ifo WORKING PAPERS

399
2023

June 2023

Prosocial Risk-Taking: Growing the Pie or Increasing your Slice?

Nina Weber

ifo INSTITUTE

Leibniz Institute for Economic Research
at the University of Munich

Imprint:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.ifo.de

An electronic version of the paper may be downloaded from the ifo website:
www.ifo.de

Prosocial Risk-Taking: Growing the Pie or Increasing your Slice?*

Abstract

Many personally risky decisions, such as innovation and entrepreneurship, have the potential to increase overall welfare by creating positive externalities for society. Rewarding such prosocial risk-taking may be an important strategy in addressing societal challenges like, for example, the climate emergency, by promoting innovation that has positive externalities for the environment. A fundamental constraint for policy makers in rewarding such behaviour are however individuals' distributive preferences. In this paper, I provide a theoretical framework and a first experimental test of how distributive preferences are affected by potential positive externalities of risky behaviour.

JEL Code: D63, D62, H23, D81, C91
Keywords: Prosocial risk-taking, externalities, distributive preferences, fairness

Nina Weber
Ludwig Erhard ifo Center for Social Market Economy
and Institutional Economics
Gartenstr. 6
90762 Fürth, Germany
nina.weber@ifo.de

* I thank my supervisor Shaun Hargreaves Heap, Alexander Cappelen, Aurelien Baillon, Bertil Tungodden, Chris Roth, Daniel Carvajal, Erik Sorensen, Frauke Stehr, Frederik Schwerter, Heidi Thysen, Kai Barron, Max Lobeck, Pia Pinger, Sam Hirshman, Thomas Meissner, and the seminar and conference participants at UC San Diego, NHH Bergen, the BBE Colloquium, the University of Cologne, and the European ESA Meeting in Bologna for helpful feedback and discussions. The experiment was preregistered via the American Economic Association registry for Randomized Controlled Trials with RCT ID AEARCTR-0009701 and was granted ethical clearance from the Research Ethics Committee at King's College London (reference number MRSP-21/22-30100). The data collection benefited from funding by the ESRC.

I Introduction

Many decisions people make have the potential to increase overall welfare by creating positive externalities. This is especially the case for decisions that are personally risky: Starting a business, developing new technologies, or investing into new ventures, are all decisions that involve uncertainty about personal gains and losses, but have potential positive externalities for wider society. For example, an entrepreneur may create knowledge spillovers through her products and services, while taking on personal risk over her own earnings. In this paper, I will refer to these decisions as prosocial risk-taking (Do et al. 2017). Prosocial risk-taking may be especially important in addressing societal challenges, such as economic inequalities or the climate emergency (Embry et al. 2019), through social entrepreneurship or by creating climate-friendly technologies. The likelihood of a person taking such prosocial risk is however affected by whether and how societies reward this decision. A growing literature on the optimal taxation of innovation is therefore concerned with how to ideally incentivise and reward innovators (e.g. Djankov et al. 2010; Da Rin et al. 2011; Stantcheva 2021a; Akcigit and Stantcheva 2020; Akcigit et al. 2022). While this literature is concerned with the optimal supply of such policies, there is so far no research on the demand for policies aimed at rewarding prosocial risk-takers. Specifically, how potential positive externalities affect distributive preferences is an open question. In this paper, I provide a simple theoretical model and a first experimental test of how distributive preferences are affected by the potential positive externalities of prosocial risk-taking.

The potential positive externalities of prosocial risk-taking can affect distributive preferences in a couple of ways which have so far not been taken into account in the literature and the experiment is designed with this in mind: The first and obvious one is that distributive preferences over income earned from risky choices could be sensitive to those potential positive externalities. While distributive preferences over income earned from risky decisions have been examined by Cappelen et al. (2013a) and others, it is plausible that these preferences would be affected if risk-taking also creates a potential benefit for others. This possibility has, however, not been examined yet. The second and less obvious consequence of the presence of positive externalities is that they create windfall gains for the individuals who receive them. Those individuals enjoy benefits which have nothing to do with their own choice of actions. Much less is known about distributive preferences with respect to windfall gains than with respect to outcomes for individuals that can be directly traced to decisions made by those individuals. For this reason, I focus on the allocation of windfall gains in my experimental design.

A large existing literature in economics is concerned with understanding the determinants of individuals' distributive preferences (Fehr and Schmidt 1999; Corneo and Grüner 2002; Falk et al. 2003; Alesina and Angeletos 2005; Klor and Shayo 2010; Cappelen et al. 2010; Alesina and Giuliano 2011; Luttmer and Singhal 2011; Cappelen et al. 2013b; Durante et al. 2014; Kuziemko et al. 2015; Alesina et al. 2018a,b; Stantcheva 2021b). The main finding, especially of the experimental strand of this literature, is that income earned through effort is redistributed less than income earned through luck (Cherry et al. 2002; Cherry and Shogren 2008; Krawczyk 2010; Cappelen et al. 2013a; Lefgren et al. 2016; Gee et al. 2017; Rey-Biel et al. 2018; Almås et al. 2020). Here, income is usually distributed by a third-party spectator with no stake in the outcomes, to isolate people's prosocial preferences. Some papers have also looked at how income from risky choices is redistributed, finding that most redistribute little even if outcomes were unlucky (Cappelen et al. 2013a; Mollerstrom et al. 2015; de Oliveira et al. 2017). This suggests that people tend to be held accountable for the choices they make when it comes to distributive preferences. While these studies have developed a good understanding of the effects of different earnings mechanisms on distributive preferences, how these preferences are affected when choices can also create externalities remains an open question. Most closely related, a small literature has looked at how a surplus is divided if subjects contributed unequal amounts to this surplus (Ruffle 1998; Rodriguez-Lara and Moreno-Garrido 2012). Here, subjects who contribute more to the surplus also tend to receive a larger payoff.

Almost all of the existing experimental studies on distributive preferences focus on how income earned through these different mechanisms is redistributed, mostly by third-party spectators, but how genuine windfalls are distributed is so far not well understood. Somewhat related, Chowdhury and Jeon (2014) study the effects of increased common show-up payments on dictator game behaviour and find that dictators distribute more if the show-up payment is larger. Importantly however, an increase in the show-up payment is not directly comparable to a windfall gain as it is still earned through participation in the experiment. Heap et al. (2021) use actual windfalls to study group identification in dictator games. How windfall gains are distributed by third-party spectators is however also an open question.

To provide a first experimental test of how potential externalities affect distributive preferences, I run an experiment consisting of decision makers and spectators. Decision makers have the option to choose a lottery or a safe payoff. In a control group, both choices only affect the own income of the decision maker. In the treatment group, decision makers face an identical choice between a lottery and a safe payoff; however, the lottery might also produce positive externalities for other, anonymous participants. Whether these potential positive externalities realise is unknown ex ante and the likelihood of them realising is independent of the outcome of the lottery.

