Discussion Paper Series

RIEB

Kobe University

DP2020-23

Risks on Others

Takaaki HAMADA
Tomohiro HARA

Revised September 5, 2022

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Risks on Others*

Takaaki Hamada† Tomohiro Hara‡

August 20, 2022

Abstract

We investigate other regarding preferences when others are involved in some risks. We call such risks "risks on others (ROO)." We design a novel experiment to capture ROO, and also theoretically define risk attitudes toward others under two dimensions: an *absolute* term, which characterizes one's risk attitudes toward others without comparing with other individuals, and a *relative* term, which compares these risk attitudes between individuals. From our experiment, we find that decision makers exhibit robust risk-averse behaviors, which contradicts the most representative linear inequality aversion models by Fehr and Schmidt (1999). Utilizing our experimental setting, which enables us to compare risk attitudes toward others between subjects, we also investigate the effect of sources of ROO on other regarding behaviors.

Keywords: Other-regarding preferences, risk preferences, inequality aversion

JEL classification: C91, D63, D64, D81, D91

*This study was funded through a JSPS (Japan Society for Promotion of Science) research grant and was also supported by the IRBs of the University of Tokyo and the University of Maryland, College Park. We would like to thank Emel Filiz-Ozbay, Erkut Ozbay, Kota Saito, Akihiko Matsui, and the seminar participants at the Kansai Game Theory Workshop, Pre-experiment Workshop at the Association of Behavioral Economics and Finance in Japan, and 23rd Experimental Social Science Conference at Meiji Gakuin University for their helpful suggestions. We are grateful to Masashi Takahashi for excellent research assistance. All remaining errors are our own.

†Faculty of Economics, Toyo university, 5-28-20, Hakusan, Bunkyo-ku, Tokyo 112-8606 Japan/Research Institute for Economics and Business Administration, Kobe University, 2-1 Rokkodai, Nada, Kobe 657-8501, Japan. Email: t.hamada.econ@gmail.com

‡Department of Economics, University of Maryland, College Park. Email: tomohara@umd.edu.

1 Introduction

Suppose you are asked to donate some cash to poor people through an unknown charitable organization. Even if you are a nice person and willing to contribute to the poor, you may become skeptical about the unknown organization, thinking "does this organization actually give my donation to poor people?" It is well-documented in the economic and psychological literature that individuals are likely unwilling to give to unfavorable or ineffective groups. For instance, Gneezy, Keenan and Gneezy (2014) document that people are less likely to donate when they find that their donations are used for administrative works rather than directly helping the poor, and Karlan and Wood (2016) find that people prefer to give more money to a charitable organization if they know the organization is more effective. These studies focus on the situation where people are aware of the state of the organizations that they are contemplating donation to. However, in reality, people cannot be perfectly sure about how their donations are used or how effective these organizations are.

This paper focuses on such uncertainties attached to others. We call them "risks on others" (hereafter ROO), where people cannot observe others' true states but do know probability distributions of states. We provide a novel experiment and theoretical framework to capture ROO. ROO appears in a wide range of decisions. For example, a college professor may not want to supervise a student when the professor is skeptical about the student's performance. People may put less effort toward volunteering activities when they have suspicions that they are working for those who are not suffering. Further, we consider a wide range of racial/ethnic/gender discrimination in the labor market as an issue related to ROO, because unobservable productivity correlated to employees' group affiliations can be considered a source of ROO.[1] Therefore, investigating ROO contributes to understanding a wide range of decision problems.

Research on other-regarding preferences so far does not provide a clear view on how to incorporate such uncertainty. Although some papers have investigated the relationship between charitable giving and risk (e.g., Brock, Lange and Ozbay 2013; Saito 2013; López-Vargas 2014; Rau and Müller

[1]When an employer does not observe full information on an employee's productivity, they infer the statistical properties of the productivity from the employee's group identities (e.g., gender and ethnicity), and treat employees differently by group. This is a classical story of statistical discrimination, which is well-documented in labor economics research (see surveys such as Neumark 2018). It is well-known that statistical discrimination and taste-based discrimination are closely associated (e.g., Bertrand and Mullainathan 2004), and the latter is considered a pattern of other-regarding preferences (Chen and Li 2009). Concepts in our paper can be related to the recently growing literature (e.g. Bohren, Haggag, Imas and Pope 2020) on the correlation between statistical (either accurate or inaccurate) and taste-based discrimination. Understanding ROO can highlight a potential mechanism bridging these two types of discrimination.

2017), most of them focus on the decision-maker's risk attitudes for their "own" payoffs, with little attention being paid to the risk attitudes for receivers' outcomes. Under existing studies, it is thus unclear how people perceive some risks attached to others.

To deeply understand ROO, the main scope and aim of this paper is to answer the following questions: How do we conceptualize and elicit risk attitude toward others' payoffs? How does a decision maker react to ROO? Are we able to capture behaviors toward ROO by existing models?

To this end, we develop a novel experiment to measure risk attitudes toward others with a theoretical framework. It is a modified version of a standard dictator game, summarized in Figure 1. We consider a dictator game with three players: a dictator (denoted as D in the figure) and two recipients that we call a "certain recipient" (denoted as C) and a "risky recipient" (denoted as R), with initial endowments denoted by e_D, e_C, and e_R for each. We have two treatment arms in our experiment, but at this point, let us consider the first treatment ("A: State-risk treatment").[2] In the example in Figure 1, the dictator initially has \$10 ($= e_D$) and confronts a decision problem on which recipient, C or R, they prefer to give $X = \$5$. Transfer X is fixed and the dictator cannot change the given amount. Lottery p is a safe lottery, where the dictator is going to give $X = \$5$ to a certain recipient who has \$0.5 ($=e_C$) as an initial endowment. In lottery q, the dictator is going to give the same amount to the risky recipient who has no endowment ($e_R = 0$) with probability 0.9, but has \$5 ($e_R = 5$) with probability 0.1.

Fixing all parameters other than e_C, we ask dictators to find an endowment e_C under which lottery p (giving to a certain recipient; C) and lottery q (giving to a risky recipient; R) are indifferent. We apply the Holt and Laury (2002) type of price list to elicit an indifferent point.[4] Our list starts with $e_C = 0$ and ends with $e_C = 5$, with some increments. As we will discuss in the theoretical part, we expect that people generally prefer to transfer to a certain recipient at the beginning of the list, and then switch to transferring to a risky recipient at some point under standard social preferences. We refer to this point as a "switching point" throughout the paper. It enables us to understand dictators' preferences between a risky individual and a non-risky one.

Along the structure of our experiment, we theoretically define and characterize the risk pref-

[2]The treatment B "Non-state-risk treatment" is explained later.

[3]Note that recipients C and R receive e_C and e_R as initial endowment, respectively, regardless of the dictator's choice. For example, in the Figure 1, the certain recipient receives $e_C = 0.5$ even if the dictator chooses lottery q.

[4]Holt and Laury (2002) change the probabilities in the list with other parameters constant, while our design keeps the probability constant and changes the initial endowment of the recipient C.

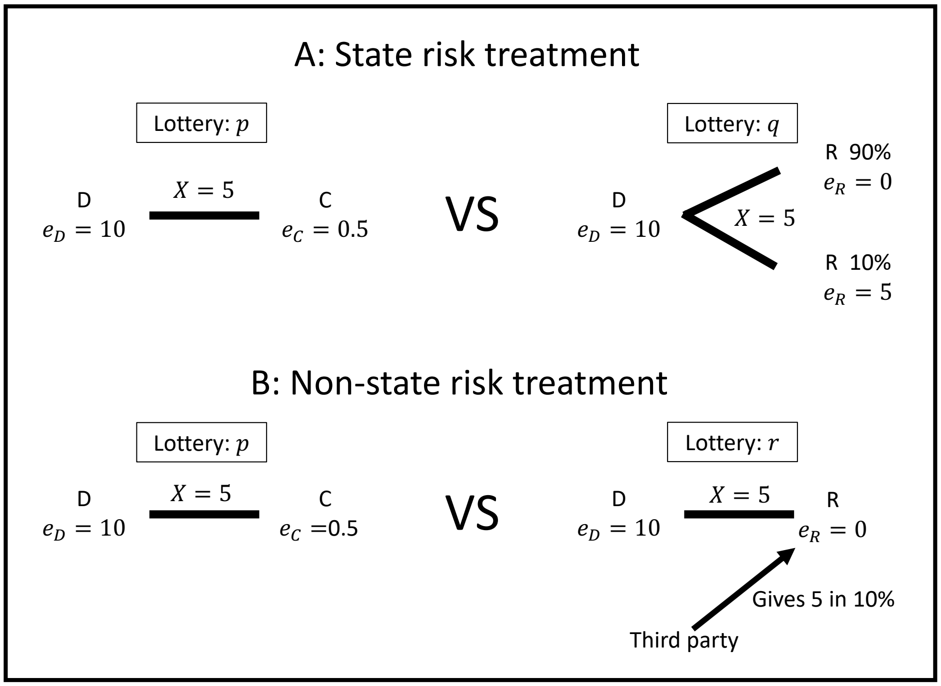


Figure 1: Simplified design of experiment [3]

erences toward others. First, we define decision-makers' risk attitudes in *absolute* terms (i.e., risk averse toward others, risk loving toward others, and risk neutral toward others) by comparing a switching point with the risky recipient's expected initial allocation.[5] Using the standard axiomatized models of inequality aversion (Fehr and Schmidt 1999; Saito 2013), we predict that dictators are risk loving toward others in five out of eight decisions in our experiment. On the other hand, in the actual experiment (conducted at the University of Maryland, College Park in 2018), we find strong risk-averse behaviors toward others. The results suggest that the well-used Fehr and Schmidt (1999)'s linear specification and its applications cannot capture people's decisions on allocations to others when recipients face risks.

Second, we define risk attitudes toward others in *relative* terms. Under our theoretical framework, the comparison of switching points enables us to compare the risk preferences toward others between individuals. This is possible under our novel experimental setting, in which multiple important parameters are well-controlled. Using this property, we compare subjects' risk attitudes by various aspects: gender, college major, race, and sources of risk.

We randomly assign subjects into two treatment groups to vary the source of the risk: the "state-

[5]For example, a decision maker is said to be risk averse toward others if the switching point is higher than the risky recipient's expected initial allocation (0.5 in the example of Figure 1). Roughly speaking, dictators who exhibit risk aversion toward others persist in giving to a certain recipient.

risk treatment" and the "non-state-risk treatment." Figure 1 illustrates these two treatments. In the "state-risk treatment," dictators are asked to compare lottery p to lottery q, while in the "non-state-risk treatment," dictators are asked to compare lottery p to lottery r. In lottery q (i.e., state-risk), the risky recipient has no endowment with probability 0.9 and \$5 with probability 0.1. This mimics a situation in which a seemingly poor person is actually rich with a small probability, which may be an unfavorable situation for a dictator. In lottery r (i.e., non-state-risk), the risky recipient has no initial endowment for sure, but has the opportunity to receive \$5 with probability 0.1 from a third party (the experimenter). In this second situation, the recipient is poor for sure, but has a small probability of receiving a good amount of transfer from a third party. For these two treatment arms, decision makers confront exactly the same level of ROO outcome-wise, but may behave differently because of the preferences regarding the source of ROO.

Our experimental results show that there is no difference in behaviors by different sources of risk on average overall. Yet, within male subjects and white subjects, people are more risk averse in the non-state-risk treatment than in the state-risk treatment. This means that our results for male and white subjects cannot be explained by the outcome-based theories. Moreover, male, STEM-major and white subjects exhibit more risk-averse behavior toward others than female and non-STEM subjects in general, indicating that risk preferences towards others are fundamentally different from standard risk preferences (that in general, women exhibit more risk aversion; see Eckel and Grossman 2008).

This paper makes three main contributions. First, our paper is the first work that explicitly defines and examines ROO. Second, we empirically find that people are consistently risk averse toward others. We show that the risk-averse behavior observed from the experiment is inconsistent with the linear form of Fehr and Schmidt (1999). As Fehr and Schmidt (1999) write, "some observations in dictator experiments suggest that there are a non-negligible fraction of people who exhibit nonlinear inequality aversion"(pp. 823), it is often said that the linear model fails to explain decision-makers' behaviors in some experiments.[6] However, existing experimental evidence so far is limited to the results under complete information in the payoffs, transfers, and others' states. To our knowledge, our result is the first clear evidence under a risky decision environment that suggests the limitation of linear inequality aversion. Third, we demonstrate the possibility that people behave differently

[6]For example, Bellemare, Kröger and Van Soest (2008) point out the limitation of linearity using ultimatum games to estimate inequality aversion, and provide an alternative form of the more flexible utility function allowing curvature.

by the source of the risk. Existing theories under risk (Brock et al. 2013; Saito 2013; López-Vargas 2014; Rau and Müller 2017) are outcome-based, meaning that sources of risk do not matter. While our results do not show strong evidence, we consider that the results for male and white subjects are suggestive evidence for the limitation of outcome-based theory. To consider policy implications boosting prosocial behaviors, it is important to elucidate how and why different sources of risk may induce different behavior.[7] We examine the potential behavioral mechanisms in our experiment from the viewpoint of *preference* and *bias* to inform future works.

The remainder of this paper is organized as follows. Section 2 describes the experimental design. In Section 3, we define risk attitudes toward others and provide the analytical framework for our experimental environment. Using our definitions and framework with existing models, Section 4 makes predictions on the dictator's behavior in our experiment. We also discuss and highlight the difference between our experiment and existing literature. Section 5 provides and interprets the results of our experiment. Section 6 discusses potential mechanisms that drive our results by modifying existing models, and Section 7 provides concluding remarks.

2 Experimental Design

The experiment has two treatment arms. For each treatment, the subjects completed eight tasks to elicit their preferences. After completing the main task, subjects solved another task for another study (not reported in this paper) and then answered simple demographic questions.

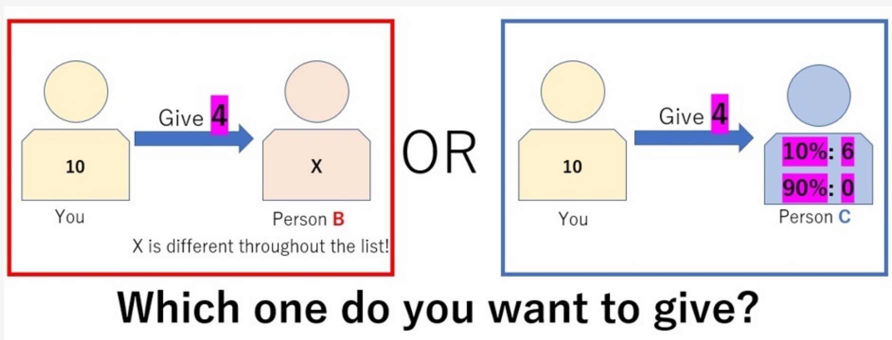
The main tasks are a modified version of a dictator game played by three subjects: one dictator and two recipients. The summarized design of the game is presented in the introduction (Figure 1). The dictator is always endowed with \$10. One of the recipients is always involved in uncertainty (we call this recipient a risky recipient, denoted by R throughout the paper and as person C in our experimental protocol), while the other is never involved in uncertainty (we call this recipient a certain recipient, denoted by C throughout the paper and as person B in our experimental protocol). We randomly assign subjects to one of the three roles in the beginning of the game without notifying them of their roles, and they play all games as if they were the dictator. At the end of the experiment, the subjects are notified of their actual roles. Payoffs are calculated

[7] For example, if a differential behavior is driven by a preference for a certain group that has unfavorable characteristics, reframing such unfavorable characteristics may be effective. If people form different beliefs toward different sources of risk, disclosing probabilities (or proportions of allocation) may be effective.

depending on the actual roles, but because subjects do not know their roles until the end of the experiment, all subjects have incentives to behave as if they are dictators. Many studies use this strategy method, "role uncertainty," to save costs in their experiments. It is known that the use of role uncertainty leads to biases such that it makes dictators more altruistic in their decisions compared to role certainty (Iriberri and Rey-Biel 2011; Mesa-Vázquez, Rodriguez-Lara and Urbano 2021; Walkowitz 2021). In our design, role uncertainty would not cause such a problem because we exclude the self-serving motivation.[8]

Initially in this section,
You (Person A) have 10,
Person B has different endowments,
Person C has 6 with 10% chance, and 0 with 90% chance.

You are going to decide to give 4 to either Person B or C from your pocket.



Many people prefer to give to Person B in the beginning of the list and then switch to give to Person C at some point. Therefore, one way to complete this list is to explore the best row to switch one option to another.

Indicate your preference by selecting corresponding button to choose either option.
Please note that regardless of your choice, Person C has 6 with 10% chance, and 0 with 90% chance as an initial endowment.

You are going to give 4 to

○	B who has **0**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **0.6**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **1.2**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **1.8**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **2.4**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **3.0**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **3.6**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **4.2**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **4.8**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **5.4**.	○	C who has 6 (10%) or 0 (90%).
○	B who has **6**.	○	C who has 6 (10%) or 0 (90%).

Figure 2: Screenshot of the actual experiment: State-risk treatment

[8] Iriberri and Rey-Biel (2011), Mesa-Vázquez et al. (2021) and Walkowitz (2021) provide clear evidence on the biases caused by role uncertainty, in which they directly examine the difference between with and without role uncertainty. The use of role uncertainty in our experiment may cause some biases, but by excluding self-serving motivation, we do not come up with any systematic biases.

