data in that setting. The column *Marginal increase* provides the contribution of the motivation listed in the Motivation column. This is calculated by subtracting the proportion of subjects who share data in the previous setting from the the proportion of subjects who share data in this setting. For example, In the Investment Frame (panel b), 71.4% of subjects share data in the trust setting while 54.8% of subjects share data in the altruism setting. The difference between these two, 16.6%, represents the proportion of subjects who are motivated by trust but not by personal benefit or altruism. The last column shows the p-value (and hence significance) of the corresponding marginal increase. For example, the marginal increase of 16.6% is statistically significant at the 5% level. The p-value of 0.018 is the resulting p-value after conducting a t-test that compares the two means, 54.8% and 71.4%. Note that we use a paired t-test which is more appropriate in this situation because each subject participates in all three settings. For the base setting, where subjects have only one motivating factor, we compare the proportion of subjects who share data (aka mean value) with 0. The results show that personal benefit is a significant motivator to share data in both frames while trust is a significant motivator in only the Investment Frame.

Table C.6: Differences across settings for data sharers

(a) Genetic Frame

Setting name	**Motivation**	**Pct who share data**	**Marginal increase**	***p*-value**
Base	personal benefit	71.8%	71.8%	0.000
Altruism	altruism	69.2%	-2.6%	0.661
Trust	trust	79.5%	10.3%	0.324

(b) Investment Frame

Setting name	**Motivation**	**Pct who share data**	**Marginal increase**	***p*-value**
Base	personal benefit	45.2%	45.2%	0.000
Altruism	altruism	54.8%	9.6%	0.210
Trust	trust	71.4%	16.6%	0.018

Table C.7 shows the contribution of different motivations for *data recipients* to protect another person's data. The first table (panel a) shows the contribution of each motivation in the Genetic Frame and the second table (panel b) shows the contribution of each motivation in the Investment Frame. The *Motivation* column shows the motivation being tested in that setting. In the recipient-altruism setting, the only motivation for protecting another person's data is altruism. In the reciprocity setting, there are two motivations for protecting another person's data: altruism and reciprocity. Any *additional* data sharing, relative to the recipient-altruism setting, must be due to reciprocity. The column *Pct who protect data* shows the proportion of subjects who choose to protect the data sharer's data in that setting. The column *Marginal increase* provides the contribution of the motivation listed in the Motivation column, which is calculated by subtracting the proportion of subjects who protect data in the previous setting from the the proportion of subjects who share data in this setting. For example, In the Investment Frame (panel b), 60% of subjects protect data in the reciprocity setting while 54.8% of subjects protect data in the recipient-altruism setting. The difference between these two, 5.2%, represents the proportion of subjects who are motivated by reciprocity only. The last column shows the p-value (and hence significance) of the corresponding marginal increase. For example, the marginal increase of 5.2% has a p-value of 0.326 and is not statistically significant. These p-values are determined by a means-comparison t-test. In the recipient-altruism setting, where subjects are motivated by altruism only, the t-test tests whether the proportion of subjects who protect data is significantly greater than zero. Note that some data recipients (specifically, those who are paired with data sharers who do not share data) do not have a decision to make in the reciprocity setting; instead, they make a decision in the recipient-altruism setting only. The t-test that compares the proportion of subjects in the recipient-altruism setting, and the proportion of subjects who protect data in the reciprocity does not include these subjects. The results show that, in both frames, altruism is a significant motivator for protecting data while reciprocity is not. Nevertheless, reciprocity has a considerably lower p-value in the Genetic Frame than in Investment Frame, suggesting that reciprocity is a still a relatively stronger motivator for physicians than it is for money managers.

Table C.7: Differences across settings for data recipients

(a) Genetic Frame

Setting name	**Motivation**	**Pct who protect data**	**Marginal increase**	***p*-value**
Recipient Altruism	altruism	51.3%	51.3%	0.000
Reciprocity	reciprocity	61.3%	10.0%	0.103

(b) Investment Frame

Setting name	**Motivation**	**Pct who protect data**	**Marginal increase**	***p*-value**
Recipient Altruism	altruism	54.8%	54.8%	0.000
Reciprocity	reciprocity	60.0%	5.2%	0.326

Tables C.8, C.9, and C.10 show regression results from the online replication of the experiment. These are presented in a similar way as the results for the lab experiment.