After decision makers have made their choices, spectators have to allocate a windfall bonus between randomly selected pairs of decision makers. Pairs are designed to always be composed of one decision maker who chose the safe option and one who chose the lottery. In a within-subject design, spectators make 20 such allocation decisions across three different conditions which are randomised in order. One of their allocation decisions is ultimately selected at random to determine the payment of a decision-maker pair. In the control condition, pairs are selected from the control group of decision makers and so no potential externalities are present. Spectators therefore make their allocation decisions knowing only whether the decision makers in each pair selected the safe or risky option and their respective income. In the ex ante treatment, decision makers are selected from the treatment group. As in the control condition, spectators know which choices decision makers made and their respective income, but are not informed about whether the potential externalities actually generated. In the ex post treatment, spectators are fully informed about decision makers choices, income, and whether the externality realised.

The results of the experiment mostly support the expectations of my theoretical model and suggest that spectators reward decision makers for prosocial risk-taking. On average, the share of the windfall bonus allocated to the risk-taker increases from 49% to 53% if risk-taking can result in potential externalities. This result is independent of the risk-taker's own income and, on average, identical for the ex post and ex ante treatment. As predicted by the theoretical model, this effect also increases as the opportunity cost of choosing the risky option increases. Specifically, as the value of the safe option increases by $1, the reward allocated to the risk-taker increases by about 28-30ct in both treatments.

reward allocated to the risk-taker increases by about 28-30ct in both treatments. While these main treatment effects suggest that individuals have a significant preference for rewarding prosocial risk-taking, the results also indicate outcome bias: Spectators primarily reward risk-takers if the externality actually realised ex post. In fact, the treatment effect disappears when focusing only on the scenarios in the ex post treatment condition where the externality did not realise. That is, despite risk-takers having no agency over the likelihood of the externality realising. Additionally, and contrary to the theoretical prediction, spectators do not compensate risk-takers more in either treatment if the personal lottery was unsuccessful. Taken together, these two additional findings indicate that while spectators reward prosocial risk-taking if outcomes are unknown, they only reward successful risk-takers ex post and do not compensate prosocial risk-taking if unsuccessful. These outcome-dependent findings are in line with a growing literature on attribution bias in distributive preferences (Gurdal et al. 2013; Brownback and Kuhn 2019; Cappelen et al. 2016; Andre 2021; König-Kersting et al. 2021).

The implications of these main findings are primarily relevant for the stability of policies aimed at rewarding prosocial risk-takers, such as entrepreneurs and innovators. While people have a preference to reward prosocial risk-takers, they only wish to do so if the prosocial risk-taker was successful in generating positive externalities, even if this outcome is the result of pure luck. In practice, this finding suggests that policies aimed at rewarding prosocial risk-takers may receive support ex ante, but will likely lose this support ex post, if externalities did not realise. For example, monetary rewards for potential innovators may be in line with people's preferences to reward prosocial risk-taking ex ante. However, the support for such a policy is likely unstable and highly dependent on the success of those innovators ex post. This would be the case, even if the successful or unsuccessful outcome could not have been predicted ex ante. An important follow-up question for future research is, therefore, how this instability of policy support, if such policies are implemented, might affect the likelihood of subsequent prosocial risk-taking by decision makers.

As further analysis of spectator beliefs and choices indicates, spectators appear to trade-off their personal fairness criteria over earned income with a desire to reward prosocial risk-taking. In practice, this means that while a spectator might be a choice egalitarian in the control condition, this preference is traded-off against a desire to increase the reward for the risk-taking subject when risk-taking can generate externalities. Theoretically, this finding suggests that the desire to reward externality-generating choices can be added as an additional term in existing fairness models (Cappelen et al. 2007; Almås et al. 2020). The additional analysis however also reveals substantial heterogeneity in distributive decision making over prosocial risk-taking and identifies five distinct reward types among spectators.

Finally, the results also indicate that decision makers take substantially more risk when risk-taking can create potential positive externalities for others. This is the case even though reputational concerns most likely do not affect choices, as all participants are anonymous and unlikely to ever interact in person. In a belief elicitation, risk takers indicate however that they expect to receive a larger share of the reward in the treatment compared to the control condition. This indicates that a potential motive for prosocial risk-taking might be the expected subsequently larger share of the windfall bonus. This question can however not be answered conclusively within the given experimental design of this paper. Spectators also underestimate how strongly decision-makers react to the possibility to generate positive externalities for others. They believe the average first-stage treatment effect to be only half its actual size. They also underestimate the importance of opportunity costs for prosocial risk-takers. Another potentially interesting question for future research would therefore be whether spectator preferences change if these incorrect beliefs about prosocial risk-taking are corrected.

The outline of this paper is as follows. Section II develops a simple and generalisable theoretical framework that incorporates externalities into distributive decision making and outlines the main hypotheses following from this. Section III describes the experimental design and section IV reports the results and further analysis of spectator beliefs. Section V concludes.

II Theoretical Framework

To provide a generalisable framework to analyse distributive choices in the presence of externalities, I propose a theoretical model that takes the size of any potential externalities and the expected costs and benefits of exposing oneself to risk into account. My basic model builds on the decision model proposed in Cappelen et al. (2007, 2013a, 2016). This model assumes that in a situation in which an impartial spectator decides on the allocation of money between two individuals, each spectator holds a specific fairness criterion which specifies a fair allocation of money to each of the two individuals. While a stakeholder would trade off their fairness criterion with their own self-interest, a spectator is exclusively motivated by their fairness criterion as self-interest cannot affect the decision of the spectator (Cappelen et al. 2013a).

There are two possible ways in which externalities can enter this model: First, as an additional fairness criterion that affects choices only in the presence of externalities. Second, as an additional term in the utility function that spectators weigh against their fairness criterion over decision makers' own income. I choose the second option in this paper, for the following reason: In most of the scenarios in which externalities are present, a desire to reward risk takers and common fairness criteria contradict each other. For example, a risk taker's own lottery might be successful and the risk taker becomes the person with the higher income in the pair. An inequality averse spectator who also wants to reward risk taking would then have to weigh her desire to reward the risk taker against her desire to decrease inequality within the pair. Depending on where she positions herself on this trade-off, she may then either allocate the bonus to equalise incomes, allocate it to the risk taker as a reward, or decide on a combination of these two options. I therefore assume that spectators' fairness criteria are independent of their desire to reward risk takers.

For a given subject, a spectator may then choose an allocation of payoff y based on the below equation. I assume for now that there are only two subjects i and j and that both subjects face the same choice set with options $a = 1$ and $a = 0$:

$$V(y_i) = -\beta\frac{(y_i - FE_i)^2}{2X} + \delta(e_i - d_i)y_i \tag{1}$$

Here, as in Cappelen et al. (2007, 2013a, 2016), β is the weight attached to fairness over earned income, FE the share of the bonus the subject should receive based on the spectator's fairness criterion, and X the total available bonus. This first term of the equation is almost identical to the model proposed in Cappelen et al. (2013a). The second term of the equation is where externalities enter the model. It captures the weight δ attached to the unallocated externalities e_i generated by choosing option $a = 1$ over $a = 0$. These externalities are assumed to be "unallocated" as the spectator has no information on the recipient of the externalities when making their choice. Spectators are also assumed to account for expected benefits in the payoffs subjects receive when choosing option $a = 1$ over $a = 0$ by reducing the reward and, vice versa, to increase the reward as the expected cost of choosing the option that creates the externality rises. This is captured by the term d_i.