In each task, subjects make decisions on 11 binary choices, allocating a portion of their endowment (X) to either the certain or the risky recipient. A screenshot of the actual experiment (for the state-risk treatment, which is explained soon) is shown in Figure 2. We use the list elicitation method, which is similar to Holt and Laury (2002) or Exley (2016). Fixing the parameters of the risky recipient, we increase the initial endowment of the certain recipient (e_C) as we go down the list. In the first half of the decision page, we provide the experimental environment. We instruct the numbers highlighted in pink in the experimental screen change over tasks, so that subjects are cautious about these numbers. Fixing the amount that the dictator gives to recipients ($4 in the example), we ask dictators which recipients they prefer to make that transfer. In the example in Figure 2, the list starts with a situation in which a dictator decides to give the transfer to either the certain recipient (person B) with no endowment or the risky recipient (person C),[9] who has 6 with a 10% probability and 0 with a 90% probability.

There have been some criticisms toward this list elicitation method; typically regarding comprehension and multiple switching points (MSPs) (for example, Charness, Gneezy and Imas 2013; Charness and Viceisza 2016). We justify using the list elicitation method by showing the existence of at most one switching point under standard inequality aversion theories (Appendix C.2).[10] Indeed, in our experiments, almost 80% of dictators exhibit consistent behaviors, while 20% are inconsistent with standard social preferences. Some fraction of inconsistent dictators are considered purely selfish since no self-serving incentive is provided in our experiment. Such pure selfish dictators may make random choices, which may end up revealing multiple switching points.[11]

As in Andreoni and Sprenger (2011), we include the following statement on our experimental screen, also shown in Figure 2: "Many people prefer to give to Person B in the beginnings of the list and then switch to give to Person C at some point. Therefore, one way to complete this list is to explore the best row to switch one option to another." This is justifiable based on standard theories (in Appendix C.2) and results from our online pilot experiments, conducted without such a statement.

[9] Note that when we explain our experimental settings in Figure 1, D denotes the dictator, C denotes the certain recipient, R denotes the risky recipient.

[10] We also show that dictators prefer to transfer to a certain recipient at the beginning of the list and then switch to the risky recipient at some point.

[11] Andersen, Harrison, Lau and Rutström (2006) suggest a possibility that multiple switching behaviors reflect indifference among the choices because of the lack of an explicit option revealing indifference. Moreover, Yu, Zhang and Zuo (2021) find that some multiple switching behaviors may be led by mistakes; a nudge protocol that gives dictators one more chance to reconsider their choices reduces multiple switching behaviors by 21% compared with standard protocol.

Table 1: Summary of risky task

	Dictator (D)	Certain Recipient (C)	Risky Recipient(R)
Panel A: State-risk treatment			
Endowments	10	e_C	0 with prob. p e_{rH} with prob. $1-p$
Transfers from a third party	0	0	0
Final allocations when D gives X to C.	$10-X$	e_C+X	0 with prob. p e_{rH} with prob. $1-p$
Final allocations when D gives X to R.	$10-X$	e_C	X with prob. p $X+e_{rH}$ with prob. $1-p$
Panel B: Non-state-risk treatment			
Endowments	10	e_C	0
Transfers from a third party	0	0	0 with prob. p e_{rH} with prob. $1-p$
Final allocations when D gives X to C.	$10-X$	e_C+X	0 with prob. p e_{rH} with prob. $1-p$
Final allocations when D gives X to R.	$10-X$	e_C	X with prob. p $X+e_{rH}$ with prob. $1-p$

To understand whether sources of risks matter for dictator's behavior, we use framing. There are two treatment arms in this experiment: the "state-risk treatment" and "non-state-risk treatment." Making all the other conditions equal, the state-risk treatment describes the situation where a risky recipient is endowed with $\$e_{rH}$ with probability $1-p$ and \$0 with probability p. By contrast, the non-state-risk treatment describes the situation where a risky recipient certainly has no initial endowment, but has a chance to receive $\$e_{rH}$ from a third party (the experimenter). In Appendix A.1, we present a screenshot of the non-state-risk treatment, which corresponds to Figure 2. Except for framing in the initial endowment of the risky recipient, there is no difference between the two treatments. Both risks are not based on the subject's behavior, but are exogenously given by the experimenter. As Eckel, Grossman and Johnston (2005) show, people prefer to give to poor people (i.e., low endowment) rather than rich people (with high endowment). For the state-risk treatment, dictators may think that they prefer to avoid giving to rich recipients, while for the non-state-risk treatment, dictators may not necessarily want to avoid those who are poor but receive some transfer from outside based on luck. Conversely, people may feel strong envy for those who win a lottery.

Table 2: Summary of parameters in the experiment

Pair	Transfer (X)	Endowment of risky-rich (e_{rH})	probability of risky-poor (p)
Pair 1	5	6	0.9
Pair 2	5	6	0.8
Pair 3	5	12	0.9
Pair 4	3	6	0.9
Pair 5	3	6	0.8
Pair 6	3	12	0.9
Pair 7	5	6	0.95
Pair 8	5	12	0.95

This table summarizes the experimental parameters. We vary transfers ($X \in \{5, 3\}$), the endowments of risky recipients when they have high endowments (or receive transfers from a third party) ($e_{rH} \in \{6, 12\}$), and the probabilities that risky recipients are poor ($p \in \{0.8, 0.9, 0.95\}$).

In such a case, dictators may become more risk averse in the non-state-risk treatment.

Table 1 summarizes the initial endowments and payoffs of both treatments, while Table 2 reports the list of parameters in our experiment. We vary (i) the amount of fixed transfers, (ii) endowment of the risky-rich, and (iii) the probability of being rich.

To understand how people evaluate ROO, one can think of simpler experimental designs. Some alternative designs are summarized in Figure 3. First, consider the simplest dictator game with risks in Figure 3-(1), where dictators decide x_d^{other}, the quantity they give to a risky recipient. While it is possible to compare x_d^{other} between treatments, which seems reasonable, it includes the motivation of an *excuse* not to give to others, which may interact with risks. As Exley (2016) shows, people use risks on others as excuses with self-serving motivation. It is thus difficult to disentangle excuse motives for risk attitudes under this simplified setting. By contrast, our experiment fixes the amounts that dictators can transfer for each round, so that we can exclude the excuse and isolate risk attitudes.[12]

Second, we can consider another experiment with a simpler lottery in Figure 3-(2) based on the outcomes. That is, we may ask a dictator to decide lottery p' with certain allocation $(me, other) = (6, 4.5)$ versus lottery q' with probabilistic allocation, $(me, other) = (6, 4)$ with 90% probability and $(me, other) = (6, 9)$ with 10% probability. While the structure of the experiment seems simpler to analyze because there are only two agents, interpreting the decision-makers' choice does not make much sense when we want to investigate the importance of ROO: choosing between the two lotteries

[12]For selfish dictators, fixing the amount they can give to recipients implies that they are indifferent to any choices in our experiment. Our results suggest that this is not the case, because dictators change the switching point based on risky recipients' risks.

means that the decision maker is also choosing the other's state itself. This will be an interesting topic for future research, but it is beyond the scope of our research questions. Our experimental design and theoretical model, on the other hand, explicitly consider safe and risky agents and ask decision makers to decide to whom to make a transfer. That is, dictators do not decide the other's state, because of the presence of both safe and risky agents.

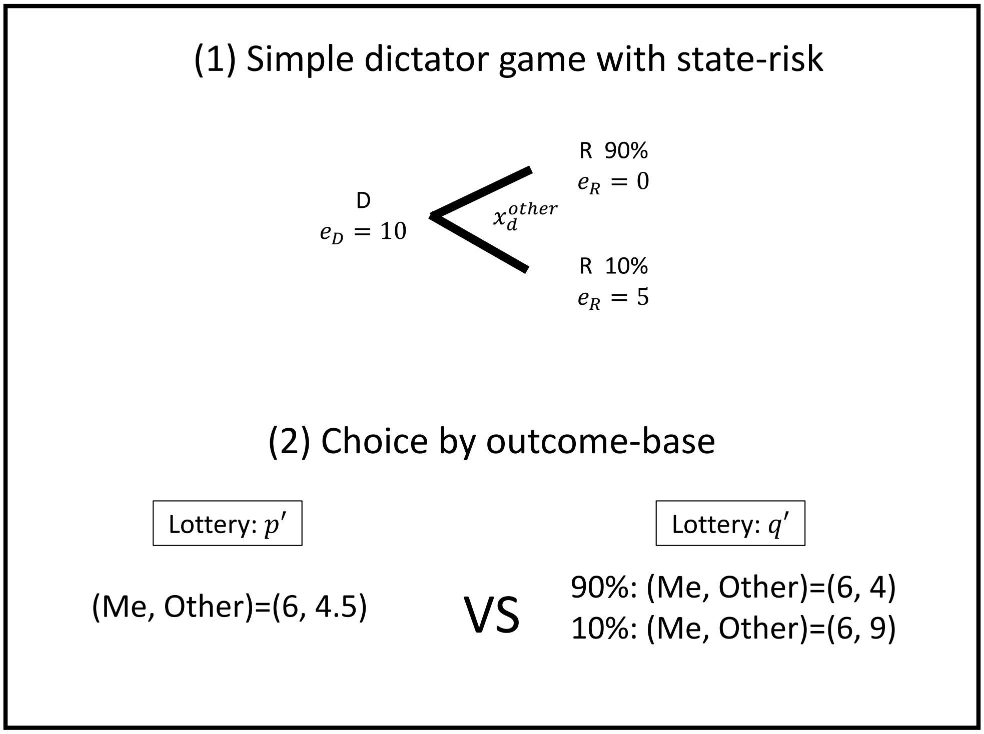


Figure 3: Potential alternative designs

3 Analytical Framework

In this section, we theoretically examine ROO using our experimental design. Specifically, we define risk attitude toward others, discuss its elicitation and propositions stemming from the definition.

3.1 Set up

Let $a = (a_D, a_C, a_R)$ be the final allocation of the three subjects, including each initial endowment, where subject D is a dictator, and C and R are the certain and risky recipients, respectively. Let $e = (e_D, e_C, e_R)$ be a vector of the initial endowments of the three subjects. For example, when a dictator allocates X to recipient C, the final allocation becomes $a = (e_D - X, e_C + X, e_R)$. For each treatment, $e_R = e_{rL}$ is realized with probability p and $e_R = e_{rH}$ is realized with probability $1-p$, where $e_{rL} = 0 < e_{rH}$.

Dictators have preferences over lotteries on the final allocations a. A lottery is denoted by L. For example, lottery $L = \{p : a_1, 1-p : a_2\}$ gives a_1 with probability p and a_2 with probability $1-p$. In our experiment, we ask dictators to give X to either recipient C or R. This is equivalent to asking dictators to choose one of the following two lotteries, $L_C^T(e_C)$ or $L_R^T(e_C)$:

$$\begin{cases} L_C^T(e_C) \equiv \{p : \underbrace{(e_D - X, e_C + X, 0)}_{final\ allocation},\ 1-p : \underbrace{(e_D - X, e_C + X, e_{rH})}_{final\ allocation}\} \\ L_R^T(e_C) \equiv \{p : (e_D - X, e_C, X),\ 1-p : (e_D - X, e_C, e_{rH} + X)\}. \end{cases} \tag{1}$$

Superscripts $T \in \{S, N\}$ refer to treatments, where S is the state-risk treatment and N is the non-state-risk treatment, and subscripts C and R refer to the dictator's choice, recipient C or R. These two lotteries depend on e_C, which varies from 0 to e_{rH} in our experiment. Note that the final allocations are the same between the two treatments S and N.

In our analytical framework, we assume a dictator has a standard social preference as follows.

Definition 1 *A dictator is said to have a standard social preference if there exists a unique cutoff $e_C^T \in (0, e_{rH})$, such that $L_R^T(e_C^T) \sim L_C^T(e_C^T)$, $L_C^T(e_C) \succ L_R^T(e_C)$ for any $e_C \in [0, e_C^T)$ and $L_R^T(e_C) \succ L_C^T(e_C)$ for any $e_C \in (e_C^T, e_{rH}]$.*

Intuitively, when a dictator needs to transfer some money, those with some social preferences will want to transfer to the poor rather than the rich. Under standard social preferences, when recipient C's initial endowment is 0 (i.e., $e_C = 0$), the dictator should prefer to transfer to recipient C ($L_C^T(0) \succ L_R^T(0)$), because recipient C is likely to be poorer than recipient R. Conversely, when recipient C's initial endowment is e_{rH} (i.e., $e_C = e_{rH}$), the dictator should prefer to transfer to recipient R ($L_R^T(e_{rH}) \succ L_C^T(e_{rH})$), because recipient R is likely to be poorer than recipient C. The definition implies that dictator's preference reverses from $L_C^T(e_C) \succ L_R^T(e_C)$ to $L_R^T(e_C) \succ L_C^T(e_C)$, with at most one switching point $e_C^T \in (0, e_{rH})$ as e_C increases from 0 to e_{rH}.

This definition can be justified by representative social preference models. We prove that a dictator following Saito (2013) have a standard social preference in Appendix C.2. However, we may not observe such preference if a dictator is entirely selfish or prefers inequality.[13]

3.2 Risk attitude toward others: Absolute evaluation

To capture the effect of ROO in our experimental environment, we formalize the risk attitude toward others. First, we evaluate the dictator's risk attitude in *absolute* terms, without comparing it with those of other dictators.

Consider two lotteries, $L_C^T(\overline{e_r})$ and $L_R^T(\overline{e_r})$ for each T, where $\overline{e_r}$ is the expected allocation held by a risky recipient without receiving a transfer from a dictator (i.e., $\overline{e_r} = p0 + (1-p)e_{rH}$). In each lottery, recipient C has $\overline{e_r}$ for sure and recipient R is expected to have $\overline{e_r}$. We define risk attitude toward others as a preference between $L_C^T(\overline{e_r})$ and $L_R^T(\overline{e_r})$.

Definition 2 *For each treatment $T \in \{S, N\}$, fix a decision problem between $L_C^T(\overline{e_r})$ and $L_R^T(\overline{e_r})$. Then, a dictator is said to be (i) risk averse, (ii) risk neutral, and (iii) risk loving toward others if (i) $L_C^T(\overline{e_r}) \succ L_R^T(\overline{e_r})$, (ii) $L_C^T(\overline{e_r}) \sim L_R^T(\overline{e_r})$, and (iii) $L_R^T(\overline{e_r}) \succ L_C^T(\overline{e_r})$, respectively.*

A dictator is said to be risk averse if they prefer a certain recipient C who certainly has $\overline{e_r}$ to a risky recipient R who is expected to have $\overline{e_r}$. Unlike the risk attitude in the standard expected utility theory, our risk attitude depends on each decision problem.[14] While this is counterintuitive, it may be natural when we capture risk attitude in relation to other regarding behavior. We demonstrate this fact using the inequality aversion model in Section 4.1.

By the definition of standard social preference, $e_C^T > \overline{e_r}$ implies $L_C^T(\overline{e_r}) \succ L_R^T(\overline{e_r})$, and $e_C^T < \overline{e_r}$

[13] In experimental studies, there are some individuals who exhibit opposite preferences, but their proportion is small. For example, Iriberri and Rey-Biel (2013) classify individual preferences into four categories: selfish, social welfare maximizer, inequality averse, and competitive. Individuals with the competitive preference, who prefer to reduce others' allocations, account for only 13% of all individuals in their experiment. Since the final allocations of dictators and the sum of final allocations are fixed, regardless of the choices by the dictators, the observed behaviors in our experiment are not driven by preferences such as selfish and social welfare maximizers. In our experiment, we exclude approximately 22% of observations from the analysis, as they exhibit choices inconsistent with the standard social preferences.

[14] That is, dictators can be classified as risk averse for some decisions and risk loving for others.

implies $L_R^T(\overline{e_r}) \succ L_C^T(\overline{e_r})$. Therefore, risk attitudes toward others are characterized by switching points e_C^T as follows.

Proposition 1 *Suppose a dictator has a standard social preference. The dictator is* (i) *risk averse,* (ii) *risk neutral,* (iii) *risk loving toward others, if and only if* (i) $e_C^T > \overline{e_r}$, (ii) $e_C^T = \overline{e_r}$, (iii) $e_C^T < \overline{e_r}$ *for each* $T \in \{S, N\}$, *respectively.*

3.3 Risk attitude toward others: Relative evaluation

Second, we evaluate dictator's risk attitude toward others in *relative* terms based on a comparison between two dictators. We consider two dictators, i and j, and describe their preferences by $\succsim_i$ and $\succsim_j$, respectively. We define *being more risk averse* as follows.

Definition 3 *For any two dictators* i *and* j, *dictator* i *is said to be more risk averse toward others than* j *if it holds* $L_C^T(e_C) \succ_i L_R^T(e_C)$ *for any* e_C *such that* $L_C^T(e_C) \succsim_j L_R^T(e_C)$.

We consider dictator i to be *more risk averse* than dictator j if they strictly prefer a safe choice, giving to recipient C, whenever subject j weakly prefers it.

We denote the switching point chosen by dictator k as e_{Ck}^T. Under the assumption of standard social preference, $e_{Ci}^T > e_{Cj}^T$ implies that dictator i is more risk averse than j. Otherwise, by the definition of more risk averse, there exists e'_C such that $L_C^T(e'_C) \succsim_j L_R^T(e'_C)$ and $L_R^T(e'_C) \succsim_i L_C^T(e'_C)$. Then, $e_{Cj}^T \geq e'_C$ and $e'_C \geq e_{Ci}^T$ hold, which contradicts $e_{Ci}^T > e_{Cj}^T$. Moreover, if dictator i is more risk averse than j, it holds that $L_C^T(e_{Cj}^T) \succ_i L_R^T(e_{Cj}^T)$, so that $e_{Ci}^T > e_{Cj}^T$. Therefore, *being more risk averse toward others* can be characterized by switching points as the following proposition.