Table C.8: Framing effects in base setting with AMT subjects

Variables	**OLS**	**OLS**	**Probit**
genetic	0.156**	0.168***	0.470***
	(0.0620)	(0.0643)	(0.179)
control failed		0.0953	0.265
		(0.0691)	(0.189)
age		0.00382	0.0111
		(0.00344)	(0.00964)
victim		0.00301	0.00481
		(0.0814)	(0.223)
female		-0.0433	-0.130
		(0.0665)	(0.185)
white		0.00338	0.0151
		(0.0868)	(0.240)
hispanic		0.156	0.417
		(0.135)	(0.366)
religious		0.0493	0.136
		(0.0317)	(0.0872)
Constant	0.294***	0.0757	-1.184***
	(0.0439)	(0.146)	(0.415)
N	239	230	230
R^2	0.026	0.066	

Standard errors in parentheses

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table C.9: Framing effects in altruism setting with AMT subjects

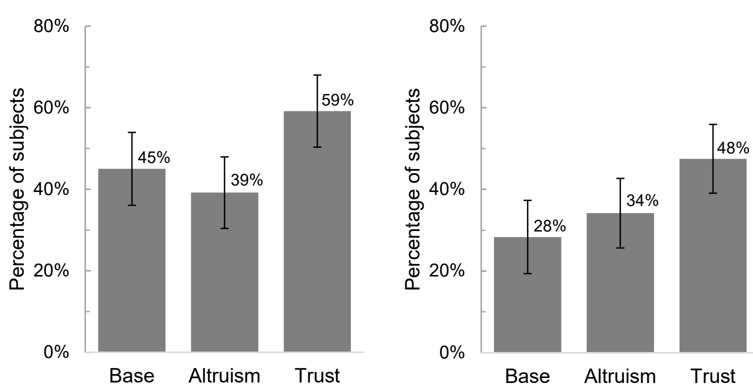
Variables	**OLS**	**OLS**	**Probit**
Genetic Frame	0.0555	0.0791	0.221
	(0.0624)	(0.0650)	(0.177)
failed control questions		0.110	0.297
		(0.0698)	(0.188)
Age		0.000987	0.00287
		(0.00348)	(0.00946)
Victim of ID theft		0.0136	0.0339
		(0.0822)	(0.222)
% Female		-0.0590	-0.161
		(0.0672)	(0.183)
% White		0.0309	0.0911
		(0.0877)	(0.239)
% Hispanic		-0.0115	-0.0275
		(0.137)	(0.364)
Religiousness (0-3 scale)		0.0733**	0.196**
		(0.0320)	(0.0856)
Constant	0.336***	0.190	-0.846**
	(0.0442)	(0.148)	(0.407)
N	239	230	230
R^2	0.003	0.042	

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure C.13: Choices made by data sharers in online version of experiment

(a) Genetic Frame (b) Investment Frame



Compare this with Figure 2 in the main paper, which shows choices made by data sharers in the lab version of the experiment.

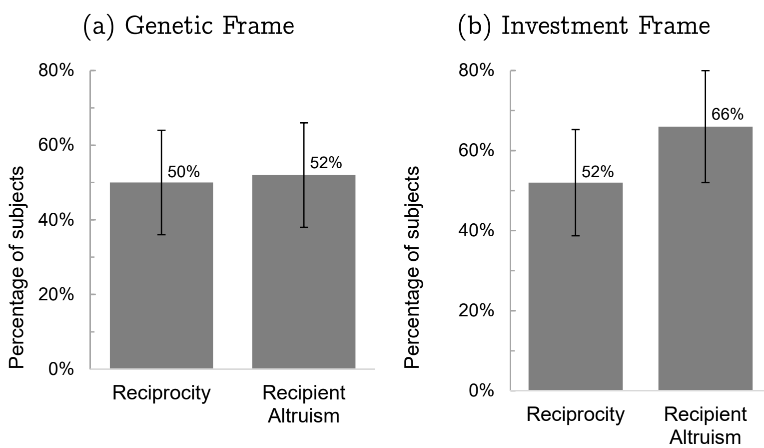
Table C.10: Framing effects in trust setting with AMT subjects

Variables	**OLS**	**OLS**	**Probit**
Genetic Frame	0.121*	0.106	0.296*
	(0.0644)	(0.0655)	(0.175)
failed control questions		0.00405	0.0115
		(0.0703)	(0.189)
Age		0.00985***	0.0284***
		(0.00351)	(0.0101)
Victim of ID theft		0.148*	0.389*
		(0.0829)	(0.225)
% Female		0.0203	0.0487
		(0.0678)	(0.182)
% White		-0.0815	-0.230
		(0.0884)	(0.237)
% Hispanic		0.206	0.604
		(0.138)	(0.388)
Religiousness (0-3 scale)		0.0511	0.142
		(0.0323)	(0.0872)
Constant	0.471***	0.0850	-1.178***
	(0.0456)	(0.149)	(0.421)
N	239	230	230
R^2	0.015	0.100	

Standard errors in parentheses

* $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$

Figure C.14: Choices made by data recipients in online version of experiment.



Compare this with Figure 3 in the main paper, which shows choices made by data recipients in the lab version of the experiment.