Given the well-documented presence of outcome bias in distributive decision making (e.g. Brownback and Kuhn 2019; Cappelen et al. 2016; Andre 2021; König-Kersting et al. 2021), the value of e_i and d_i is likely going to vary depending on whether outcomes have already been realised or not. Spectators may also differ in their tendency to display outcome bias. Specifically, e_i is then given as:

$$e_i = \begin{cases} E[\pi(a_i = 1) - \pi(a_i = 0)] & \text{if } \pi \text{ is unknown} \\ \rho[\pi(a_i = 1) - \pi(a_i = 0)] + (1 - \rho)E[\pi(a_i = 1) - \pi(a_i = 0)] & \text{if } \pi \text{ is known} \end{cases} \tag{2}$$

Here, π is the size of the total unallocated externalities. If π is unknown, e_i is equal to the difference in the expected size of the total unallocated externality π when subject i chooses option $a_i = 1$ as opposed to $a_i = 0$. If however outcomes are already realised and spectators know the value of π, then they may display some degree of outcome bias. This is captured by a a value of $\rho > 0$. A spectator who displays complete outcome bias, i.e. does not take ex ante options into account at all, has a value of $\rho = 1$. Equally, d_i is given as:

$$d_i = \begin{cases} E[y_i(a_i = 1) - y_i(a_i = 0)] & \text{if } y_i \text{ is unknown} \\ \rho[y_i(a_i = 1) - y_i(a_i = 0)] + (1-\rho)E[y_i(a_i = 1) - y_i(a_i = 0)] & \text{if } y_i \text{ is known} \end{cases} \quad (3)$$

Similarly, if the actual payoff of the risk taking subject y_r is yet unknown, d_i is equal to the difference in expected payoffs y_i for subject i when choosing option $a = 1$ as opposed to $a = 0$. If however spectators know the outcome of the personal lottery of the risk taking subject, then they may display again some degree of outcome bias as captured by a value of $\rho > 0$. In practice, this suggests that spectators assess the cost of the risky decision as lower (higher), if the personal lottery was successful (unsuccessful), and reduce (increase) their reward accordingly.

If spectators are motivated by a particular fairness view over the own income of the decision makers, they may follow a choice egalitarian or ex ante fairness criterion (Cappelen et al. 2013a) and decide on an even split of the bonus. Alternatively, spectators may follow an inequality averse or ex post criterion (Fehr and Schmidt 1999) and decide to equalise outcomes.

To ensure that spectators can follow either of these allocation rules while also rewarding risk-takers, the size of the windfall bonus exceeds the maximum possible level of inequality within the pair, which equals \$3.5, as well as the maximum value for $e_i - d_i$, which is \$2.5.[1] Given this, spectators have a large enough choice set so that I can observe interior solutions of the proposed decision model for all subjects. The optimal choice of y for a spectator is then given by:

$$y^* = FE + \frac{\delta}{\beta}(e_i - d_i) \quad (4)$$

In practice, this predicts that a spectator who is motivated both, by a particular fairness criterion and by potential externalities, would allocate to subject i the amount the given fairness criterion suggests, as well as an externality reward depending on the relative im-

[1] The externality has an expected value of \$1 as the \$2 only realise in 50% of cases irrespective of the outcome of the lottery. If a spectator has to redistribute between two subjects whereby one chose the lottery and one took \$3.5, the maximum safe option available, then this implies $d = -\$1.5$ and so $e_i - d_i = \$1 + \$1.5$.

portance of δ over β. I can thus estimate the reward parameter δ by comparing spectator allocations between treatment and control for given values of e_i and d_i.

II.1 Hypotheses

The theoretical model outlined in section II allows me to make several predictions about spectator decisions. First, if $e_i - d_i$ is positive, which is the case for all decisions in which the safe option is equal to \$1.50 or larger, then the model predicts that a spectator who has a positive reward parameter δ will increase y_r, the share allocated to the subject choosing the lottery, in the presence of potential positive externalities in the ex ante treatment condition. Outcome bias cannot influence spectator choices in this treatment as the realised values of $\pi(a_i = 1) - \pi(a_i = 0)$ and $y_r(a_i = 1) - y_r(a_i = 0)$ are unknown and therefore only the expected values can influence e_i and d_i. These theoretical predictions lead to H1a:

> **Hypothesis 1a:** y_r, the share of the bonus allocated to the subject choosing the lottery, is higher in the ex ante treatment than in the control condition.[2]

The model further predicts that for larger values of the safe option A, the value of $e_i - d_i$ increases. This means that spectators who have a positive reward parameter δ will reward risk takers more in the presence of potential positive externalities if the expected value of the lottery chosen decreases relative to the available safe option.[3] H1b follows:

> **Hypothesis 1b:** The treatment effect in H1a increases as the value of the safe option A increases.

As outlined in section II, an extensive literature on financial decision making as well as distributive preferences for choices that are affected by luck, points to the importance of outcome bias in spectator (or principal) decision making (e.g. Gurdal et al. 2013; Brownback and Kuhn 2019; Cappelen et al. 2016; Andre 2021; König-Kersting et al. 2021). Outcome

[2] That is, if the difference between the expected size of the externality and the expected personal benefit from choosing the risky option is positive or zero. Therefore, H1 will primarily be tested for decisions where the safe option > \$1.

[3] This effect could be nonlinear as it may be particularly strong in those cases where the EMV of the lottery is smaller than the EMV of the safe option (when the safe option is > \$2). That is, because risk takers incur an actual cost to their own expected payoff in those instances which spectators may wish to compensate. For simplicity, such non-linearity is not included in the theoretical model.

bias is captured in the theoretical model by a value of $\rho > 0$. Assuming that spectators display some degree of outcome bias in their decision making, the model predicts a larger treatment effect on y_r when the externality realised than when it did not. That is, because spectators may not take into account the ex-ante choice set available to decision makers but focus mostly on the ex-post monetary outcomes. H2a follows:

> **Hypothesis 2a:** The ex post treatment effects in H1a and H1b are larger when the externality realised than when it did not.

Given that the probabilities of the externality realising and the lottery being successful for the risk-taking subject are independent in the proposed experimental design, I am also able to test whether the outcome of the personal lottery, irrespective of whether the externality realised, affects y_r, the share of the bonus allocated to the subject choosing the lottery. The theoretical model here predicts that if the value of $\rho > 0$, the positive treatment effect on y_r predicted in H1a will be larger if the personal lottery was unsuccessful compared to when it was successful. That is, because for a positive value of ρ, the value of $e_i - d_i$ increases if the lottery was unsuccessful and decreases if it was successful. H2b follows:

> **Hypothesis 2b:** The ex post treatment effects in H1a and H1b are larger when the personal lottery was unsuccessful than when it was successful.