Proposition 2 *Suppose dictator* i *and* j *have a standard social preference. Dictator* i *is more risk averse toward others than* j *if and only if* $e_{Ci}^T > e_{Cj}^T$ *holds.*

3.4 An application of relative evaluation: state-risk vs non-state-risk

We can apply the relative evaluation to the comparison between the two treatments in our experiment. Since final allocations are the same between the two treatments, if average switching points are statistically different between them, dictators have different risk attitudes toward others by treatment. When we denote the average switching point in treatment T as $\overline{e_C^T}$, Proposition 2 implies the following.

Claim 1 *Suppose dictators have a standard social preference. If* $\overline{e_C^S} > (<) \overline{e_C^N}$ *is observed, dictators are more risk averse in the state-risk (non-state-risk) treatment on average. If* $\overline{e_C^S} = \overline{e_C^N}$ *is observed, there is no difference between state-risk and non-state-risk in dictators' risk attitudes toward others.*

Using the relative evaluation, we also examine the effects of subjects' characteristics—gender, college major, and race—on their risk attitudes toward others.

4 Model analysis and predictions using existing theories

Using existing theoretical models, we analyze our experimental environment and predict behaviors. We also discuss extension of these models to explain our experimental results in Section 6. Let $u(a)$ be the utility of a dictator for a final allocation a, and assume the following additive separable form:

$$u(a) = v(a_D) + \sum_{k\in\{C,R\}} g_k(a_D - a_k), \tag{2}$$

where $v : \mathbb{R}_+ \to \mathbb{R}_+$, $g_k : \mathbb{R} \to \mathbb{R}_-$. We assume that $g_k(0) = 0$, $g'_k(x) < 0$ if $x > 0$ and $g'_k(x) > 0$ if $x < 0$, which are the basic properties of inequality aversion models. One of the most representative forms of u is the following function, suggested by Fehr and Schmidt (1999):

$$u^{FS}(a) = a_D + \sum_{k\in\{C,R\}} \Big[-\alpha_k \max[a_D - a_k, 0] - \beta_k \max[a_k - a_D, 0]\Big], \tag{3}$$

where $\alpha_k \geq 0$, $\beta_k \geq \alpha_k$, and $\beta_k \leq 1$ for each k. Following Fudenberg and Levine (2011), Brock et al. (2013), and Saito (2013), we assume that the utility for lottery depends on both ex-ante and ex-post fairness concerns. Let the utility for lottery L be $U(L)$. We assume that there exists an increasing function $W(y, z)$ in both y and z, and $U(L)$ is described as follows:

$$U(L) = W(u(\mathbb{E}_L[a]), \mathbb{E}_L[u(a)]), \tag{4}$$

where $\mathbb{E}_L$ is the expectation operator by a lottery L. $\mathbb{E}_L[u(a)]$ is the expected utility of the final outcomes, which expresses the dictator's ex-post perspective for inequality. $u(\mathbb{E}_L[a])$ is the utility of the expected outcome, which expresses the dictator's ex-ante perspective for inequality. It has been experimentally observed that dictators consider not only the standard expected utility but also the utility for expected outcomes when fairness concerns exist (Brock et al. 2013). One special case of the utility function (4) is the axiomatized form of Saito (2013):

$$U^{Saito}(L) = \delta u^{FS}(\mathbb{E}_L[a]) + (1-\delta)\mathbb{E}_L[u^{FS}(a)], \tag{5}$$

where $\delta \in (0,1)$, which captures the relative weighting of ex-ante fairness.

4.1 Model analysis: Absolute evaluation

Here, we analyze ROO in *absolute* terms using the general utility function (4), and provide predictions induced from the model. We assume that $g_C(\cdot) = g_R(\cdot) \equiv g(\cdot)$, which is reasonable in our experiment because dictators are matched with anonymous recipients.[15] As in the previous sections, $\overline{e_r}$ denotes an expected allocation for risky recipient without the transfer (i.e., $\overline{e_r} = p0 + (1-p)e_{rH}$). In this section, we fix recipient C's initial endowment as $\overline{e_r}$ (i.e., $e_C = \overline{e_r}$). In this case, the preference for lottery can be characterized by only ex-post fairness concerns as follows.

Lemma 1 *Consider lotteries $L_C^T(\overline{e_r})$ and $L_R^T(\overline{e_r})$ in each treatment T. Suppose a dictator's preference is represented by utility function (4) and $g_C(\cdot) = g_R(\cdot) \equiv g(\cdot)$. Then, $L_C^T(\overline{e_r}) \succsim L_R^T(\overline{e_r})$ if and only if $\mathbb{E}_{L_C^T(\overline{e_r})}[u(a)] \geq \mathbb{E}_{L_R^T(\overline{e_r})}[u(a)]$.*

This means that the preference between $L_C^T(\overline{e_r})$ and $L_R^T(\overline{e_r})$ is characterized by expected utility because these lotteries are indifferent with respect to the ex-ante fairness concern under the assumption of $g_C = g_R$. This is because the expected final outcomes of $L_C^T(\overline{e_r})$ and $L_R^T(\overline{e_r})$ are $(e_D - X, \overline{e_r} + X, \overline{e_r})$ and $(e_D - X, \overline{e_r}, \overline{e_r} + X)$, respectively, hence $g_C = g_R$ implies $u\left(\mathbb{E}_{L_C^T(\overline{e_r})}[a]\right) = u\left(\mathbb{E}_{L_R^T(\overline{e_r})}[a]\right)$. Lemma 1 holds under the utility function (4) in general, which is one notable property of our experimental design. By Lemma 1, we directly obtain the following.

Lemma 2 *Suppose a dictator follows utility function (4) and $g_C(\cdot) = g_R(\cdot) \equiv g(\cdot)$. Then, the dictator is (i) risk averse, (ii) risk neutral, and (iii) risk loving toward others if and only if (i) $G > 0$, (ii) $G = 0$, and (iii) $G < 0$, respectively, where*

$$\begin{aligned} G \equiv \mathbb{E}_{L_C^T(\overline{e_r})}[u(a)] - \mathbb{E}_{L_R^T(\overline{e_r})}[u(a)] = & \underbrace{g(e_D - 2X - \overline{e_r}) + pg(e_D - X) + (1-p)g(e_D - X - e_{rH})}_{=\mathbb{E}_{L_C^T(\overline{e_r})}[u(a)]\,:\,\textit{utility from giving } C \textit{ and not giving } R \textit{ a transfer}} \\ & - \underbrace{\{pg(e_D - 2X) + (1-p)g(e_D - 2X - e_{rH}) + g(e_D - X - \overline{e_r})\}}_{=\mathbb{E}_{L_R^T(\overline{e_r})}[u(a)]\,:\,\textit{utility from giving } R \textit{ and not giving } C \textit{ a transfer}}. \end{aligned} \tag{6}$$

[15] However, this does not hold if dictators have a preference or cognitive bias for other's state.

Equation G is defined as the net benefit of transferring to recipient C. The first term includes not only the direct utility of transferring to C (i.e., $g(e_D - 2X - \overline{e_r})$) but also the expected disutility of not transferring to R (i.e., $pg(e_D - X) + (1-p)g(e_D - X - e_{rH})$). Similarly, the second term includes the expected utility of transferring to R and the disutility of not transferring to C. Note that self-serving concern $v(a_D)$ does not matter because the dictator's payoff does not depend on their choices in our experiment. Equation (6) can be rewritten as follows, which is useful for the incoming analysis:

$$G = \mathbb{E}_{L^T_{C(\overline{e_r})}}[u(a)] - \mathbb{E}_{L^T_{R(\overline{e_r})}}[u(a)] = \underbrace{g(e_D - 2X - \overline{e_r}) - pg(e_D - 2X) - (1-p)g(e_D - 2X - e_{rH})}_{differences\ in\ expected\ utility\ of\ giving\ C\ and\ R\ a\ transfer}$$
$$-\{\underbrace{g(e_D - X - \overline{e_r}) - pg(e_D - X) - (1-p)g(e_D - X - e_{rH})}_{differences\ in\ expected\ utility\ of\ not\ giving\ C\ and\ R\ a\ transfer}\}. \quad (7)$$

Here we consider an example of the linear form by Fehr and Schmidt (1999) and a parameter set $(e_D, X, e_{rH}, p) = (10, 5, 6, 0.9)$ in our experiment, which is described in Figure 4.

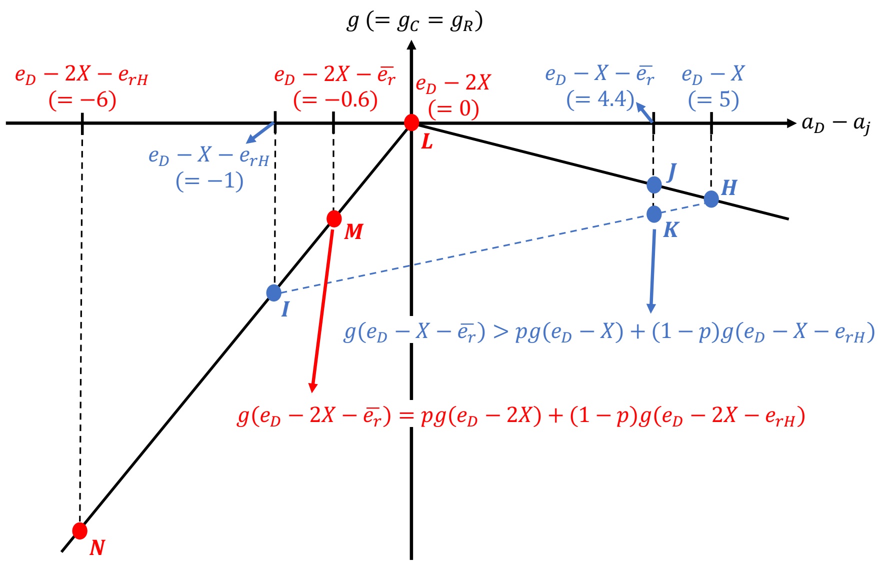


Figure 4: (Case 1) linear assumption in each domain

The first term of G in equation (7) is the difference between the utility transferring to C (the utility from $a_D - a_C = e_D - 2X - \overline{e_r} = -0.6$ with probability 1, which is point M in the figure) and the expected utility of transferring to R (the utility from $a_D - a_R = e_D - 2X$ with probability p, point L, and $a_D - a_R = e_D - 2X - e_{rH}$ with probability $1-p$, point N). When g is the linear form of Fehr and Schmidt (1999), these two utilities are the same and cancel each other out, and hence the first term of G is 0. Graphically, we can observe that both are expressed at point M in the figure. Similarly, the second term of G in (7) is the difference between the utility of not transferring to C ($a_D - a_C = e_D - X - \overline{e_r} = 4.4$, point J) and the expected utility of not transferring to R ($a_D - a_R = e_D - X$ with probability p, point H, and $e_D - X - e_{rH}$ with probability $1-p$, point I). The former is expressed at the level of point J, and the latter is at point K in Figure 4, so that the second term of G is positive. Therefore, G is negative, and we can observe that dictators are risk loving by Lemma 2.

The risk-loving behavior stems from the term of recipients who do not receive any transfer in dictators' utility (i.e., the second term of G). Moreover, $g(\cdot)$ is concave as a whole domain, which makes the second term of G positive. An alternative utility representation is to let dictators ignore the recipients without transfer (i.e., the second term of G). In this case, we expect risk-neutral behaviors by dictators.[16]

Here, we state a characterization result for the risk attitude toward others in Fehr and Schmidt (1999)'s form.

Proposition 3 *Suppose that (1) $e_D - X \geq X$, (2) $e_D - X < e_{rH} + X$, and (3) $e_D - X < e_{rH}$ under $e_D - 2X \leq \overline{e_r} \leq e_D - X$. Moreover, suppose a dictator's preference is represented by utility function (5), and $g_C = g_R$ (i.e., anonymity). Then, the dictator is (i) risk averse, (ii) risk neutral, and (iii) risk loving toward others if and only if (i) $\overline{e_r} < e_D - (1+p)X$, (ii) $\overline{e_r} = e_D - (1+p)X$, and (iii) $\overline{e_r} > e_D - (1+p)X$, respectively.*

Proof: See Appendix C.3. $\square$

Condition (1), $e_D - X \geq X$, implies that a dictator has a higher final allocation than a recipient with no initial endowment receiving transfers. Condition (2), $e_D - X < e_{rH} + X$, implies that a dictator has a lower final allocation than a recipient who initially gains e_{rH} and receiving transfers. Finally, (3), $e_D - X < e_{rH}$ under $e_D - 2X \leq \overline{e_r} \leq e_D - X$, is a technical condition that excludes peculiar outcomes which come from the linearity of Fehr and Schmidt (1999).[17] Our experimental parameters satisfy conditions (1), (2), and (3).

Intuitively, we can see that risk attitudes are characterized by $\overline{e_r}$, which is the initial endowment of recipient C in this analysis. If recipient C's initial endowment, $\overline{e_r}$, is relatively low (case (i)),

[16] It is also inconsistent with our experimental results.

[17] Otherwise (i.e., $e_D - 2X \leq \overline{e_r} < e_{rH} \leq e_D - X$), risk attitudes do not depend on $\overline{e_r}$, while in the proposition, we attempt to characterize risk attitudes by using $\overline{e_r}$.

the dictator is more likely to prefer choosing recipient C (i.e., $L^T_C(\overline{e_r})$), so that we can consider the dictator as *risk averse*. As $\overline{e_r}$ increases, the dictator is more likely to prefer recipient R (i.e., $L^T_R(\overline{e_r})$) and switches from $L^T_C(\overline{e_r})$ to $L^T_R(\overline{e_r})$ at $\overline{e_r} = e_D - (1+p)X$. Importantly, Proposition 1 holds regardless of the parameters of Fehr and Schmidt (1999)'s utility function.[18]

Using Proposition 3, the risk attitudes of dictators following Fehr and Schmidt (1999) in our experiment are shown by Table 3:

Pairs	e_D	X	e_{rH}	p	risk attitude
Pair 1	10	5	6	0.9	*risk loving*
Pair 2	10	5	6	0.8	*risk loving*
Pair 3	10	5	12	0.9	*risk loving*
Pair 4	10	3	6	0.9	*risk averse*
Pair 5	10	3	6	0.8	*risk averse*
Pair 6	10	3	12	0.9	*risk averse*
Pair 7	10	5	6	0.95	*risk loving*
Pair 8	10	5	12	0.95	*risk loving*

Table 3: Prediction of risk attitude based on utility function by (5)

Under the assumption of Fehr and Schmidt (1999)'s linear form, dictators are predicted to be risk loving in 5 pairs out of 8 in our experiment. From these results, we obtain the following prediction.

Prediction 1 *Suppose a dictator's preference is represented by utility function (4) and $g_C = g_R$. Then Fehr and Schmidt (1999)'s form predicts Table 3. If they are not observed, this implies dictators do not follow the form of Fehr and Schmidt (1999).*

4.2 Model analysis: relative evaluation between state-risk and non-state-risk

Next, we analyze sources of risk by comparing two treatments: state-risk and non-state-risk. Using the general utility function (4), since it is an outcome-based model, it predicts no statistical difference between two treatments. This implies the source of the risk does not affect decision-makers' choices.

Prediction 2 *Suppose a dictator's preference is represented by utility function (4). Then, there is no statistical difference between state-risk and non-state-risk. Otherwise, it implies that dictators do not follow outcome-based utility function (4).*

[18] This property holds under restricted experimental parameters. Generally, the risk attitude can be affected by Fehr and Schmidt's parameters.

4.3 Relation to previous literature

There are some theoretical and experimental studies on the intersection of social preferences and risks. Experimental evidence by Brock et al. (2013) shows that, when we consider the risk or uncertainty toward other people, we consider both ex-ante and ex-post fairness. This implies that it might not be sufficient to consider the expected utility of representative other-regarding models, such as Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2002). This problem was theoretically examined by Fudenberg and Levine (2011), axiomatized by Saito (2013), and generalized to a more flexible functional form by López-Vargas (2014) and Rau and Müller (2017).[19]

Our contributions to the literature are as follows. As a first contribution, we build a methodology to examine ROO. Existing studies on social preference and risk focus on risk attitudes for decision-makers' own payoffs. For example, some studies investigate the relationship between decision-maker's risk attitudes for their own payoffs and inequality aversion (López-Vargas 2014, Rau and Müller 2017). Hitherto, less attention has been paid to the risk attitudes toward other's payoffs, and existing studies do not provide experimental and theoretical methodologies to properly capture ROO. We thus develop a novel experimental design to elicit decision-makers' risk attitudes toward others from two aspects: *absolute* and *relative* evaluations. Our concepts of risk attitude toward others are simply characterized by the switching points chosen by dictators.

The second contribution is on the *absolute* evaluation of the risk attitudes toward others. We show the first evidence on how people behave towards others with risks, and we find that decision makers are consistently risk averse toward others. In the theoretical predictions, the axiomatized form of Saito (2013), a direct extension of Fehr and Schmidt (1999) to risky environments,[20] predicts that dictators are risk loving toward others in 5 out of 8 sets of parameters in our experiment. Our experimental results suggest that Fehr and Schmidt (1999)'s linear specification is not capable of making reasonable predictions for risk attitudes toward others. Bellemare et al. (2008) also show that Fehr and Schmidt (1999)'s linear form does not fit the behaviors in their ultimatum game, suggesting an increasing and concave utility function of the disadvantageous payoff difference. How-

[19] In addition, some studies experimentally examine people's preference for fairness and redistribution in the situation where allocation among agents depends on their risk-taking behaviors (Krawczyk 2010, Cappelen, Konow, Sørensen and Tungodden 2013, Mollerstrom, Reme and Sørensen 2015, Akbaş, Ariely and Yuksel 2019). They also focus on ex-ante and ex-post fairness, but decision makers in their experiments are not exposed to ROO.