While not central to the main research question of this study, the first stage of the experimental design also allows me to test whether the choices made by decision makers are affected by the potential positive externalities of the risky option. This is, in fact, a topic which has received very little attention in the experimental literature so far. To the best of my knowledge, only one paper has tested the effects of externalities on risky choices (de Oliveira 2021). She finds that while subjects are less willing to make risky choices in the presence of negative externalities, they are not more willing to expose themselves to risk to create positive externalities for others. Given that the externalities in her experimental design are certain, this would suggest that subjects who are faced with only potential positive externalities, as is the case in my design, may be even less likely to respond to these externalities with more risk seeking behaviour. Importantly, the decision makers in this study are however aware of the third stage of the experiment, and know that a spectator will judge their choices before awarding a windfall bonus. Even though they might therefore not be motivated to choose the risky option due to altruism (de Oliveira 2021), they may nonetheless expect spectators to judge the risk taking behaviour favourably when positive externalities could result from their choice. H3 follows:

> **Hypothesis 3:** Decision makers are more risk seeking in the treatment than in the control group.

H3 also provides a further justification for providing spectators with hypothetical as well as real decision scenarios, given that the distribution of choices is likely to be different in the treatment as opposed to the control group.

III Experimental Design

My basic experimental design follows in particular Cappelen et al. (2013a) by asking an impartial spectator to decide on a fair allocation of a monetary bonus between two decision makers. The impartial spectator frame hereby allows me to isolate distributive preferences by eliminating the possibility of any selfish considerations affecting the decision making. Prior to spectators making their distributive choice, decision makers have to choose between a lottery and a safe income. In the treatment condition, choosing the lottery does not just yield a potentially high reward for the decision maker if successful, but might also result in positive externalities for the participant pool. The probability of the externality realising and the personal lottery being successful are however independent of each other for two reasons: First, this aspect of the design allows for a clean estimation of the parameters of my decision model outlined in II. Second, risky choices may also have positive externalities even if the outcome of the choice itself is unsuccessful. In the case of entrepreneurship, an example of this are knowledge spillovers (Acs et al. 2009).

Importantly, this positive externality will not benefit the two decision makers in the pair but will be used to reward randomly chosen participants from other, unrelated studies. This is a crucial aspect of the design as it allows me to hold both, expected payoffs and the degree of inequality within pairs, constant. It also ensures that there are no strategic reasons for decision makers to expose themselves to risk to generate the externality beyond a potential desire to be received as altruistic by the spectators. In the experimental instructions, I also do not refer to "externalities" at any point to avoid any positive or negative connotation subjects may have with the word.

Finally, the expected monetary value (EMV) of the risky option is, in some decisions, lower than the EMV of the safe option. While these risky options are strictly dominated for a risk averse or risk neutral subject when no potential externalities are present, this aspect of the design allows me to test whether decision makers become more risk seeking in the presence of externalities; and, how spectators value such choices.

I depart from most existing studies by asking spectators not to redistribute the earned income of the pair but to allocate an additional monetary bonus between the two decision makers. This addresses a concern arising from the literature on distributive preferences over income earned from different choices: Given that a substantial proportion of spectators can be classified as acting according to a choice egalitarian or ex ante fairness criterion (e.g. Cappelen et al. 2013a), one would likely observe little baseline redistribution in a design that asks spectators to redistribute existing earnings as opposed to distributing an additional bonus. By asking spectators to allocate an additional monetary amount, I am therefore able to observe the distributive preferences of all spectators. Additionally, while most of the experimental literature on distributive preferences focuses on subjects' choices over endowments, preferences over windfall gains have been studied less. As the existence of positive externalities means that many people's income will, at least partly, consist of what from the perceptive of that individual are windfalls, I focus on spectators' distributive choices over a windfall gain.

Stage 1: Participants recruited via Prolific Academic participate in either the control or treatment condition that determines individual payoffs.

Control: Subjects are asked to decide between the following two options seven times, one choice being randomly selected for payment, each time with a slightly different value for option 1:

Option 1: \$0.5/\$1/\$1.5/\$2/\$2.5/\$3/\$3.5

Option 2: A 50% chance to receive \$4 and a 50% chance to receive \$0.

Treatment: Subjects are asked to decide between the following two options seven times, one choice being randomly selected for payment, each time with a slightly different value for option 1:

Option 1: \$0.5/\$1/\$1.5/\$2/\$2.5/\$3/\$3.5

Option 2: A 50% chance to receive \$4 and a 50% chance to receive \$0. Irrespective of the outcome of the lottery, there is a 50% chance that an externality of \$2 will generate which will be used to reward two randomly chosen participants from other studies.

Stage 2: Participants are paired based on the procedure outlined in IV.2 and one of the seven decisions subjects made is randomly chosen to determine payoffs. Each pair consists of a subject that chose the lottery and a subject that chose the safe option for the randomly selected decision. This is however unknown to subjects prior to making their choice, to ensure that subjects cannot make strategic decisions about the likely composition of the pair. Importantly, only subjects within the same treatment condition are matched.

Stage 3: Impartial Spectators, who have not participated in the first two stages of the experiment, are asked to allocate a windfall bonus of \$4 between pairs of decision makers in one control and two treatment conditions. Each spectator makes allocation decisions in all three conditions, but the order in which spectators see these conditions is randomized:

Control condition: Spectators allocate the \$4 between pairs of decision makers from the control group in stage 1. They receive full information of the choices made and the resulting earnings of each decision maker.

Ex ante treatment condition: Spectators allocate the \$4 between pairs of decision makers from the treatment group in stage 1. They receive full information of the choices made and on the resulting earnings of the decision makers but no information on whether the externality realised or not.

Ex post treatment condition: Spectators allocate the \$4 between pairs of decision makers from the treatment group in stage 1. They receive full information of the choices made, the resulting earnings of the decision makers, and on whether the externality realised or not.

Figure 1 shows example scenarios spectators might face during the experiment. Importantly, spectators only compare the decisions of two subjects for the same choice set, i.e., the choices made when faced with the same safe and risky option. They also do not have the option to communicate with the participants.

Overall, spectators make 20 such allocation decisions - 5 in the control condition, 5 in the ex ante treatment condition, and 10 in the ex post treatment condition - with 19 being hypothetical and one resulting in actual payoffs for a participant pair. Spectators however do not know which of the allocation decisions will result in actual payoffs when making their decisions.[4]

IV Results

The main experiment was conducted via Prolific Academic in July 2022 with a total sample size of 180 spectators and 360 decision makers. The experimental design and all tests conducted in the following analysis were pre-registered via the American Economic Association registry for Randomized Controlled Trials with RCT ID AEARCTR-0009701. The average time subjects took to complete the experiment was 15.64 minutes. The average earnings were $5.52 which corresponds to an average hourly rate of $21.18.

IV.1 Prosocial Risk-Taking

Before analyzing spectator choices, I first turn to the results of the decision maker stage. Figure 2 reports the proportion of decision makers choosing the risky option for a given value of the safe option. The blue bars indicate the amount of risk taking in the control conditions and the green bars the corresponding amount of risk taking in the treatment condition. In both conditions, decision makers become decreasingly likely to choose the risky option as the value of the safe option increases. This preliminary result is encouraging as it indicates decision makers understand the choices they are asked to make and rationally decrease their

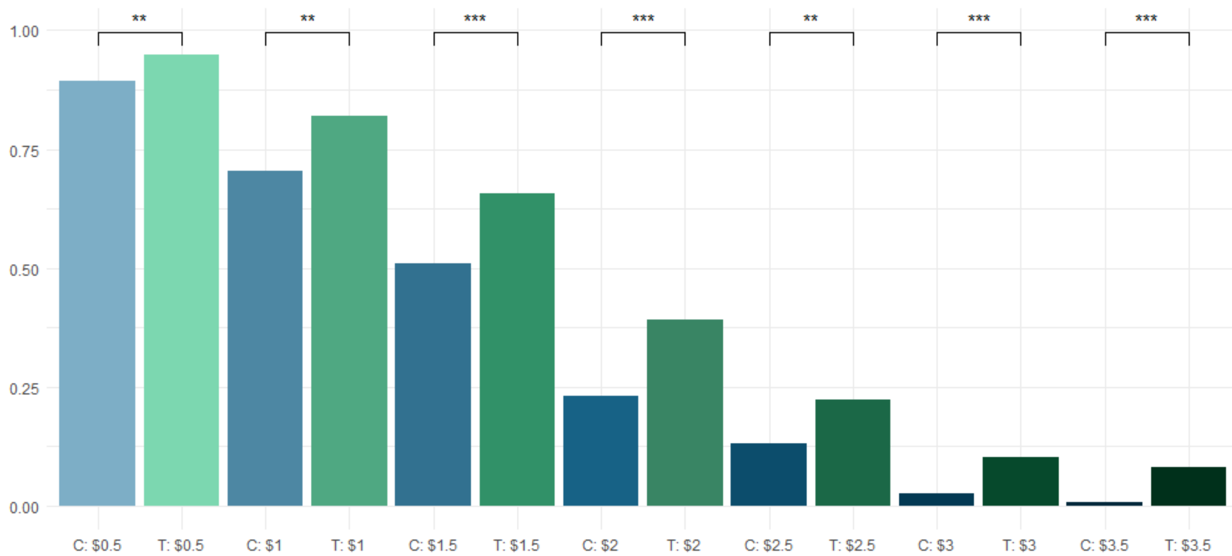
[4]This method is commonly used in experimental designs to increase the number of observed choices without affecting the behaviour of the subjects (Charness et al. 2016).

Figure 1: Example Spectator Scenarios in Control & Treatments

Notes: These figures illustrate example screens spectators face during the experiment. The top left panel shows a control condition scenario, the top right panel shows an ex ante treatment condition scenario, and the two bottom panels show ex post treatment scenarios with different outcomes for the lottery over the externality. For comparability, the personal lottery decision makers faced and its outcome is identical and positive in all four cases.

risk taking as the relative benefit of doing so also decreases. A share of 11.61% of decision makers in the control condition always choose the safe option, while none choose the risky option throughout. On average, decision makers in the control group are risk averse.

Figure 2: Proportion of decision makers choosing the risky option (EMV=$2)



Notes: The figure reports the proportion of decision makers choosing the risky option with an expected value of $2, given the amount of the safe option and treatment condition. For example, "C: $1" reports the proportion of decision makers choosing the risky option with EMV=$2 if the safe option is equal to $1 and there is no possibility to generate externalities. Significance levels indicate the results of simple t-test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

As figure 2 illustrates, the proportion of decision makers choosing the risky option in the treatment condition increases significantly for all values of the safe option compared to the control condition. This difference is also statistically significant in the aggregate ($p<0.001$) in a Pearson's chi-squared test. The results of the first stage of the experiment therefore provide strong support for hypothesis 3:

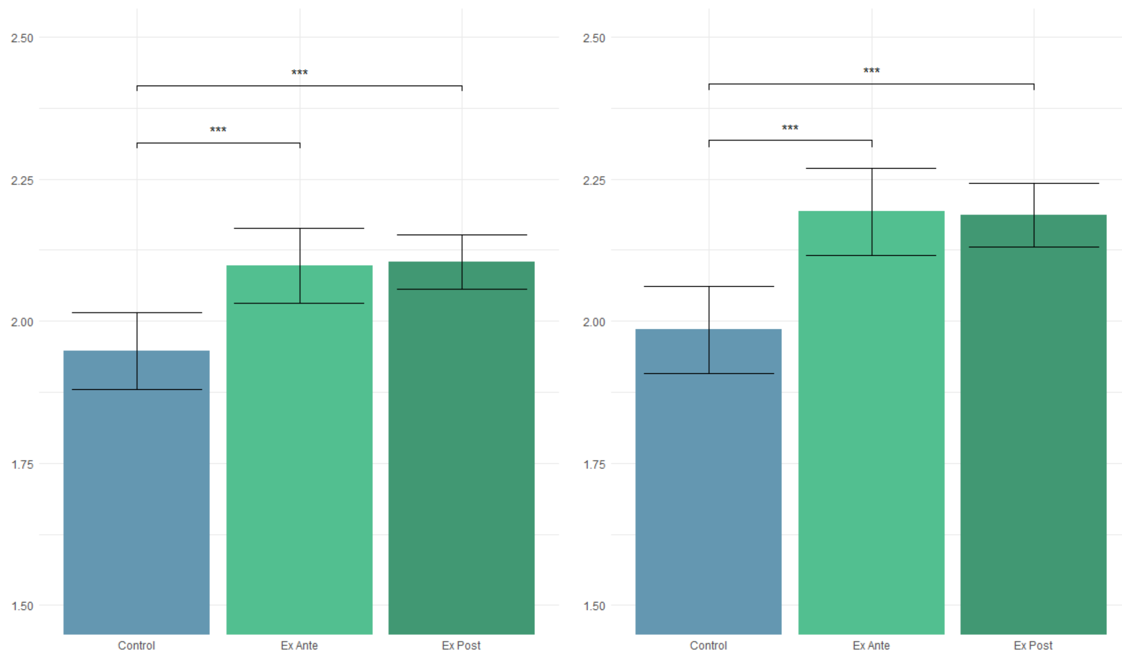
> **Result 1:** Decision makers are significantly more likely to take risk in the treatment than in the control condition.

Table A1 in appendix section A reports a balance test by random treatment assignment of decision makers. As not all variables are perfectly balanced between the two samples, table A2 also reports treatment effects controlling for those unbalanced variables. Result 1 is robust to all of the conducted tests. In fact, the unbalanced variables actually make it less likely that a treatment effect would be observed. Therefore, the reported result is likely an underestimate of the true first stage treatment effect.

IV.2 Spectator Allocation Decisions

Figure 3 summarises the mean share of the bonus allocated to the risk-taking subject by treatment condition. While spectators allocate, on average, $1.95 to the risk-taker in the control condition, this value increases to $2.10 in the ex ante and ex post treatment conditions. The differences between the control and respective treatment conditions are highly statistically significant. The right part of figure 3 restricts the same analysis to those scenarios

Figure 3: Share allocated to risk taker by treatment



Notes: The figure reports the share of the total bonus of $4 allocated to the subject who chose the risky option by treatment condition. The figure on the left includes all scenarios while the figure on the right includes only those scenarios where the safe option was larger than $1. Significance levels indicate the results of non-parametric Wilcoxon rank-sum tests. *** p<0.01, ** p<0.05 , * p<0.1.

where the safe option was larger than $1. As outlined in section II.1, the theoretical framework predicts a positive treatment effect primarily if the difference between the expected size of the externality, which equals $2, and the expected personal benefit from choosing the risky option, is positive or zero. That is, because choosing the lottery in scenarios where the safe option is equal to $1 or less is a strictly dominated choice for a risk neutral decision maker even if the potential externalities are taken into account. Spectators may therefore not consider such decisions as prosocial risk-taking.