[20] To be precise, Saito (2013) axiomatizes a more general form which includes Fehr and Schmidt (1999) as a special case.

ever, their subjects make decisions under a non-risky environment, not under a risky environment including ROO. Our result implies that the nonlinear inequality aversion would work better even in risky decision environments.

The last contribution is on the *relative* evaluation of risk attitudes toward others. In our experiment, we examine two sources of risks by comparing two treatments. As mentioned above, existing models predict that there will be no difference between two treatments, because they are all outcome-based theories. In contrast, we leave room to observe the difference through the effect of other's state. Many experiments in non-risky decision environments show that people have preferences regarding others' initial endowments, states, or types. People prefer to give donations to known charitable organizations over anonymous others (Eckel and Grossman 1996), those with revealed social status over those without it (Charness and Gneezy 2008), poor over rich (Eckel et al. 2005), and even artificially created "minimal group" members in a lab over non-members (Chen and Li 2009). We naturally infer from the above-cited papers that others' states have some peculiar effects on subjects' behavior even in our risky environment. To the best of our knowledge, Fong and Oberholzer-Gee (2011) is the only exception that incorporates risk into unfavorable agents. The recipients of their dictator game are a disabled group (preferred to give) and drug users (not preferred to give). They ask dictators how much they want to transfer with and without uncertainty. When confronting an uncertain situation, they provide an option to eliminate uncertainty, and some dictators pay for that option. However, Fong and Oberholzer-Gee (2011) do not explicitly state the degree of uncertainty (i.e., dictators do not know the proportion of unfavorable recipients in the uncertain treatment). Consequently, they also cannot identify whether the composition of the unfavorable group or risk itself matter when people are willing to pay to avoid risks. Our experimental design can clearly identity the effects of the other's states. We also explore the potential mechanisms behind the behavioral difference toward different types of risks from the viewpoints of *preference* and *cognitive bias* for others' states in Section 6.

In terms of *relative* evaluation, we also examine behavioral differences by demographic characteristics (i.e. gender, college major, and race). One interesting observation from our experiment is that male subjects show more risk-averse attitudes toward others than female subjects do. As many existing studies show that women are more risk averse than men (see Eckel and Grossman 2008), our findings suggest that ROO may function differently from standard risk attitudes toward own.

5 Experimental results

The experiment was conducted in October and November 2018 at the computer laboratory in the University of Maryland, College Park. We had 16 sessions with 192 subjects in total (12 participants per session and 96 participants by treatment); all participants were undergraduate students at the University of Maryland. Including the instruction, participants spent 30 to 45 minutes at the laboratory, and earned \$18.11 on average, which include a participation fee of \$7. The maximum earnings in our experiment were \$25.25 and the minimum earnings were \$9.5.[21] The entire session was conducted on a computer, and oTree (Chen, Schonger and Wickens 2016) was used to program the experiment.

5.1 Main results

First, we simply compare the expected values of risky recipients' allocations without transfers (i.e., $\overline{e_r} = p0 + (1-p)e_{rH}$) and mean switching points to examine ROO in *absolute* terms for each lottery. To recap our concepts, the switching point corresponds to the point when dictators switch from giving to a certain recipient to giving to a risky recipient. Experimentally, the switching point is defined as the first point after the dictator switches their choice. This comparison enables us to evaluate individuals' risk attitudes, as shown in Proposition 1. Individuals who do not exhibit the standard social preferences (defined in Definition 1) are excluded from the analyses. That is, those who have more than 2 switching points or those who preferred risky recipients in the first part of the list and switched to certain recipients in the later part of the list are excluded. This reduces the number of observations by approximately 20%.

Results of absolute evaluation is presented in Table 4. For each set of parameters, the results exhibit strong and robust risk aversion, because the mean switching points are much greater than $\overline{e_r}$. As discussed in Section 4.1 and summarized in Prediction 1, Fehr and Schmidt (1999)'s utility function *cannot* explain this phenomenon: observed strong risk aversion suggests applying nonlinearity for the utility function. One potential solution is to allow curvature in the utility function, as suggested in Bellemare et al. (2008).[22] Note that this result is not likely to be observed from

[21] To fully control the environment, we restricted the number of participants to 12 per session. To guarantee this number, we invited more than 12 subjects per session and paid a participation fee of \$7 for those who came to the lab after the 12th subject.

[22] We cannot back-up the functional form of the utility from this experiment with a structural estimation, because there is no variation in own allocations. That is, we fixed own allocations as well as the amount of transfers in our

the "random" choices of subjects. If subjects are completely selfish and do not care for others' allocations, the choices in our experiment do not matter for subjects because we fix the amount of transfers (and therefore the amounts of dictators' final allocations). Let n be the nth option in the experimental lists (within 11 options), where dictators switch their behaviors. We test the hypothesis that the averages of n are jointly the same across different sets of parameters. The F-test rejects the hypothesis ($F = 3.82$, $p = 0.0009$), which suggests that decision makers behave differently across different sets of parameters. Additionally, among all subjects, only 15 subjects chose the same turning points for all the eight sets of parameters. Fourteen of them chose the edge of the options, and one of them kept choosing the 5th turning point throughout the games. The rest of the participants (92.3% of all subjects, conditional on those who were consistent with standard social preferences) choose different turning points by different sets of parameters.

Table 4: Main results: Absolute evaluation

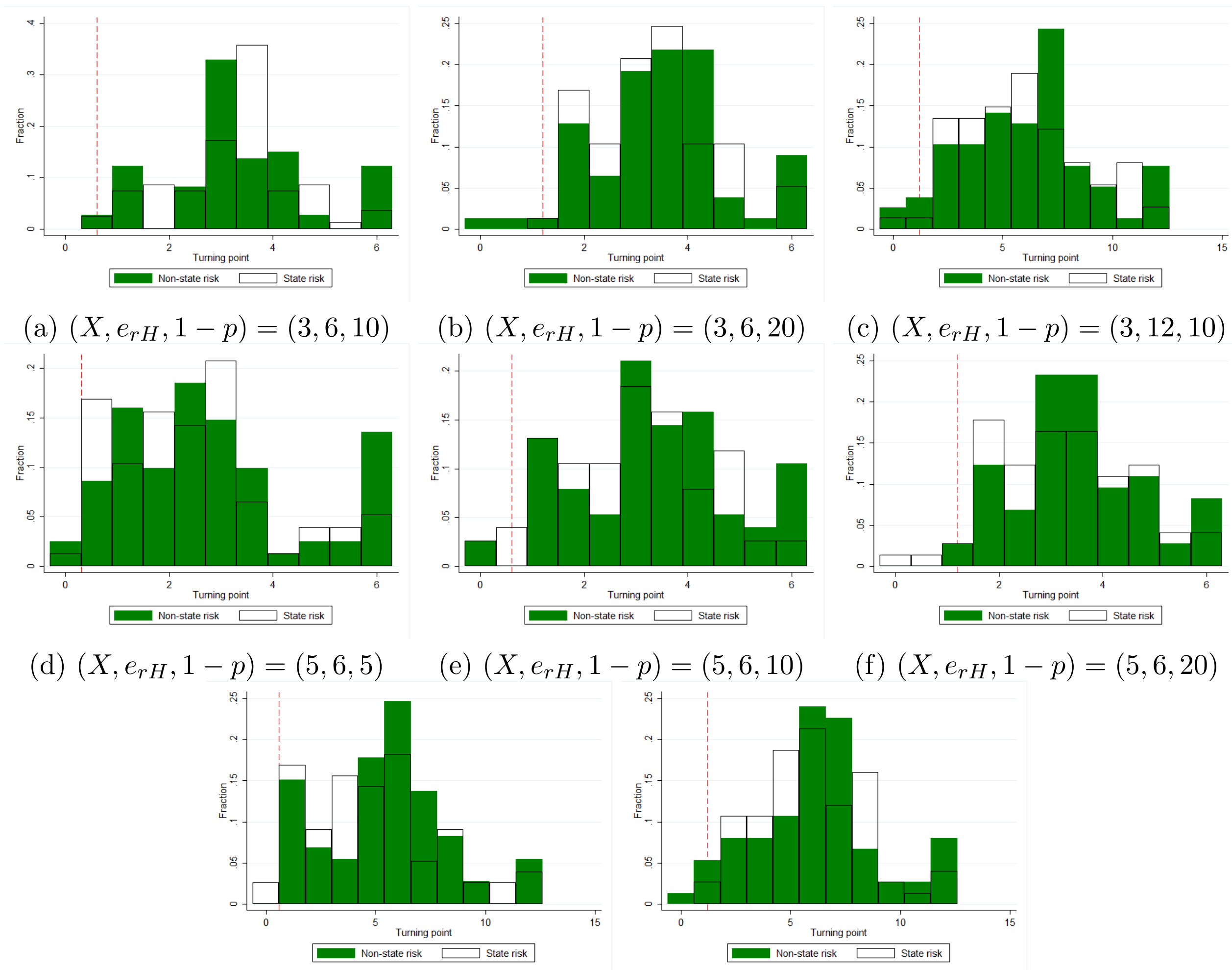
	Parameters					Results		
	e_D	X	e_{rH}	p	$\overline{e_r}$	Mean Switching Point	S.D.	N
Pair 1	10	5	6	0.9	0.6	3.17	1.55	152
Pair 2	10	5	6	0.8	1.2	3.41	1.34	146
Pair 3	10	5	12	0.9	1.2	6.12	2.79	150
Pair 4	10	3	6	0.9	0.6	3.34	1.38	154
Pair 5	10	3	6	0.8	1.2	3.46	1.30	155
Pair 6	10	3	12	0.9	1.2	6.06	2.98	152
Pair 7	10	5	6	0.95	0.3	2.68	1.73	158
Pair 8	10	5	12	0.95	0.6	5.22	3.02	150

$\overline{e_r} = p0 + (1-p)e_{rH}$. Individuals who do not exhibit the standard social preferences (defined in Definition 1) are excluded from the analyses. That is, those who have more than 2 switching points or those who preferred rich recipients in the first part of the list and switched to risky recipients in the later part of the list are excluded.

Second, we present distributions of switching points in Figure 5 under each treatment. This enables us to examine risk attitudes toward others in *relative* terms, in terms of the behavioral differences by sources of risk. The green bars represent frequencies of switching points in the non-state-risk treatment ($T = N$), and the clear (white) bars represent those in the state-risk treatment ($T = S$). The vertical red dashed lines represent the expected values of risky recipients' allocations without transfers (i.e., $\overline{e_r} = p0 + (1-p)e_{rH}$) for each lottery. Again, we exclude those who exhibit

experiment. This is going to be a trade-off between allowing the potential to excuse (Exley 2016) to avoid transferring to others using ROO. We believe this will be an interesting research agenda in the future.

behaviors inconsistent with standard social preferences.



(a) $(X, e_{rH}, 1-p) = (3, 6, 10)$ (b) $(X, e_{rH}, 1-p) = (3, 6, 20)$ (c) $(X, e_{rH}, 1-p) = (3, 12, 10)$

(d) $(X, e_{rH}, 1-p) = (5, 6, 5)$ (e) $(X, e_{rH}, 1-p) = (5, 6, 10)$ (f) $(X, e_{rH}, 1-p) = (5, 6, 20)$

(g) $(X, e_{rH}, 1-p) = (5, 12, 5)$ (h) $(X, e_{rH}, 1-p) = (5, 12, 10)$

Figure 5: Main results: dictators' decisions by each parameter[23]

These figures again visually present that most dictators are risk averse, meaning that their switching points are much greater than $\overline{e_r}$. Second, we compare the two treatments to examine risk attitudes in *relative* terms by comparing behaviors between treatments. We do not observe any clear differences between these treatments. In addition to the visual evidence in Figure 5, regression

[23]Figures illustrate the main results. The green bars represent the frequencies of switching points for non-state-risk treatment and the clear (white) bars represent those for the state-risk treatment. Each small figure is labeled with experimental parameters: transfers from dictators (X), allocations before transfers for risky recipients when they endow or receive some amounts (e_{rH}), and probabilities that risky recipients endow or receive some amounts $(1-p)$. The vertical red dashed lines represent the expected values of risky recipients' expected allocations without transfers (i.e., $\overline{e_r} = p0 + (1-p)e_{rH}$) for each lottery.

analyses in Table 5 and a simple t-test for checking the differences in the means of the switching points in Table 6 also suggest behaviors are indistinguishable between the two treatments.

Table 5: Regression analyses Dependent variable: Switching point of dictator's choice

	(1)	(2)	(3)	(4)
Other Risk	-0.308	-0.380	-0.0640	-0.155
	(0.250)	(0.446)	(0.212)	(0.345)
Endowment of Rich	0.471***	0.475***	0.464***	0.461***
	(0.0272)	(0.0430)	(0.0225)	(0.0321)
Prob. being Rich	0.0506***	0.0442***	0.0565***	0.0550***
	(0.00728)	(0.0110)	(0.00571)	(0.00801)
Endowment of Rich × Other-Risk		-0.00875		0.00722
		(0.0544)		(0.0450)
Prob. Rich × Other-Risk		0.0128		0.00279
		(0.0145)		(0.0114)
Constant	-0.117	-0.0816	-0.473*	-0.426
	(0.251)	(0.367)	(0.206)	(0.264)
Observations	1217	1217	1133	1133
R^2	0.267	0.268	0.335	0.335

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors are clustered by individual level. Specifications (1) and (2) include all observations except those who exhibit inconsistent preferences, while specifications (3) and (4) further exclude those who kept answering either left side or right side throughout the list.

In Table 5, we estimate the following regression:

$$Switching_i = \alpha + \beta_1 OtherRisk_i + \beta_2 EndowRich_i + \beta_3 ProbRich_i + \varepsilon_i, \tag{8}$$

where $Switching_i$ is the actual value of the switching points (in \$), $OtherRisk_i$ is a dummy variable taking 1 if the treatment is state-risk ($T = S$), $EndowRich_i$ is an endowment of a risky recipient without the transfer when the recipient is rich (i.e. e_{rH}), and $ProbRich_i$ is the probability that the risky recipient is rich $(1 - p)$. If there is a behavioral difference between two treatments, we should observe $\beta_1 \neq 0$, which is the parameter of interest. Additionally, if dictators exhibit standard social preferences, we should observe $\beta_2 > 0$ and $\beta_3 > 0$, because dictators should prefer safe recipients when the endowments of risky-rich recipients are high and the probability that risky recipients are rich is high as well.

Further, we include interaction terms for treatment status and other variables as follows:

$$\begin{aligned} Switching_i = \alpha &+ \beta_1 OtherRisk_i + \beta_2 EndowRich_i + \beta_3 ProbRich_i \\ &+ \beta_4 OtherRisk_i \times EndowRich_i + \beta_5 OtherRisk_i \times ProbRich_i + \varepsilon_i. \end{aligned} \tag{9}$$

The results of estimating equation (8) are presented in columns (1) and (3) of Table 5, and the results of estimating equation (9) are in columns (2) and (4). Columns (1) and (2) use all subjects, except those who exhibit behaviors that are inconsistent with standard social preferences. The subjects used in these regressions correspond to 79.32% of all observations. Moreover, columns (3) and (4) exclude those who prefer to give one particular recipient. That is, we exclude those who keep giving to a certain (or risky) recipient throughout the entire list. The subjects used in these regressions correspond to 73.76% of all observations. For all regressions, we do not observe any differences in behaviors by treatment status (i.e., $\beta 1$, $\beta 4$, and $\beta 5$ are statistically indistinguishable from 0), though β_1 is negative for all specifications.

In Table 6, we compare the means of the switching points (in USD) by treatment status. For any set of parameters, we observe that switching points are higher in the non-state-risk treatment, though differences are not statistically significant.

Overall, our findings can be summarized as follows. First, people exhibit robust risk aversion toward others, as opposed to the existing models such as Fehr and Schmidt (1999). Second, we observe that switching points are higher in the non-state-risk treatment, although the differences are statistically insignificant. In the next section, we conduct heterogeneity analyses and show that these differences are driven by particular groups in our sample.

5.2 Heterogeneity by group

As proposed in Section 3.3, we can use switching points to compare risk attitudes between subjects by their characteristics. As an application, we have already shown risk attitudes by different sources of risk. In this subsection, we further examine whether subjects exhibit different risk attitudes toward others by gender, college major, and racial group.

Table 7 shows means, standard errors and differences in the values of switching point by each demographic characteristics interacted with two treatments. We find that male, STEM-major, and white students exhibit higher switching points on average than female, non-STEM, and non-white students. This means that those population groups are more risk averse toward others. For gender

Table 6: Means and differences for each experiment

Parameters (x, e_r, p)	State-risk	Non-State-risk	Difference
(3, 6, 10)	3.28 (0.138)	3.41 (0.179)	-0.129 (0.224)
N	81	73	
(3, 6, 20)	3.36 (0.138)	3.53 (0.157)	-0.172 (0.209)
N	78	77	
(3, 12, 10)	5.97 (0.330)	6.15 (0.350)	-0.19 (0.485)
N	78	74	
(5, 6, 5)	2.50 (0.182)	2.85 (0.205)	-0.35 (0.275)
N	77	81	
(5, 6, 10)	2.95 (0.166)	3.39 (0.187)	-0.442 (0.250)
N	76	76	
(5, 6, 20)	3.28 (0.160)	3.56 (0.152)	-0.279 (0.221)
N	73	73	
(5, 12, 5)	4.92 (0.350)	5.53 (0.345)	-0.615 (0.350)
N	77	73	
(5, 12, 10)	5.97 (0.296)	6.27 (0.347)	-0.304 (0.456)
N	75	75	

Standard errors are in parentheses.

and race, differences are mostly driven by the non-state-risk treatment. Indeed, within male subjects and white subjects, participants are more risk averse under the non-state-risk treatment: differences with the state-risk treatment are 0.49 and 0.65 respectively. Perhaps, those population group may avoid those who obtain a lot just by a luck (such as winners for a lottery jackpot) more than those who obtain a lot by themselves (such as entrepreneurs). The results for gender are particularly interesting, because many existing studies show that women are more risk averse than men (see Eckel and Grossman (2008) for a survey). Our results suggest that risk attitudes toward others may function differently from standard risk attitudes toward oneself. In Table A.1 in the Appendix, we also show the interactions with these population characteristics and other experimental parameters. Results suggest that the behavioral differences between STEM- and non-STEM- major students come from the perception toward probabilities: STEM-students, who are supposed to have more exposure to probabilities, are more sensitive to the changes in p, the probability of others being rich. Finally, we also present means, standard errors and differences in the value of the switching point by each set of parameter in Appendix A.2 to A.4 for each demographic characteristic.