In line with this prediction, the treatment effect on the mean share of the bonus allocated to the risk-taker is also larger in the right panel of figure 3. Here, spectators allocate, on average, $1.98 to the risk-taker in the control condition, $2.19 to those in the ex ante treatment condition, and equally $2.19 to those in the ex post treatment condition. On average, risk takers therefore receive more than half of the bonus in the treatment condition in scenarios where the safe option is larger than $1. These treatment effects are again highly statistically significant and provide support for Hypothesis 1a:

Result 2: The share of the bonus allocated to the subject choosing the lottery is significantly higher in both treatment conditions than in the control condition.

While hypothesis 1a only focused on the comparison between the control and ex ante treatment conditions, the main result also holds for the ex post treatment condition. In fact, the difference between the ex post and ex ante treatment conditions is insignificant in both panels of figure 3. This main finding is also robust to various additional tests reported in section B of the appendix. Specifically, the treatment effect is stronger if spectators are more confident in their allocation decision and if they performed better when answering the understanding questions in the beginning of the experiment. They are also robust to including a large number of control variables as reported in table B4 of the Appendix. Although it is not entirely obvious that the order in which spectators make their allocation decisions should matter, table A3 in appendix section A reports a balance test by the first treatment condition block spectators were randomly assigned to see. As not all variables are perfectly balanced between those three groups of spectators, table A4 also reports treatment effects controlling for those unbalanced variables. The results are robust to all of these additional specifications.

As outlined in the theoretical framework, hypothesis 1b predicts that the treatment effect increases as the value of the safe option increases. That is, because the opportunity cost of choosing the lottery also increases for the risk-taking subject. Table 1 reports treatment interaction effects with the level of the safe option. Given that the theoretical framework predicts that the treatment effect should primarily be observed as the value of the safe option increases above $1, the results are reported for the full set of scenarios and for those with the safe option above $1 specifically. All the coefficients in table 1 are highly significant and

Table 1: Treatment interactions with level of safe option

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment x Level	0.157***	0.154***	0.137***	0.133***
	(0.022)	(0.019)	(0.031)	(0.029)
Constant	1.788***	1.788***	1.848***	1.848***
	(0.039)	(0.039)	(0.072)	(0.072)
Observations	3,700	3,700	2,641	2,641

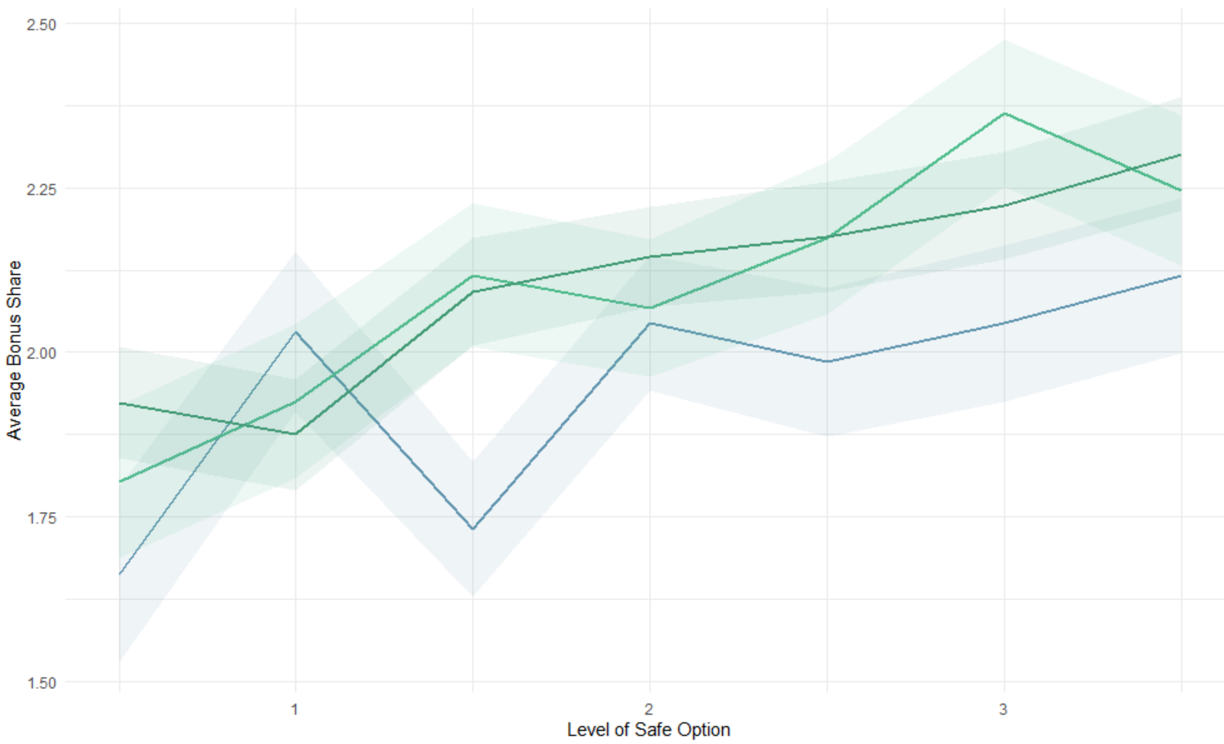
Notes: Estimates come from linear regressions. A one unit increase in the level of the safe option equals an increase of $0.50. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

positive, suggesting that the treatment effect increases as the value of the safe option increases. The coefficients are somewhat smaller when focusing only on those scenarios where the value of the safe option is above $1, but this difference is not statistically significant. Using the smaller estimates, these results suggest that as the value of the safe option increases by $0.50, the share of the bonus allocated to the risk taker increases, on average, by $0.14. In other words, 28% of the increase in the safe option is on average given as a bonus to the risk taker.

Result 3: The increase in the share of the bonus given to the risk taker in the treatment conditions increases as the value of the safe option increases.

Figure 4 provides a further, and more detailed, visual test of this hypothesis. The blue line indicates the average share of the bonus given to the risk-taking subject by different values of the safe option A. The light green line does the same for the ex ante treatment and the dark green line for the ex post treatment. Shaded areas indicate 95% confidence intervals. While the two treatment condition lines overlap quite closely for all values of the safe option A, the control condition values are lower for almost all values of the safe option. Interestingly, there

Figure 4: Increase in share allocated to risk taker by treatment



Notes: The figure reports the share of the total bonus of $4 allocated to the subject who chose the risky option by treatment condition for each value of the safe option A. As in all figures, the blue line corresponds to the control condition, the light green line corresponds to the ex ante treatment condition and the dark green line corresponds to the ex post treatment condition.

does not appear to be a linear increase in the difference between the control and treatment values. The values appear mostly indistinguishable until the safe option equals $1.5. At this point, the share allocated to the risk taker in the control condition drops significantly compared to the treatment conditions. The difference then becomes insignificant again at $2 but remains weakly significant and mostly stable until the safe option equals $3.5. Most of the interaction effect reported in table 1 therefore appears to be driven by the increase in the treatment effect after the safe option increases to $1.5.

IV.2.1 Outcome Bias over Lottery and Externalities

As outlined in section II, my theoretical framework predicts that outcome bias will show up twofold in spectators' allocation decisions: First, as outcome bias over externalities, and second, as outcome bias over personal lottery outcomes. Specifically, hypothesis 2a predicts that the treatment effect found previously will be larger if the positive externalities realised compared to if they did not. Hypothesis 2b predicts that the effect will also be larger if the personal lottery of the risk-taking subject was unsuccessful as opposed to successful. In other words, while spectators may reward prosocial risk-taking in itself, this reward is expected to be larger if the positive externalities actually realised and if the risk-taking subject was unlucky in their personal lottery. Table 2 provides a first test of these predictions focusing only on the ex post treatment, where outcomes are known. The coefficients are interaction effects of an observation being in the ex post treatment condition as opposed to the control and a particular outcome. Focusing first on whether the externality realised or

Table 2: Ex Post treatment effect interactions with outcomes

	Amount allocated to Risk-Taker			
	Externality		Lottery	
	✓	x	✓	x
Treatment x Outcome	0.225***	0.049*	0.165***	0.156***
	(0.049)	(0.049)	(0.056)	(0.059)
Constant	1.947***	1.947***	1.656***	2.233***
	(0.034)	(0.034)	(0.045)	(0.049)
Observations	2,775	2,775	2,775	2,775

Notes: Estimates come from linear regressions. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

not, the coefficients show significant outcome bias in the share allocated to the risk taker. If the externality realised, risk takers are rewarded with, on average, $0.23 more in the ex post treatment condition compared to the control condition. This is a 12% increase on the reward risk-takers receive in the control condition. If the externality however did not realise, there is no conventionally significant increase in the reward allocated to the risk-taker in the ex post treatment condition compared to the control condition. This means that a decision maker who took a personal risk that could have created positive externalities but by chance did not, is rewarded the same as a decision maker who took a personal risk with no possible benefit for others. This finding provides strong support for H2a:

Result 4: The treatment effect is significantly larger and only statistically significant if the externality realised ex post.

Moving to the interaction effects between the ex post treatment and the outcome of the personal lottery, there is no evidence of outcome bias. While there is a positive treatment effect for both potential outcomes of the personal lottery, this effect is not significantly larger or different in the scenarios where the personal lottery did not realise. In fact, if anything, the treatment effect is somewhat larger if the personal lottery did realise; however, the two coefficients are not significantly different from each other. This result suggest that, contrary to the predictions of the theoretical framework, a decision maker who took a personal risk that could have created positive externalities for others, does not get compensated if the personal risk did not pay off, relative to a decision maker who took a personal risk with no possible benefit for others.[5] Hypothesis 2b can therefore not be supported:

Result 5: The treatment effect is not significantly larger if the personal lottery was unsuccessful ex post; if anything, it is somewhat smaller.

In summary, the main results suggest that the possibility to generate positive externalities significantly increases prosocial risk taking and that, on average, spectators reward prosocial risk-takers by increasing the reward they allocate to them by 8%. This reward also increases, as the personal cost to prosocial risk-taking increases. However, if outcomes are already known, spectators only reward those risk-takers who, by chance, actually generated positive benefits for others. Prosocial risk-takers who did not generate externalities by chance but faced the same choice set ex ante, get rewarded the same as risk-takers in the control condition where externalities could not be generated at all. Spectators also, on average, do not compensate prosocial risk-takers more than risk-takers in the control group if the personal risk did not pay off. In the following, I will test how to best organise the data at the individual-level and whether beliefs of spectators can explain the distributive choices made.

[5] It should be noted that the comparison of the $treatment \times outcome$ interactions is not identical here to that for the outcomes over the externality. That is, as there was no possibility for externalities to generate in the control condition and so the comparison in the first two columns is each to the control condition where no externalities are present. On the other hand, column three and four each report the $treatment \times outcome$ interaction effect compared to the same outcome in the control condition. This technical difference is however not relevant for the interpretation of the results.

IV.3 Reward Types and Fairness Criteria

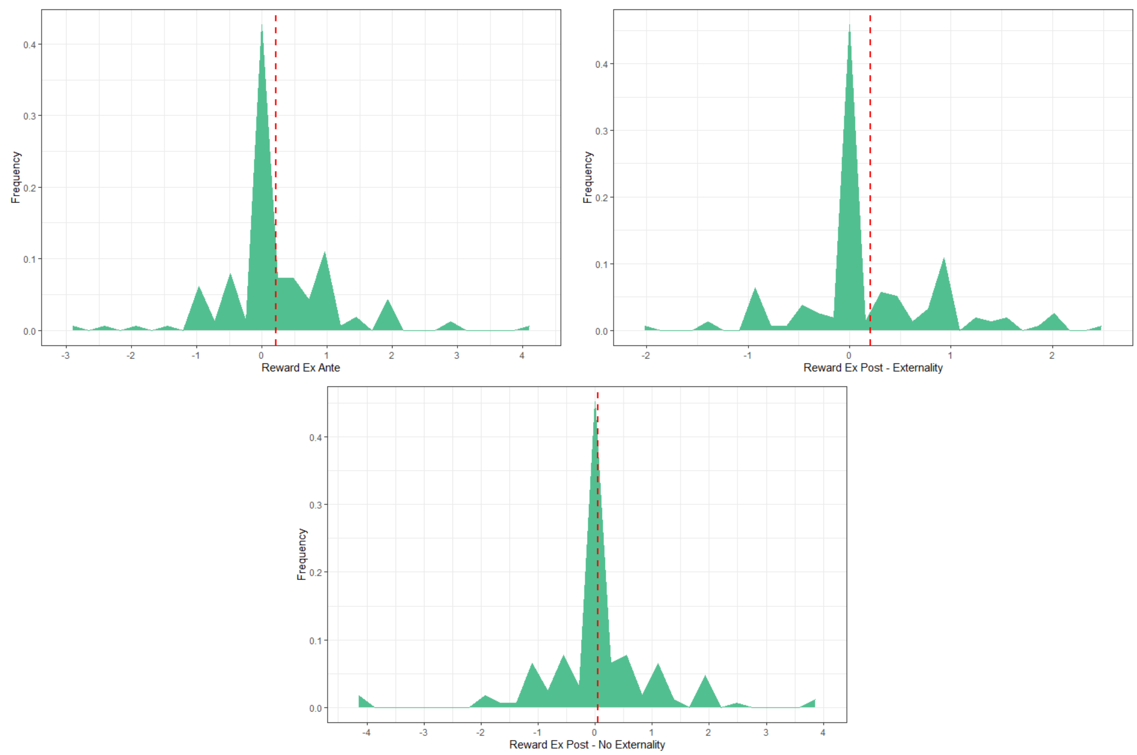
The within-subject design of the experiment allows me to estimate spectator-level treatment effects and reward types to gain a better understanding of how potential externalities affect distributive decision making. Figure 5 reports distributions of these individual treatment effects for the ex ante as well as the ex post treatment, divided into scenarios where the externality realised and those where it did not. Each value is calculated by comparing a spectator's choices between scenarios in the control and treatment conditions which are identical in the level of the safe option and the outcome of the lottery for the risk-taking subject. The difference in the share given to the risk-taking subject between the respective treatment and control scenarios is then averaged separately for each spectator and the resulting individual-level treatment effects are reported in the figure. The red dotted line reports the average individual-level treatment effect.