Table 7: Means and differences in switching points by gender, college major, and race

	(1) State-risk	(2) Non-State-risk	(1) + (2) Combined	(1) − (2) Difference
	Panel A: Gender			
Male	3.36	3.85	3.62	-0.49
	(0.128)	(0.147)	(0.098)	(0.196)*
Female	3.54	3.41	3.47	0.131
	(0.141)	(0.145)	(0.101)	(0.202)
Difference	-0.18	0.44	0.14	
	(0.189)	(0.209)*	(0.141)	
	Panel B: College Majors			
Non-STEM	3.31	3.56	3.44	-0.25
	(0.118)	(0.122)	(0.085)	(0.170)
STEM	3.71	3.86	3.78	-0.15
	(0.154)	(0.198)	(0.123)	(0.248)
Difference	-0.40	-0.30	-0.34	
	(0.194)*	(0.224)	(0.148)*	
	Panel C: Race			
Non-white	3.47	3.33	3.39	0.14
	(0.149)	(0.136)	(0.100)	(0.202)
White	3.45	4.09	3.73	-0.65
	(0.118)	(0.159)	(0.097)	(0.194)***
Difference	0.02	-0.77	-0.35	
	(0.188)	(0.209)***	(0.140)*	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Standard errors are in parentheses.

6 Discussion

Here, we extend existing theories described in section 4 to explore the potential mechanism in our experiment. We focus on the risk-averse behavior (absolute evaluation) and the behavioral difference between state-risk and non-state-risk treatments (relative evaluation).

6.1 Discussion on absolute evaluation

As in Section 4.1, we use the general utility function (4), assume that $g_C(\cdot) = g_R(\cdot) \equiv g(\cdot)$ and consider a parameter set $(e_D, X, e_{rH}, p) = (10, 5, 6, 0.9)$ in our experiment. In section 4.1, we consider an example of the linear form by Fehr and Schmidt (1999), in which the risk-loving behavior is predicted. To explain the risk-averse behavior observed in the experiment, we now consider the function g to be concave in the envy domain as shown in Figure 6.

Using equation (7), we determine the risk attitude in the example of Figure 6. The first term

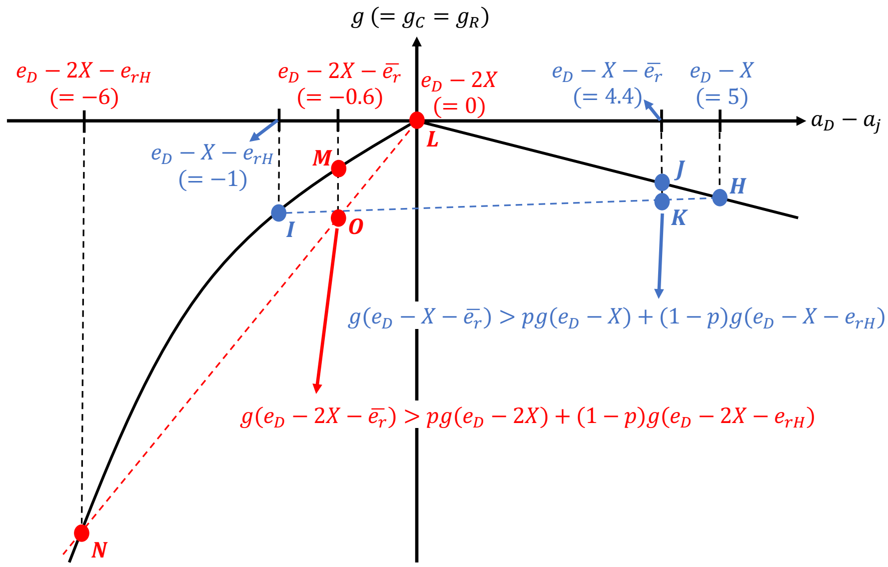


Figure 6: (Case 2) concave in envy domain and linear in guilt domain

of G in the equation (7) is the difference between the value of g at point M and the value at point O, which is positive. This is due to the concavity of g in the envy domain. The second term of G is the difference between the value at point J and that at point K, which is positive but smaller than the first term.[24] As shown in Figure 6, the first term of G is larger than the second, so that the dictator is risk averse toward others in this example. The more concave form of g in the envy domain implies more risk aversion. Bellemare et al. (2008) experimentally find that their subjects have an increasing and concave utility function in the envy domain in their ultimatum game under a non-risky decision environment. Therefore, the findings of Bellemare et al. (2008) are consistent with the risk-averse behaviors in our experiment.

The functional form in the guilt domain also affects risk attitude. In particular, the more convex the function in the guilt domain is, the larger the value of G is because the second term of G becomes smaller. Dictators become more risk averse in such cases.

6.2 Discussion on relative evaluation: state-risk vs non-state-risk

To explore the difference between state-risk and non-state-risk treatments, we consider two

[24]This, of course, depends on the concavity in the envy domain. As the envy domain becomes closer to linear, the utility function becomes similar to that of Fehr and Schmidt (1999), and the first term of G becomes smaller than the second.

potential mechanisms: *preference* and *cognitive bias* toward others.

6.2.1 Preference

First, let us consider a potential mechanism through decision-maker's *preferences* toward recipients using the form of Fehr and Schmidt (1999). Let α_k^T and β_k^T be degrees of guilt and envy toward recipient k in treatment $T \in \{S, N\}$, respectively.[25] We consider that the parameters of envy or guilt can vary depending on the situation (framing) in each treatment, even though the final outcome is exactly the same. We have the following prediction.

Prediction 2-1 *Suppose a dictator's preference is represented by the utility function* (5). *If a dictator's preference satisfies* $\alpha_k^S \neq \alpha_k^N$ *and/or* $\beta_k^S \neq \beta_k^N$ *for some recipient* k, $\overline{e_C^S} \neq \overline{e_C^N}$ *can be observed. Especially, if it holds that* $\alpha_{rH}^T > \alpha_{rH}^{T'}$ *and* $\beta_{rH}^T < \beta_{rH}^{T'}$, *then* $\overline{e_C^{T'}} > \overline{e_C^T}$ *can be observed, where* $T, T' \in \{S, N\}$ *and* $T \neq T'$.

Proof: See Appendix C.4 for the proof of the second statement. □

The first statement says that if parameters of envy or guilt vary depending on the treatment, the average values of switching points can differ between two treatments. The second statement shows an example of the first statement, which says that if the guilt concern is lower and the envy concern is higher toward recipient rH (recipient R with $e_R = e_{rH}$) in the treatment T' than in the treatment T, then the dictators show more risk-averse behaviors in the treatment T', that is $\overline{e_C^{T'}} > \overline{e_C^T}$.[26] If the dictators have such preferences, they are more likely to avoid making a transfer to the recipient R with $e_R = e_{rH}$ in the treatment T', because they feel more envy and less guilt toward them. Then, the dictators in the treatment T' will be more persistent in giving to recipient C, which leads to the higher switching points, $\overline{e_C^{T'}} > \overline{e_C^T}$.

An example of such preferences is presented in Figure 7. In the figure, even though the difference in an outcome between the dictator and the recipient rH, that is $a_D - a_{rH}$, is same (is equal to x) in both treatment T and T', the dictator suffers more losses from envy in the treatment T' than in the treatment T, that is $g(x|T) > g(x|T')$. Similarly, the dictator's losses from guilt are lower in the treatment T', that is $g(y|T') > g(y|T)$.

Now, let us consider what induces the difference in preferences by treatment. The first is the recipient's initial endowment. As we have provided multiple examples, people may prefer transferring

[25] Again, S means state-risk treatment and N means non-state-risk treatment.

[26] Cox, Friedman and Sadiraj (2008) introduce this formulation as "more altruistic than (MAT)."

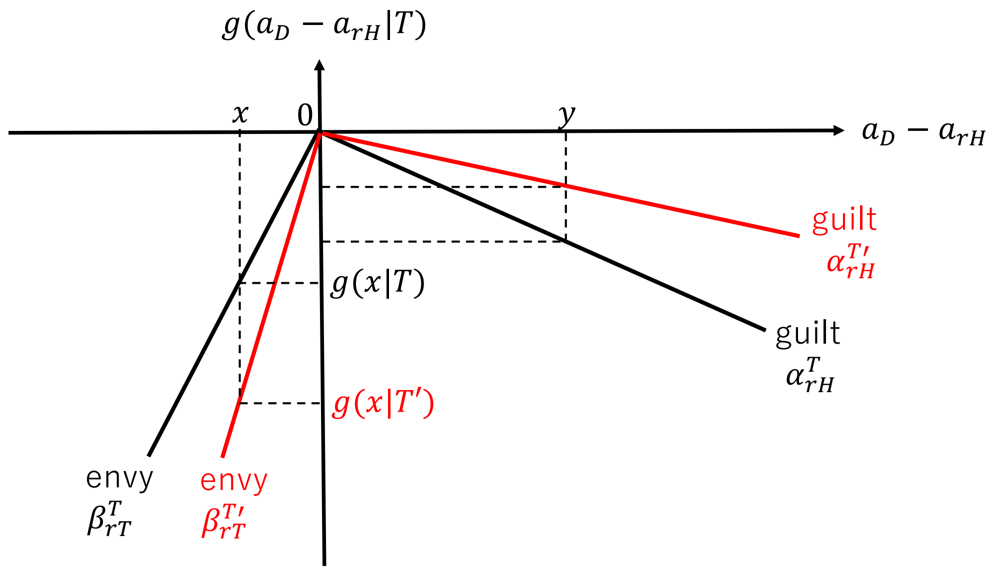


Figure 7: An example that the guilt concern is lower and the envy concern is higher toward recipient rH in the treatment T' than in the treatment T

to others with low initial endowment over transferring to others with high initial endowment. It implies that the parameters can depend on recipient's initial endowment: $\alpha_k = \alpha_k(e_k)$ and $\beta_k = \beta_k(e_k)$. Especially, it may hold that $\alpha^N_{rH} = \alpha_{rH}(0) > \alpha_{rH}(e_{rH}) = \alpha^S_{rH}$ and $\beta^N_{rH} = \beta(0) < \beta(e_{rH}) = \beta^S_{rH}$. That is, the dictator's guilt concern weakens and the envy concern strengthens in the state-risk treatment where the recipient has a higher initial endowment. In this case, by Prediction 2-1, we can observe $\overline{e^S_C} > \overline{e^N_C}$.

The second is the recipient's opportunity to make new money. People may avoid giving to a lucky person who won the lottery without their own effort, which implies that the parameters can vary depending on recipient's luck. Our framing in the non-state-risk treatment may give an impression that the recipient rH is a lucky person compared to the recipients in the state-risk treatment. In this case, $\alpha^S_{rH} > \alpha^N_{rH}$ and $\beta^S_{rH} < \beta^N_{rH}$ may hold. That is, the dictator's guilt concern weakens and the envy concern strengthens in the non-state-risk treatment. By Prediction 2-1, we can observe $\overline{e^N_C} > \overline{e^S_C}$. Our results for male and white subjects are more consistent with this second explanation.

6.2.2 Cognitive bias

Next, let us consider a potential mechanism through decision-maker's *cognitive bias* toward recipients. While each dictator knows the objective probability on the risky recipient, they may form own subjective probabilities, which may vary depending on the treatment. This captures a dictator's psychological evaluation of the true objective probability, whose concept comes from probability weighting in prospect theory (Kahneman and Tversky 1979). Let $\pi(s|T)$ be a dictator's subjective probability when the objective probability is s. We have the following prediction.

Prediction 2-2 *Suppose a dictator's preference is represented by the utility function (5). If a dictator's subjective probability satisfies $\pi(1-p|S) \neq \pi(1-p|N)$, then $\overline{e_C^S} \neq \overline{e_C^N}$ can be observed, where $1-p$ is an objective probability of $e_R = e_{rH}$. Especially, if it holds $\pi(1-p|T') > \pi(1-p|T)$, then $\overline{e_C^{T'}} > \overline{e_C^T}$ can be observed, where $T, T' \in \{S, N\}$ and $T \neq T'$.*

Proof: See Appendix C.5 for the proof of the second statement. □

The first statement says that if the subjective probability of $e_R = e_{rH}$ varies depending on the treatment, the average values of switching points can differ between two treatments. The second statement shows an example of the first statement, which says that if the subjective probability is higher in the treatment T' than in the treatment T, then the dictators are more risk averse in the treatment T', that is $\overline{e_C^{T'}} > \overline{e_C^T}$. If the dictators have such cognitive bias, they are more likely to avoid making a transfer to the recipient R in the treatment T', because they perceive a higher chance to feel envy in the treatment T'. Then, the dictators in the treatment T' will be more persistent in giving to recipient C, which leads to $\overline{e_C^{T'}} > \overline{e_C^T}$.

An example of such cognitive bias is presented in Figure 8. For the same objective probability $1-p$, the dictator has more bias in the treatment T' than in the treatment T, that is $\pi(1-p|T') > \pi(1-p|T)$, when $1-p$ is small. This probability weighting follows the idea of rank-dependent utility developed by Quiggin (1982) and Schmeidler (1989).[27]

As in the analysis of *preference*, we consider that recipient's initial endowment and luck may induce the difference in cognitive bias by treatment. For example, when a recipient's initial endowment is high, dictators may overreact to the high initial endowment and over-evaluate the objective probability $1-p$, which implies $\pi(1-p|S) > \pi(1-p|N)$. In this case, by Prediction 2-2, we can observe $\overline{e_C^S} > \overline{e_C^N}$. Moreover, when a recipient has a chance to win a lottery, dictators may overreact to the luck of the recipient and over-evaluate the objective probability $1-p$, which implies $\pi(1-p|N) > \pi(1-p|S)$. This captures that the dictators do not prefer a lucky person. In this case, by Prediction 2-2, we can observe $\overline{e_C^N} > \overline{e_C^S}$.

[27] However, we cannot naturally apply their idea to inequality aversion models, which poses theoretical challenges.

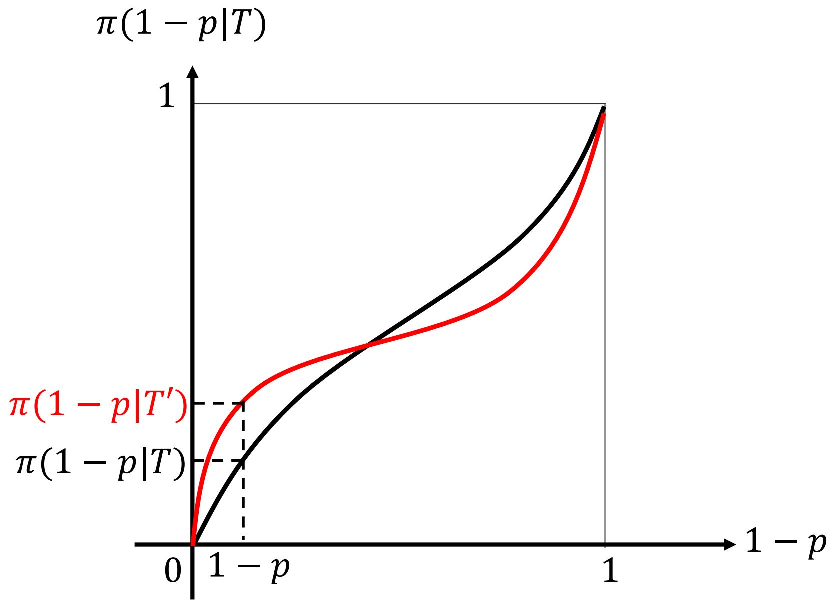


Figure 8: An example of cognitive bias for other's initial endowment

Before running our experiment, we predicted $\overline{e_C^S} > \overline{e_C^N}$ to be observed through the initial endowment effect, as naturally inferred by existing experimental papers (e.g. Eckel et al. 2005). However, as reported in the results, there is no difference between two treatments on average, and conversely $\overline{e_C^N} > \overline{e_C^S}$ is observed for male and white subjects. Our results imply that a dictator may want to avoid a lucky person who won the lottery without his own effort through preference or bias. Our framing in the non-state-risk treatment may provide a sense of unfairness with respect to the opportunity to make new money compared with that in state-risk treatment. Our experiment cannot distinguish whether the results are driven by the preference or bias. Identifying the source of behavior may be an important future research agenda.

7 Concluding remarks

In this paper, we propose a novel experiment to capture risks on others (ROO). While ROO appear in a wide range of economic activities, such as charitable giving and discrimination, previous studies have not investigated the theoretical and empirical properties of ROO. We theoretically define risk attitudes toward others under two dimensions: an *absolute* term, which characterizes one's risk attitudes toward others without comparing with other individuals, and a *relative* term, to compare risk attitudes between individuals.

We introduce a modified dictator game with three agents (a dictator, certain, and risky recipients) to capture ROO. In this environment, dictators face a trade-off between transferring to certain and risky recipients. In the experiment, we elicit the point at which dictators switch from transferring to a certain recipient to a risky one. At the switching point, dictators are indifferent between transferring to risky and safe recipients. From this experimental structure, we can identify the risk attitudes toward others in *absolute* terms to classify whether individuals are risk averse, risk neutral, or risk lover. We then compare the switching points between individuals to evaluate the risk attitudes toward others in *relative* terms. We compare the two treatments to disclose the effect of sources of ROO. In the "state-risk treatment," a risky recipient has a high initial endowment with a small probability, and in the "non-state-risk treatment," the risky recipient does not have any endowment for sure but has a small chance of receiving transfers from a third party.

We conduct experiments to examine ROO at the University of Maryland, College Park in 2018, and obtain following observations. First, for the risk aversion toward others in *absolute* terms, we find that people are consistently risk averse toward others. These results cannot be explained by Fehr and Schmidt (1999), which assume linearity in their utility functions. We suggest allowing curvature on the utility function as suggested by Bellemare et al. (2008), instead of the linear specification. Second, we do not find any difference in behaviors by the different sources of risks on average overall when we examine ROO in *relative* terms. Yet, within male subjects and white subjects, people are more risk averse in the non-state-risk treatment than in the state-risk treatment. This is a suggestive evidence that people perceive risks differently by states of risks, which is inconsistent to the current theories based on outcomes. As people avoid risks more in the non-state-risk treatment in our experiment, people perhaps want to avoid a lucky person who won the lottery without his own effort. Yet, we could not pin down underlying mechanisms behind such a behavior through our experiment, which can be an important research agenda in the future. Moreover, male, STEM-major and white subjects exhibit more risk aversion toward others than female subjects do in general. The result for gender implies that risk preferences towards others are fundamentally different from standard risk preferences, that women are more likely to exhibit risk-averse behaviors (Eckel and Grossman 2008).

We also note that different states of risks, compared in our experiment, are just examples of ROO. There are various different types of ROO. One may use our experimental framework and conduct some new experiment which has more specific context on types of risks and specifications of unfavorable groups as in Fong and Oberholzer-Gee (2011).

Bibliography

Akbaş, Merve, Dan Ariely, and Sevgi Yuksel, "When is inequality fair? An experiment on the effect of procedural justice and agency," *Journal of Economic Behavior & Organization*, 2019, *161*, 114–127.

Andersen, Steffen, Glenn W. Harrison, Morten Igel Lau, and E. Elisabet Rutström, "Elicitation using multiple price list formats," *Experimental Economics*, 2006, *9* (4), 383–405.

Andreoni, James and Charles Sprenger, "Uncertainty Equivalents: Testing the Limits of the Independence Axiom," NBER Working Papers 17342, National Bureau of Economic Research, Inc 2011.

Bellemare, Charles, Sabine Kröger, and Arthur Van Soest, "Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities," *Econometrica*, jul 2008, *76* (4), 815–839.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 2004, *94*, 991–1013.

Bohren, J Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope, "Inaccurate statistical discrimination," Technical Report, National Bureau of Economic Research 2020.

Bolton, Gary E and Axel Ockenfels, "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 2000, *90* (1), 166–193.

Brock, J Michelle, Andreas Lange, and Erkut Y Ozbay, "Dictating the Risks – Experimental Evidence on Norms of Giving in Risky Environments," *American Economic Review*, 2013, *103* (1), 415–437.

Cappelen, Alexander W., James Konow, Erik Ø. Sørensen, and Bertil Tungodden, "Just Luck: An Experimental Study of Risk-Taking and Fairness," *American Economic Review*, June 2013, *103* (4), 1398–1413.

Charness, Gary and Angelino Viceisza, "Three Risk-elicitation Methods in the Field - Evidence from Rural Senegal," *Review of Behavioral Economics*, July 2016, *3* (2), 145–171.

— **and Matthew Rabin**, "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 2002, *117* (3), 817–869.

— **and Uri Gneezy**, "What's in a name? Anonymity and social distance in dictator and ultimatum games," *Journal of Economic Behavior and Organization*, 2008, *68* (1), 29–35.

— **, — , and Alex Imas**, "Experimental methods: Eliciting risk preferences," *Journal of Economic Behavior & OrganizationOrganization*, 2013, *87*, 43–51.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Chen, Yan and Sherry Xin Li, "Group Identity and Social Preferences," *American Economic Review*, 2009, *99* (1), 431–457.

Cox, James C., Daniel Friedman, and Vjollca Sadiraj, "Revealed Altruism," *Econometrica*, 2008, *76* (1), 31–69.

Eckel, Catherine C. and Philip J. Grossman, "Altruism in Anonymous Dictator Games," *Games and Economic Behavior*, 1996, *16* (2), 181–191.

Eckel, Catherine C and Philip J Grossman, "Men, women and risk aversion: Experimental evidence," *Handbook of experimental economics results*, 2008, *1*, 1061–1073.

Eckel, Catherine C., Philip J. Grossman, and Rachel M. Johnston, "An experimental test of the crowding out hypothesis," *Journal of Public Economics*, 2005, *89* (8 (Special issue)), 1543–1560.

Exley, Christine L., "Excusing selfishness in charitable giving: The role of risk," *Review of Economic Studies*, 2016, *83* (2), 587–628.

Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, aug 1999, *114* (3), 817–868.

Fong, Christina M. and Felix Oberholzer-Gee, "Truth in giving: Experimental evidence on the welfare effects of informed giving to the poor," *Journal of Public Economics*, 2011, *95* (5-6), 436–444.

Fudenberg, Drew and David K Levine, "Fairness, risk preferences and independence: Impossibility theorems," *Journal of Economic Behavior and Organization*, 2011, *81* (2), 606–612.

Gneezy, Uri, Erizabeth A. Keenan, and Ayelet Gneezy, "Avoiding overhead aversion in charity," *Science*, 2014, *346* (6209), 632–635.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Iriberri, Nagore and Pedro Rey-Biel, "The role of role uncertainty in modified dictator games," *Experimental Economics*, 2011, *14* (2), 160–180.

_ and _ , "Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do?," *Quantitative Economics*, 2013, *4* (3), 515–547.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–292.

Karlan, R Dean and Daniel H Wood, "The effect of effectiveness: Donor response to aid effectiveness in a direct mail fundraising experiment," *Journal of Behavioral and Experimental Economics*, 2016, *66*, 1–8.

Krawczyk, Michał, "A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution," *Journal of Public Economics*, 2010, *94* (1), 131–141.

López-Vargas, Kristian, "Risk Attitudes and Fairness: Theory and Experiment," *Job market paper*, 2014.

Mesa-Vázquez, Ernesto, Ismael Rodriguez-Lara, and Amparo Urbano, "Standard vs random dictator games: On the effects of role uncertainty and framing on generosity," *Economics Letters*, 2021, *206*, 109981.

Mollerstrom, Johanna, Bjørn-Atle Reme, and Erik Ø. Sørensen, "Luck, choice and responsibility — An experimental study of fairness views," *Journal of Public Economics*, 2015, *131*, 33–40.

Neumark, David, "Experimental Research on Labor Market Discrimination," *Journal of Economic Literature*, sep 2018, *56* (3), 799–866.

Quiggin, John, "A Theory of Anticipated Utility," *Journal of Economic Behavior & Organization*, 1982, *3* (4), 323–343.

Rau, Holger and Stephan Müller, "Decisions under Uncertainty in Social Contexts," *Working Paper*, 2017.

Saito, Kota, "Social Preferences under Risk: Equality of Opportunity versus Equality of Outcome," *American Economic Review*, 2013, *103* (7), 3084–3101.

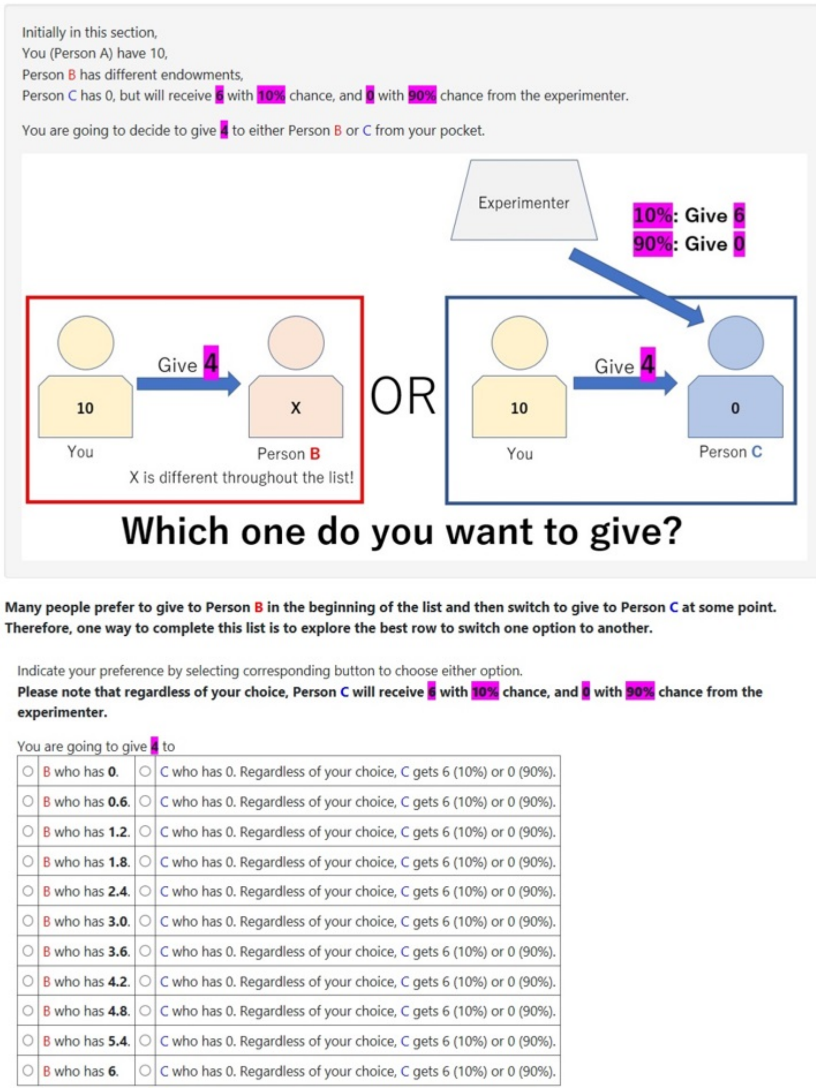
Schmeidler, David, "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 1989, *57* (3), 571.

Walkowitz, Gari, "Dictator game variants with probabilistic (and cost-saving) payoffs: A systematic test," *Journal of Economic Psychology*, 2021, *85*, 102387.

Yu, Chi Wai, Y. Jane Zhang, and Sharon Xuejing Zuo, "Multiple Switching and Data Quality in the Multiple Price List," *The Review of Economics and Statistics*, 03 2021, *103* (1), 136–150.

Appendix

A Additional figures and tables


Initially in this section,
You (Person A) have 10,
Person B has different endowments,
Person C has 0, but will receive 6 with 10% chance, and 0 with 90% chance from the experimenter.

You are going to decide to give 4 to either Person B or C from your pocket.

Many people prefer to give to Person B in the beginning of the list and then switch to give to Person C at some point. Therefore, one way to complete this list is to explore the best row to switch one option to another.

Indicate your preference by selecting corresponding button to choose either option.
Please note that regardless of your choice, Person C will receive 6 with 10% chance, and 0 with 90% chance from the experimenter.

You are going to give 4 to

○	B who has **0**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **0.6**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **1.2**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **1.8**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **2.4**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **3.0**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **3.6**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **4.2**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **4.8**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **5.4**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).
○	B who has **6**.	○	C who has 0. Regardless of your choice, C gets 6 (10%) or 0 (90%).

Figure A.1: Screenshot of the actual experiment: Non-state-risk treatment

Table A.1: Regression analyses with group heterogeneity. Dependent variable: Switching point of dictator's choice

	(1) Gender	(2) STEM major	(3) Race
Other risk	-0.672	-0.231	0.0458
	(0.376)	(0.289)	(0.329)
e_c	0.437***	0.481***	0.469***
	(0.0389)	(0.0328)	(0.0419)
percent of rich	0.0499***	0.0427***	0.0446***
	(0.00895)	(0.00934)	(0.0116)
Female × Other-Risk	0.792		
	(0.479)		
Female × Endowment of Rich	0.0776		
	(0.0537)		
Female × Prob. being Rich	0.000738		
	(0.0153)		
Gender	-0.853		
	(0.493)		
STEM-major× Other-Risk		-0.215	
		(0.565)	
STEM-major× Endowment of Rich		-0.0296	
		(0.0585)	
STEM-major× Prob. being Rich		0.0233	
		(0.0147)	
STEM-major		0.0537	
		(0.549)	
White× Other-Risk			-0.651
			(0.500)
White× Endowment of Rich			0.00356
			(0.0547)
White× Prob. being Rich			0.0114
			(0.0148)
White			0.0492
			(0.511)
Constant	0.256	-0.133	-0.134
	(0.358)	(0.305)	(0.368)
Observations	1209	1217	1217
R^2	0.277	0.269	0.272

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2: Means and differences for each parameter by gender

	State-risk			Non-state-risk		
Parameters (x, e_r, p)	Male	Female	Difference	Male	Female	Difference
(3, 6, 10)	3.08 (0.211)	3.50 (0.172)	-0.419 (0.278)	3.42 (0.264)	3.40 (0.223)	0.018 (0.378)
N	43	37		32	30	
(3, 6, 20)	3.26 (0.197)	3.46 (0.195)	-0.20 (0.280)	3.47 (0.200)	3.62 (0.258)	-0.15 (0.280)
N	46	32		42	34	
(3, 12, 10)	5.92 (0.390)	6.04 (0.594)	-0.12 (0.680)	6.18 (0.408)	6.10 (0.707)	0.08 (0.772)
N	45	29		54	24	
(5, 6, 5)	2.50 (0.283)	2.47 (0.226)	0.03 (0.371)	2.92 (0.280)	2.76 (0.304)	0.16 (0.416)
N	46	35		41	35	
(5, 6, 10)	2.69 (0.223)	3.24 (0.250)	-0.55 (0.334)	3.59 (0.282)	3.13 (0.213)	0.46 (0.378)
N	41	34		44	32	
(5, 6, 20)	3.21 (0.229)	3.31 (0.228)	-0.093 (0.325)	3.81 (0.223)	3.21 (0.177)	0.60 (0. 302)
N	39	33		42	31	
(5, 12, 5)	4.43 (0.510)	5.53 (0.467)	-1.11 (0.706)	5.69 (0.507)	5.32 (0.429)	0.37 (0.705)
N	42	34		43	30	
(5, 12, 10)	5.35 (0.391)	6.62 (0.438)	-1.26 (0.586)	6.37 (0.494)	6.12 (0.458)	0.25 (0.712)
N	39	35		45	30	

Standard errors are in parentheses.

Table A.3: Means and differences for each parameter by college major (STEM or not)

	State-risk			Non-state-risk		
Parameters (x, e_r, p)	STEM	Non-STEM	Difference	STEM	Non-STEM	Difference
(3, 6, 10)	3.15 (0.232)	3.38 (0.171)	-0.217 (0.282)	3.77 (0.406)	3.27 (0.190)	0.51 (0.394)
N	32	49		21	52	
(3, 6, 20)	3.37 (0.227)	3.36 (0.175)	0.016 (0.288)	3.35 (0.227)	3.38 (0.175)	0.016 (0.283)
N	31	46		23	55	
(3, 12, 10)	6.04 (0.594)	5.92 (0.390)	0.12 (0.680)	6.10 (0.707)	6.18 (0.408)	-0.08 (0.772)
N	29	45		24	54	
(5, 6, 5)	2.50 (0.300)	2.50 (0.231)	-0.002 (0.375)	2.76 (0.426)	2.89 (0.230)	-0.13 (0.447)
N	30	47		25	56	
(5, 6, 10)	2.83 (0.259)	3.06 (0.217)	-0.19 (0.348)	3.54 (0.377)	3.32 (0.215)	0.22 (0.406)
N	29	47		23	53	
(5, 6, 20)	3.11 (0.194)	3.39 (0.280)	-0.28 (0.330)	3.71 (0.335)	3.50 (0.168)	0.22 (0. 338)
N	28	45		21	52	
(5, 12, 5)	4.56 (0.550)	5.16 (0.454)	-0.60 (0.719)	4.97 (0.778)	5.76 (0.370)	-0.80 (0.762)
N	30	47		21	52	
(5, 12, 10)	5.96 (0.476)	5.97 (0.382)	-0.13 (0.609)	6.25 (0.708)	6.28 (0.391)	-0.032 (0.748)
N	30	45		24	51	

Standard errors are in parentheses.

Table A.4: Means and differences for each parameter by race

	State-risk			Non-state-risk		
Parameters (x, e_r, p)	Non-White	White	Difference	Non-White	White	Difference
(3, 6, 10)	3.63 (0.208)	3.00 (0.175)	0.63 (0.270)	3.35 (0.225)	3.47 (0.281)	-0.12 (0.361)
N	36	45		36	37	
(3, 6, 20)	3.61 (0.247)	3.15 (0.147)	0.474 (0.274)	3.56 (0.212)	3.47 (0.236)	0.11 (0.320)
N	34	43		40	38	
(3, 12, 10)	6.04 (0.543)	5.91 (0.414)	0.12 (0.670)	6.12 (0.488)	6.18 (0.519)	-0.06 (0.772)
N	32	42		39	39	
(5, 6, 5)	2.68 (0.293)	2.36 (0.230)	0.32 (0.367)	2.68 (0.293)	2.36 (0.230)	0.32 (0.367)
N	30	47		25	56	
(5, 6, 10)	3.20 (0.250)	2.78 (0.220)	0.44 (0.335)	3.26(0.259)	3.54 (0.273)	-0.27 (0.187)
N	39	37		32	44	
(5, 6, 20)	3.33 (0.181)	3.78 (0.240)	-0.44 (0.302)	3.23 (0.181)	3.14 (0.240)	-0.08 (0. 302)
N	36	37		31	42	
(5, 12, 5)	5.26 (0.569)	4.66 (0.438)	0.60 (0.706)	5.40 (0.420)	5.67 (0.549)	-0.276 (0.695)
N	34	43		36	37	
(5, 12, 10)	6.56 (0.502)	5.53 (0.346)	1.04 (0.590)	97 (0.486)	6.55 (0.495)	-0.59 (0.695)
N	32	43		36	39	

Standard errors are in parentheses.

B Experimental protocol for non-state-risk treatment

Page 1

Thank you for your participation. Before we start the experiment, please read the paper titled "consent form" on the table carefully. Once you finish reading and agree with the contents, please sign the document. Please do not forget to write initials and date on the top of the page.

Please do not talk to each other, and please do not eat or drink inside the lab. Also, please do not use your phones or any other electric devices. If you have any question or problems, please raise your hand.

We are going to start the experiment. Please look at the monitor. This is an experiment in decision-making. You are going to make multiple choices. There will be two parts. The first part has 8 rounds, and the second part has 6 rounds. Entire experiment will take less than an hour. From the next page, we will explain how you are going to make your decisions in the first part. After the first part is complete, we will go through the instructions for the second part.

All earnings for this experiment will be in Experiment Currency Unit (ECU). At the end of the experiment, ECU will be converted to US Dollars at a rate of 1ECU=$0.80. You will be paid a guaranteed show-up fee of $7 in addition to your earnings for this experiment if you complete the experiment. You will be paid your earnings privately in cash before you leave the lab today.

Please click "next" once you understand the procedure.

Page 2

We are going to ask some clarification questions after each instruction. Please listen to the instruction carefully. In the first part, you are going to be matched with 2 other people in this room. You will not know who you are matched with. In a group, there are 2 roles, (Person A), and Receivers (Person B and C). You are going to be assigned one of those them in equal chance. Person A can decide his own and others' earnings. Decisions by Persons B and C do not matter for the final earnings. Their earnings are exclusively based on the decisions of Person A.

At the time of decision making, you don't know which role you are actually assigned, but you are going to make choices as if you are Person A, the decision maker. At the end of the experiment, we will notify you of your actual role. If your role turns out to be Person A, one out of 8 rounds will be randomly selected, and your choice will be implemented. If your role turns out to be Person B or C, then your choices do not matter, but you may receive some earnings depending on the choice of Person A in your group.

Please answer the questions with your best knowledge. Please click next once you answer all the questions.

Page 3

Now, let's see an example of your task. Please keep in mind that the unit of all the payoffs and numbers in this experiment is ECU, which will be converted to US Dollars at a rate of 1ECU=$0.80.

In each round, you will make 11 decisions. Each decision always involves a choice between two options, "giving to Person B" or "giving to Person C". Here is an example. Please see the picture below. I will take a few second to check the picture. You (Person A) have 10, and you are going to give 4 to either Person B or C.

(A part of Figure A.1)

In this case, you (Person A) have 10. Person B has X, and Person C has 0. Regardless of your choice, Person C will receive 6 with 10% chance, and 0 with 90% chance from the experimenter. Person B's initial endowment (X) is different in each decision.

If you (Person A) choose to give to B, the final earnings of Person B will be 4+X. Person C will not receive any transfer from you, so that Person C's final earnings will be 6 with 10% chance and 0 with 90% chance. If you (Person A) choose to give to C, the final earnings of Person C will be 10 with 10% chance and 4 with 90% chance. Person B will not receive any transfer from you, so that Person B's final earnings are X.

Please make sure that the person you decided not to make transfer DOES NOT lose other earnings from other sources. For instance, if you decide to give to B, C still earns 6 with 10% chance and 0 with 90% chance from the experimenter. If you decide to give to C, B still earns X which was her initial endowment.

Note that, no matter what, you (Person A) must give 4 to someone, so your (person A's) final earnings are the same whether you choose to give to B or C.

Please answer the questions with your best knowledge. Please click next once you answer all the questions.

Page 4

The following is the actual screen you will see in each round. Please check the picture for a while.

(The same picture as Figure A.1)

You are required to make decisions for all 11 choices. As you see, the right side is always same for all 11 choices. Moreover, regardless of your choice, the amount you are going to give is the same for all choices. On the left side, the ONLY differences among 11 choices are the endowment of Person B, which increases by some amounts as you go down the list. In this example, it increases by 0.6.

You are going to repeat this list 8 times (rounds) in total, each time with 11 choices. Final earnings for Part 1 will be determined in the following way:

1. Your actual role will be announced.
2. Among the 8 rounds, one of them is selected by a computer in the equal chance.
3. Among the 11 choices in the round chosen by a computer, one of them is selected by a computer in the equal chance. The choice of the person assigned to be Person A in your group will be implemented and your earnings for Part 1 will be based on that decision.

Please click "next" once you understand the procedure.

Page 5

Before starting the main experiment, let's experience a practice session! This is to make sure you understand how the decision environment works. You will not be paid for any of the decisions in this practice session.

Please make decisions by clicking following 11 choices. Please click "next" once you make all the choices.

Page 6

Real decisions start from the next page. There are 8 rounds in total. Your final payoffs are going to be determined by the choices of Person A in your group. Please be careful that various numbers (e.g. amount you are going to give to others) are going to change by each round. Numbers highlighted by purple in the following picture will change.

(Picture here)

It is in your best interest to behave in each decision as if it is the decision for which you will be paid. This is because you are making decisions in each of the following 8 rounds as if you were Person A. If you are ready, please press "Next", and start making your choices.

C Proofs of Propositions

C.1 Preliminary

This section presents some key equations and notations, which are used to show the existence and uniqueness of the switching point, and to prove Proposition 1 and additional propositions which support Prediction 2-1 and 2-2.

The utility function (3) is rewritten as follows:

$$u^{FS}(a) = a_D + \sum_{k\in\{C,R\}} \Big(\beta_k - \sigma(a_D - a_k)(\alpha_k + \beta_k)\Big)(a_D - a_k), \tag{C.1}$$

where $\sigma(x) = \mathbb{1}(x \geq 0)$. Using (C.1) and (5), the utility of choosing lottery $L_k^T(e_C)$ ($k \in \{C, R\}$, $T \in \{S, N\}$), which is defined in Section 3.1, is rewritten as follows:

$$\begin{aligned} U^{Saito}(L_k^T(e_C)) \;&=\; (e_D - X) + \Big(\beta_C - \sigma(d(C,k))(\alpha_C + \beta_C)\Big)d(C,k) \\ &+\; \delta\Big(\beta_r - \sigma(\mathbb{E}_{L_k^T(e_C)}[d(R,k)])(\alpha_r + \beta_r)\Big)\mathbb{E}_{L_k^T(e_C)}[d(R,k)] \\ &+\; (1-\delta)\mathbb{E}_{L_k^T(e_C)}\left[\Big(\beta_R - \sigma(d(R,k))(\alpha_R + \beta_R)\Big)d(R,k)\right], \end{aligned} \tag{C.2}$$

where $d(l,k) = a_D - a_{lk}$ is the difference in the final allocations between the dictator and the recipient, with a_{lk} being the final allocation of recipient $l \in \{C, R\}$ when the dictator gives X to $k \in \{C, R\}$. α_r is the dictator's *ex-ante* guilt parameter toward risky recipient R. In other words, it is the guilt parameter for risky recipient R who is expected to have $\overline{e_r}$ initially or by the transfer. α_R is the dictator's *ex-post* guilt parameter to risky recipient R; this becomes $\alpha_R = \alpha_{rL}$ when recipient R has $e_{rL}(= 0)$ (with probability p), and $\alpha_R = \alpha_{rH}$ when recipient R has e_{rH} initially or by transfer (with probability $1-p$). β_r and β_R are defined similarly.[28]

As an example, let us consider parameters $(e_D, X, e_{rH}, p) = (10, 5, 6, 0.9)$, and a lottery $L_R^T(e_C)$,

[28] We distinguish recipient R from ex-ante and ex-post perspectives, and assign a parameter to each of them. Of course, it's possible those parameters coincide, that is $\alpha_r = \alpha_{rL} = \alpha_{rH}$ or $\beta_r = \beta_{rL} = \beta_{rH}$.

where a dictator gives X to a recipient R. When $k = R$, we can see the followings:

$$
\begin{aligned}
d(C,R) &= (e_D - X) - e_C = 5 - e_C, \\
d(R,R) &= (e_D - X) - (e_{rL} + X) = 0 \quad \textit{with prob}\ p = 0.9, \\
d(R,R) &= (e_D - X) - (e_{rH} + X) = -6 \quad \textit{with prob}\ 1-p = 0.1, \\
\mathbb{E}_{L_R^T(e_C)}[d(R,R)] &= p\{(e_D - X) - (e_{rL} + X)\} + (1-p)\{(e_D - X) - (e_{rH} + X)\} = -0.6, \\
\sigma(\mathbb{E}_{L_R^T(e_C)}[d(R,R)]) &= \sigma(-0.6) = 0.
\end{aligned}
$$

Here, we consider $e_C = 1$ as an example. Then, $\sigma(d(C,R)) = \sigma(5-1) = 1$. Therefore, we can calculate $U^{Saito}(L_R^T(e_C))$ as follows:

$$
\begin{aligned}
U^{Saito}(L_R^T(e_C)) &= (e_D - X) + \Big(\beta_C - \sigma(d(C,R))(\alpha_C + \beta_C)\Big) d(C,R) \\
&+ \delta\Big(\beta_r - \sigma(\mathbb{E}_{L_R^T(e_C)}[d(R,R)])(\alpha_r + \beta_r)\Big)\mathbb{E}_{L_R^T(e_C)}[d(R,R)] \\
&+ (1-\delta)\mathbb{E}_{L_R^T(e_C)}\Bigg[\Big(\beta_R - \sigma(d(R,R))(\alpha_R + \beta_R)\Big) d(R,R)\Bigg] \\
&= (10-5) + \Big(\beta_C - 1 \times (\alpha_C + \beta_C)\Big) \times 4 \\
&+ \delta\Big(\beta_r - 0 \times (\alpha_r + \beta_r)\Big) \times (-0.6) \\
&+ (1-\delta)\Bigg[0.9 \times \Big((\beta_{rL} - 1 \times (\alpha_{rL} + \beta_{rL})) \times 0\Big) + 0.1 \times \Big((\beta_{rH} - 0 \times (\alpha_{rH} + \beta_{rH})) \times (-6)\Big)\Bigg] \\
&= 5 - 4\alpha_C - 0.6\delta\beta_r - 0.6(1-\delta)\beta_{rH}.
\end{aligned}
$$

Now, we define $V(e_C|T) \equiv U^{Saito}(L_C^T(e_C)) - U^{Saito}(L_R^T(e_C))$ for each $T \in \{S, N\}$ as the difference in utilities transferring to recipients C and R. Then, we can see:

$$
V(e_C|T) = v_C(e_C|T) - \delta v_r(\overline{e_r}|T) - (1-\delta)\mathbb{E}[v_R(e_R|T)], \tag{C.3}
$$

where[29]

$$
\begin{aligned}
v_C(e_C|T) &\equiv \Big(\beta_C - \sigma(d(C,C))(\alpha_C+\beta_C)\Big)d(C,C) - \Big(\beta_C - \sigma(d(C,R))(\alpha_C+\beta_C)\Big)d(C,R),\\
v_r(\overline{e_r}|T) &\equiv \Big(\beta_r - \sigma(\mathbb{E}_{L_R^T(e_C)}[d(R,R)])(\alpha_r+\beta_r)\Big)\mathbb{E}_{L_R^T(e_C)}[d(R,R)]\\
&\quad -\Big(\beta_r - \sigma(\mathbb{E}_{L_C^T(e_C)}[d(R,C)])(\alpha_r+\beta_r)\Big)\mathbb{E}_{L_C^T(e_C)}[d(R,C)],\\
v_R(e_R|T) &\equiv \Big(\beta_R - \sigma(d(R,R))(\alpha_R+\beta_R)\Big)d(R,R) - \Big(\beta_R - \sigma(d(R,C))(\alpha_R+\beta_R)\Big)d(R,C).
\end{aligned}
$$

Equivalently,

$$
v_C(e_C|T) = \begin{cases} \alpha_C X & if \;\; e_C < e_D - 2X \\ \alpha_C(e_D - X - e_C) + \beta_C(e_D - 2X - e_C) & if \;\; e_D - 2X \le e_C \le e_D - X \\ -\beta_C X & if \;\; e_D - X < e_C \end{cases} \tag{C.4}
$$

$$
v_r(\overline{e_r}|T) = \begin{cases} \alpha_r X & if \;\; \overline{e_r} < e_D - 2X \\ \alpha_r(e_D - X - \overline{e_r}) + \beta_r(e_D - 2X - \overline{e_r}) & if \;\; e_D - 2X \le \overline{e_r} \le e_D - X \\ -\beta_r X & if \;\; e_D - X < \overline{e_r} \end{cases} \tag{C.5}
$$

$$
v_R(e_R|T) = \begin{cases} \alpha_R X & if \;\; e_R < e_D - 2X \\ \alpha_R(e_D - X - e_R) + \beta_R(e_D - 2X - e_R) & if \;\; e_D - 2X \le e_R \le e_D - X \\ -\beta_R X & if \;\; e_D - X < e_R \end{cases} \tag{C.6}
$$

where $e_R = e_{rL}(=0)$ with probability p and $e_R = e_{rH}$ with probability $1-p$. $v_C(e_C|T)$ is the net utility from certain recipient C, when the dictator make a transfer to recipient C. $v_r(\overline{e_r}|T)$ is the net utility from risky recipient R in the sense of *ex-ante* fairness, and $v_R(e_R|T)$ is the net utility from recipient R in the sense of *ex-post* fairness, when the dictator makes a transfer to recipient R.

[29] For the second term, $v_r(\overline{e_r}|T)$, note that $\mathbb{E}_{L_C^T(e_C)}[d(R,C)] = p\{(e_D-X)-0\}+(1-p)\{(e_D-X)-e_{rH}\} = e_D-X-\overline{e_r}$. Similarly, $\mathbb{E}_{L_R^T(e_C)}[d(R,R)] = e_D - 2X - \overline{e_r}$. Therefore, the second term depends on $\overline{e_r}$.

C.2 Existence and uniqueness of the switching point

In this subsection, we show the existence and uniqueness of the switching point under standard social preferences, as summarized in Proposition C.1. Since our theoretical analyses evaluate the properties of the switching point, existence and uniqueness are key conditions to justify our approach.

It is essential for the proof that $v_C(e_C|T)$, which is the net utility from transferring to a certain recipient C that has e_C for certain, is a decreasing function in e_C.

Proposition C.1 *Suppose $e_{rL} = 0$, $e_D - 2X \geq 0$, $e_D - 2X < e_{rH}$ and $e_D - X > \overline{e_r}$. Also suppose a dictator's preference is represented by utility function (C.2) by Saito (2013), and $\alpha_C = \alpha_r = \alpha_{rL} = \alpha_{rH} \equiv \alpha$, $\beta_C = \beta_r = \beta_{rL} = \beta_{rH} \equiv \beta$ (anonymity assumption). Then, there exists a unique switching point $e_C^T \in (0, e_{rH})$ such that $L_C^T(e_C^T) \sim L_R^T(e_C^T)$.*

Proof: From equations (C.3), (C.4), (C.5), (C.6), and the anonymity assumption, we can see

$$V(e_C|T) \equiv U^{Saito}(L_C^T(e_C)) - U^{Saito}(L_R^T(e_C)) = v_C(e_C|T) - \delta v_r(\overline{e_r}|T) - (1-\delta)\mathbb{E}[v_R(e_R|T)]. \quad \text{(C.7)}$$

Since v_C is continuous and weakly decreasing in e_C, and v_r and v_R do not depend on e_C, $V(e_C|T)$ is continuous and weakly decreasing in e_C (an example is shown in Figure C.1). Therefore, it suffices to show that $(i) V(0|T) > 0$ and $(ii) V(e_{rH}|T) < 0$. In the following, we omit notation T.

Using equations (C.7), (C.4), (C.5), (C.6), $e_{rL} = 0$, and assumptions $e_D - 2X \geq 0$, $e_D - 2X < e_{rH}$ and $e_D - X > \overline{e_r}$, we obtain the followings:

The proof of (i):

$$\begin{aligned}
V(0) &= v_C(0) - \delta v_r(\overline{e_r}) - (1-\delta)\{p v_R(e_{rL}) + (1-p)v_R(e_{rH})\} \\
&\geq \alpha X - \delta\alpha X - (1-\delta)\{p\alpha X + (1-p)(\alpha(e_D - X - e_{rH}) + \beta(e_D - 2X - e_{rH}))\} \\
&= (1-\delta)(1-p)\{\alpha X - \alpha(e_D - X - e_{rH}) - \beta(e_D - 2X - e_{rH})\} \\
&= -(1-\delta)(1-p)(\alpha+\beta)(e_D - 2X - e_{rH}) > 0.
\end{aligned}$$

In the second line, $v_r(\overline{e_r})$ is at most αX when $e_D - X > \overline{e_r}$, and $v_R(e_{rH})$ is at most $\alpha(e_D - X - e_{rH}) + \beta(e_D - 2X - e_{rH})$ when $e_D - 2X < e_{rH}$.

The proof of (ii):

$$
\begin{aligned}
V(e_{rH}) &= v_C(e_{rH}) - \delta v_r(\overline{e_r}) - (1-\delta)\left\{p v_R(e_{rL}) + (1-p) v_R(e_{rH})\right\} \\
&= (1-(1-\delta)(1-p)) v_C(e_{rH}) - \delta v_R(\overline{e_r}) - (1-\delta) p v_r(e_{rL}) \quad (\because\ v_C(e_{rH}) = v_R(e_{rH})) \\
&\leq (1-(1-\delta)(1-p))\{\alpha(e_D - X - e_{rH}) + \beta(e_D - 2X - e_{rH})\} \\
&\quad -\delta\left\{\alpha(e_D - X - \overline{e_r}) + \beta(e_D - 2X - \overline{e_r})\right\} - (1-\delta) p \alpha X \\
&= (1-(1-\delta)(1-p))(\alpha+\beta)(e_D - 2X - e_{rH}) + (1-(1-\delta)(1-p))\alpha X \\
&\quad -\delta(\alpha+\beta)(e_D - X - \overline{e_r}) + \delta\beta X - (1-\delta) p \alpha X \\
&= (p(1-\delta)+\delta)(\alpha+\beta)(e_D - 2X - e_{rH}) - \delta(\alpha+\beta)(e_D - 2X - \overline{e_r}) \\
&= p(1-\delta)(\alpha+\beta)(e_D - 2X - e_{rH}) - \delta(\alpha+\beta)(e_{rH} - \overline{e_r}) < 0
\end{aligned}
$$

In the third line, $v_C(e_{rH})$ is at most $\alpha(e_D - X - e_{rH}) + \beta(e_D - 2X - e_{rH})$ when $e_D - 2X < e_{rH}$, and in the fourth line $v_r(\overline{e_r})$ is at least $\alpha(e_D - X - \overline{e_r}) + \beta(e_D - 2X - \overline{e_r})$ when $e_D - X > \overline{e_r}$.

□

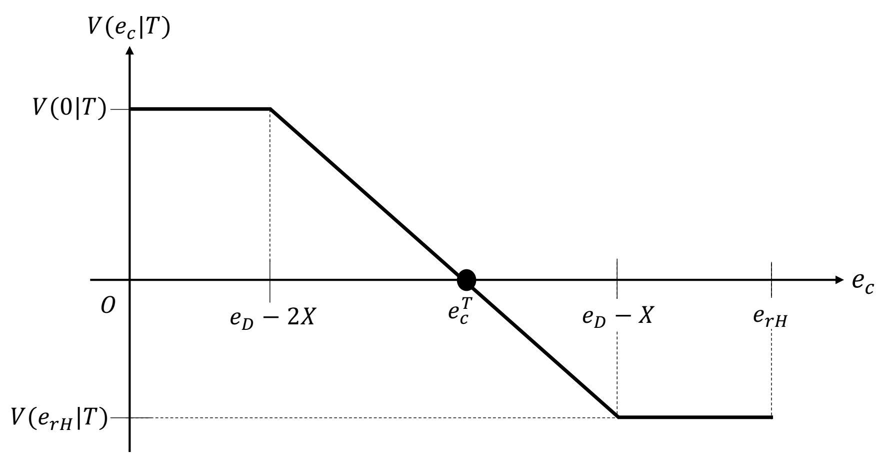


Figure C.1: An example of a unique cutoff in the case of $e_D - X < e_{rH}$

C.3 Proof of proposition 1

Proof: When a dictator has utility function (3), we can rewrite equation (6) (G in Lemma 2) as follows:

$$
\begin{aligned}
G &= g(e_D - 2X - \overline{e_r}) - g(e_D - X - \overline{e_r}) \\
&\quad + \{pg(e_D - 2X) + (1-p)g(e_D - 2X - e_{rH})\} - \{pg(e_D - X) + (1-p)g(e_D - X - e_{rH})\} \\
&= v_C(\overline{e_r}|T) - \mathbb{E}[v_R(e_R|T)] \\
&= v_C(\overline{e_r}|T) - pv_R(0|T) - (1-p)v_R(e_{rH}|T), \qquad \text{(C.8)}
\end{aligned}
$$

where $v_C(\overline{e_r}|T)$ and $v_R(e_R|T)$ are defined in Appendix C.1. Moreover, based on the anonymity assumption, we assume $\alpha_C = \alpha_r = \alpha_{rL} = \alpha_{rH} \equiv \alpha$, $\beta_C = \beta_r = \beta_{rL} = \beta_{rH} \equiv \beta$. Therefore, each component of equation (C.8) is as follows:

$$
v_C(\overline{e_r}|T) = \begin{cases} \alpha X & if \;\; \overline{e_r} < e_D - 2X \\ \alpha(e_D - X - \overline{e_r}) + \beta(e_D - 2X - \overline{e_r}) & if \;\; e_D - 2X \leq \overline{e_r} \leq e_D - X \\ -\beta X & if \;\; e_D - X < \overline{e_r} \end{cases} \qquad \text{(C.9)}
$$

$$
v_R(0|T) = \begin{cases} \alpha X & if \;\; 0 < e_D - 2X \\ \alpha(e_D - X) + \beta(e_D - 2X) & if \;\; e_D - 2X \leq 0 \leq e_D - X \\ -\beta X & if \;\; e_D - X < 0 \end{cases} \qquad \text{(C.10)}
$$

$$
v_R(e_{rH}|T) = \begin{cases} \alpha X & if \;\; e_{rH} < e_D - 2X \\ \alpha(e_D - X - e_{rH}) + \beta(e_D - 2X - e_{rH}) & if \;\; e_D - 2X \leq e_{rH} \leq e_D - X \\ -\beta X & if \;\; e_D - X < e_{rH} \end{cases} \qquad \text{(C.11)}
$$

Note that, by the assumption $e_D - 2X \geq 0$, it holds that $v_R(0|T) = \alpha X$. Additionally, based on the assumption $e_D - 2X < e_{rH}$, it holds that $v_R(e_{rH}|T) = \alpha(e_D - X - e_{rH}) + \beta(e_D - 2X - e_{rH})$ or $-\beta X$.

Now, we characterize the sign of G by $\overline{e_r}$. Let us consider the following three cases: $(a) e_D - X < \overline{e_r}$, $(b) e_D - 2X \leq \overline{e_r} \leq e_D - X$, and $(c) \overline{e_r} < e_D - 2X$.

(a): In this case, $v_C(\overline{e_r}|T) = -\beta X$. Moreover, $e_D - X < \overline{e_r}$ implies $e_D - X < e_{rH}$ because $\overline{e_r} < e_{rH}$, so that $v_R(e_{rH}|T) = -\beta X$. Thus, $G = -\beta X - p\alpha X - (1-p)(-\beta X) = -p(\alpha + \beta)X < 0$.

(b): In this case, we have $G = \alpha(e_D - X - \overline{e_r}) + \beta(e_D - 2X - \overline{e_r}) - p\alpha X - (1-p)v_R(e_{rH}|T) = (\alpha + \beta)(e_D - 2X - \overline{e_r}) + (1-p)\{\alpha X - v_R(e_{rH}|T)\}$. By assumption, we have $e_D - X < e_{rH}$ under $e_D - 2X \leq \overline{e_r} \leq e_D - X$.[30] Then, it holds that $G = (\alpha+\beta)(e_D - (1+p)X - \overline{e_r})$, so that the sign of $(e_D - (1+p)X - \overline{e_r})$ determines the sign of G. We can see that $(i) G > 0$, $(ii) G = 0$, $(iii) G < 0$ if and only if $(i) \overline{e_r} < e_D - (1+p)X$, $(ii) \overline{e_r} = e_D - (1+p)X$, $(iii) \overline{e_r} > e_D - (1+p)X$.

(c): In this case, $G = \alpha X - p\alpha X - (1-p)v_R(e_{rH}|T) = (1-p)\{\alpha X - v_R(e_{rH}|T)\}$. Then we can see that $G = -(1-p)(\alpha+\beta)(e_D - 2X - e_{rH}) > 0$ when $e_{rH} \leq e_D - X$, and $G = (1-p)(\alpha+\beta)X > 0$ when $e_D - X < e_{rH}$.

From (a), (b), and (c), we can see that $(i) G > 0$, $(ii) G = 0$, and $(iii) G < 0$ if and only if $(i) \overline{e_r} < e_D - (1+p)X$, $(ii) \overline{e_r} = e_D - (1+p)X$, $(iii) \overline{e_r} > e_D - (1+p)X$. By Lemma 2, the proof is completed. □

Figure C.2 summarizes the characterization of G by $\overline{e_r}$.

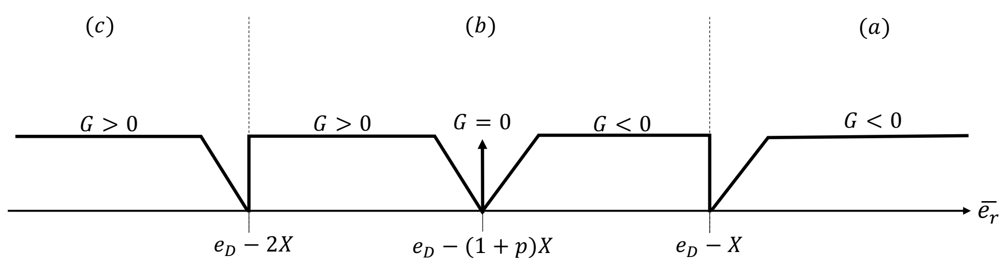


Figure C.2: Characterization of G

Table C.5 shows the determinations of G and risk attitudes under our experimental parameters.

[30] If $e_{rH} \leq e_D - X$, then by $e_D - 2X < e_{rH}$, $G = (\alpha+\beta)(e_D - 2X - \overline{e_r}) - (1-p)(\alpha+\beta)(e_D - 2X - e_{rH}) = p(\alpha+\beta)(e_D - 2X)$ (note that $\overline{e_r} = (1-p)e_{rH}$). Therefore, in this case G does not depend on $\overline{e_r}$, which is a peculiar result driven by the linear form of Fehr and Schmidt (1999).

Pairs	e_D	X	e_{rH}	p	$\overline{e_r}$	$e_D-(1+p)X$	G	risk attitude
Pair 1	10	5	6	0.9	0.6	0.5	$G<0$	*risk loving*
Pair 2	10	5	6	0.8	1.2	1	$G<0$	*risk loving*
Pair 3	10	5	12	0.9	1.2	0.5	$G<0$	*risk loving*
Pair 4	10	3	6	0.9	0.6	4.3	$G>0$	*risk averse*
Pair 5	10	3	6	0.8	1.2	4.6	$G>0$	*risk averse*
Pair 6	10	3	12	0.9	1.2	4.3	$G>0$	*risk averse*
Pair 7	10	5	6	0.95	0.3	0.25	$G<0$	*risk loving*
Pair 8	10	5	12	0.95	0.6	0.25	$G<0$	*risk loving*

Table C.5: Determinations of G and risk attitudes in our experiment

C.4 Proof of prediction 2-1

Here, we use the notations in the main text and in Appendix C.1. For prediction 2-1, we prove the following:

Proposition C.2 *Suppose a dictator's preference is represented by utility function (C.2) and the dictator has a preference such that* $\alpha_{rH}^{T} > \alpha_{rH}^{T'}$ *and* $\beta_{rH}^{T} < \beta_{rH}^{T'}$. *Further, suppose* $\alpha_k^T = \alpha_k^{T'}$ *and* $\beta_k^T = \beta_k^{T'}$ *for any* $k \in \{C, r, rL\}$.[31] *If* $e_C^T \in (0, e_{rH})$ *holds,*[32] *then* $e_C^{T'} > e_C^T$.

Proof: Since a dictator has utility function (C.2), we can apply equations (C.3), (C.4), (C.5), and (C.6). By assumptions for parameters, we have $v_C(e_C|T) = v_C(e_C|T')$, $v_r(\overline{e_r}|T) = v_r(\overline{e_r}|T')$, $v_{rL}(e_{rL}|T) = v_{rL}(e_{rL}|T')$, and $v_{rH}(e_{rH}|T) > v_{rH}(e_{rH}|T')$. Therefore, by equation (C.3), we can see that $V(e_C|T') > V(e_C|T)$ for all $e_C \in [0, e_{rH}]$. Since both $V(e_C|T)$ and $V(e_C|T')$ are continuous and decreasing in e_C, $e_C^T \in (0, e_{rH})$ implies $e_C^{T'} > e_C^T$. $\square$

Figure C.3 illustrates an example of $V(e_C|T)$ and $V(e_C|T')$. The black line corresponds to $V(e_c|T)$ and the red line corresponds to $V(e_c|T')$. As shown in the figure, $V(e_C|T') > V(e_C|T)$ holds for all $e_C \in [0, e_{rH}]$, so that the switching point in the treatment T' becomes larger than that of treatment T.

[31] We can have the same result even in the assumption that $\alpha_r^T > \alpha_r^{T'}$ and $\beta_r^T < \beta_r^{T'}$.

[32] We prove the existence of a switching point e_C under the anonymity assumption, that is, $\alpha_C = \alpha_r = \alpha_{rL} = \alpha_{rH}$, $\beta_C = \beta_r = \beta_{rL} = \beta_{rH}$. In this analysis, the parameters differ by recipient, so that existence is not always guaranteed.

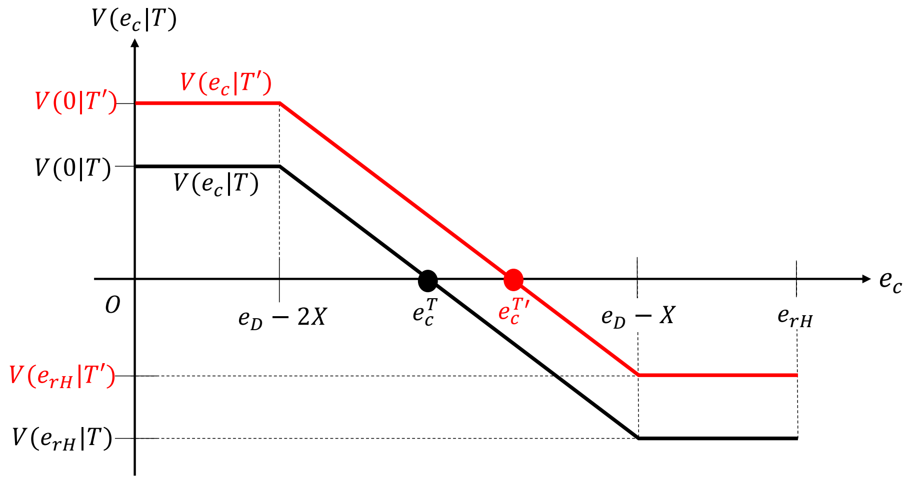


Figure C.3: Effect of preference for other's state in the case of $e_D - X < e_{rH}$

C.5 Proof of prediction 2-2

Here, we again use the notations in the main article and Appendix C.1. For Prediction 2-2, we prove the following:

Proposition C.3 *Suppose a dictator's preference is represented by utility function (C.2) and the dictator has a cognitive bias such that $\pi(1-p|T') > \pi(1-p|T)$. Further, suppose $e_D - 2X \geq 0$ and $e_D - 2X + \frac{\alpha_{rH}}{\alpha_{rH}+\beta_{rH}}X < e_{rH}$. If $e_C^T \in (0, e_{rH})$ holds, then $e_C^{T'} > e_C^T$.*

Proof: Since a dictator has utility function (C.2), we can apply equations (C.3), (C.4), (C.5), and (C.6). The subjective expected value of recipient R's allocations without transfers is denoted by $\tilde{e_r}^T = \pi(1-p|T)e_{rH}$ for each T. Note that $\tilde{e_r}^{T'} > \tilde{e_r}^T$ because $\pi(1-p|T') > \pi(1-p|T)$. First, it holds that $v_C(e_C|T) = v_C(e_C|T')$ because it does not depend on subjective probabilities. Next, because $\tilde{e_r}^{T'} > \tilde{e_r}^T$, $v_r(\cdot|T) = v_r(\cdot|T')$, and v_r is non-increasing, it holds that $v_r(\tilde{e_r}^T|T) \geq v_r(\tilde{e_r}^{T'}|T')$. Finally, $0 \leq e_D - 2X$ implies $v_{rL}(0|T) = v_{rL}(0|T') = \alpha_{rL}X$. Moreover, $e_D - 2X + \frac{\alpha_{rH}}{\alpha_{rH}+\beta_{rH}}X < e_{rH}$ implies $v_{rH}(e_{rH}|T) = v_{rH}(e_{rH}|T') < 0$,[33] so that $\pi(1-p|T') > \pi(1-p|T)$ implies $\pi(1-p|T)v_{rH}(e_{rH}|T) > \pi(1-p|T')v_{rH}(e_{rH}|T')$. Therefore, from equation (C.3), we can see that $V(e_C|T') > V(e_C|T)$ for all $e_C \in [0, e_{rH}]$, and $e_C^T \in (0, e_{rH})$ implies $e_C^{T'} > e_C^T$. □

[33] $e_D - 2X + \frac{\alpha_{rH}}{\alpha_{rH}+\beta_{rH}}X < e_{rH}$ implies $\alpha_{rH}(e_D - X - e_{rH}) + \beta_{rH}(e_D - 2X - e_{rH}) < 0$ and $e_D - 2X < e_{rH}$. Then, from equation (C.6), we can see that $v_{rH}(e_{rH}|T) < 0$.