The distributions of individual-level treatment effects show a similar pattern for each of the three treatment scenarios. A bit less than half of all spectators have an individual-level treatment effect of zero in each treatment scenario. The remaining majority of spectators show striking heterogeneity in how they reward prosocial risk-taking, with individual-level treatment effects ranging from -\$4 to +\$4. As reported previously in the main results, there is a positive average treatment effect in the ex ante treatment and in the ex post treatment if the externality realised. There is, on average, no treatment effect in the ex post treatment if the externality did not realise. The red dotted line indicating the average is here almost exactly at \$0.

IV.3.1 Reward Types

The individual-level treatment effects for each spectator further allow me to organise spectators into reward types. Given that this paper is a first experimental test of how externalities affect distributive choices, this is also the first attempt at specifying such reward types. Each reward type proposed below can be modeled using the theoretical framework outlined previously by adjusting the reward parameter δ and the outcome bias parameter ρ. Table 3 reports the share of spectators who can be classified as a particular reward type and the

Figure 5: Distribution of Individual-level Treatment Effects



Notes: Each figure displays a density plot of individual-level treatment effects of all spectators by treatment condition. The ex post treatment condition is divided into scenarios where the externality realised and those, where it did not. Spectator-specific treatment effects are equal to the average difference in the absolute share of the bonus given to the risk-taker comparing control and treatment conditions where the safe option and the lottery outcome are identical.

average bonus share given to the risk-taker by treatment condition for each type. The table also reports the share of spectators within each group of reward types who can be classified as choice egalitarian or inequality averse based on their control decisions. Finally, the table reports the average level of confidence spectators of a particular type indicated to have when making their distributive choices in the treatment scenarios.

Consistent with the heterogeneity in individual-level treatment effects reported in figure 5, there is also substantial heterogeneity in reward types. At least five distinct types can be identified with a substantial share of spectators within each category. About 13% of spectators do not reward prosocial risk-taking at all and have an individual-level treatment effect of zero across all treatment conditions. A larger share, about 18% of spectators, always reward prosocial risk-taking, irrespective of whether the decision is made ex ante or ex post

Table 3: Reward Types

	Spectator Share	Avg. Bonus Ex Ante	Avg. Bonus Ex Post 0	Avg. Bonus Ex Post 1	Choice Egalitarian	Inequality Averse	Confidence
Always Zero	12.97%	0.00	0.00	0.00	54.17%	16.67%	8.34
Always Positive	18.38%	1.15	1.10	1.15	2.94%	29.41%	7.42
Ex Ante Only	13.51%	0.84	-0.69	-0.22	12%	16%	7.93
Success Reward	20.54%	0.17	-0.43	0.76	13.16%	26.32%	7.36
Failed Reward	12.97%	0.01	0.73	-0.31	8.33%	45.83%	7.08
No Type	21.62%	-0.46	-0.02	-0.07	10%	40%	7.66

Notes: Reward Types are based on individual-level treatment effects for the ex ante condition, ex post condition where the externality realised, and the ex post condition where the externality did not realise. Choice Egalitarian and Inequality Averse report the share of spectators that follow either fairness criteria with 80% consistency in the control. Here, Inequality Averse refers to weak inequality aversion as specified in table 4. Confidence reports the average level of self-reported confidence in the distribution decisions within all treatment conditions.

and whether the externality actually realised ex post. 13.50% of spectators only reward prosocial risk-taking ex ante. This group of spectators even has a slightly negative average treatment effect in the ex post scenarios. The largest share of spectators, about 20.50%, can be classified as rewarding success as they have a positive individual treatment effect in the ex post scenarios where the externality realised. This group also has a negative average treatment effect for scenarios the externality did not realise, suggesting that these spectators even punish prosocial risk-takers who were unsuccessful in generating the externality. The size of this negative average treatment effect is more than twice as large as the main treatment effect reported in figure 3. About 15% of spectators behave in exactly the opposite way and reward prosocial risk-takers ex post if the externality did not generate, but do not reward them if it did generate. Finally, a bit less than 22% of spectators cannot be categorised into any of the previously mentioned reward types. This group of spectators appears to reward prosocial risk-takers less in the ex ante than in the control condition but appear to have an individual-level treatment effect of around 0 in the ex post conditions. While, on average, prosocial risk-taking is therefore rewarded primarily in the ex ante and ex post treatment if the externality generated, these main results conceal significant heterogeneity in reward types across spectators.

Table 3 also reports interesting heterogeneity in fairness criteria over earned income across reward types. Specifically, the majority of spectators who never reward prosocial risk-taking are also choice egalitarians in the control condition. On the other hand, almost half of all spectators who reward prosocial risk-takers if the externality did not realise are inequality averse in the control condition. The largest group of spectators who can be classified as a reward type, those who reward prosocial risk-taking if the externality realised, do not show a distinct pattern in their fairness criteria. If anything, these spectators seem to be less likely to be inequality averse than most of the others.

IV.3.2 Consistency of Fairness Criteria

As discussed in section II, spectators are assumed to follow a personal fairness criterion over earned income in the control condition which they trade-off against a desire to reward prosocial risk-takers in the treatment conditions. This suggests that the share of spectators who can be classified as following particular fairness criteria over earned income should be lower in the treatment as opposed to the control scenarios, as spectators are motivated by more than their personal criteria in the treatment conditions.

Table 4 reports the number of subjects who follow a particular fairness criterion in the control condition as well as the share of those subjects who are consistent in following this criterion across treatments. A spectator is classified as following a particular fairness criterion if at least 80% or 60% of the decisions within a treatment condition are made in line with the particular criterion. This translates to at least four or three of five decisions in the control and ex ante condition and at least eight or six of ten decisions in the ex post condition, respectively. Out of a total of 185 spectators, 28 spectators can be classified as being choice egalitarians in 80% of control condition scenarios. This number increases to 52 if the consistency requirement is reduced to 60%. Much fewer spectators can be classified as strictly inequality averse, meaning that they equalise outcomes. 14 and 18 spectators follow this criterion under the 80% and 60% consistency, respectively. As strict inequality aversion is somewhat difficult to implement correctly, given that spectators have to calculate how much of the $4 bonus each decision makers would have to receive, I also test for the consistency of weak inequality aversion. This preference is specified as simply giving more to the decision maker who initially has lower earnings and is therefore much easier to implement. Here, 41 and 87 spectators can be classified as following weak inequality aversion in their decision making under the 80% and 60% consistency requirement, respectively.

Table 4: Consistency of fairness criteria

	Fairness criterion		
	Choice Egalitarian	Strict Inequality Aversion	Weak Inequality Aversion
80% consistency			
Ex Ante	64.29%	50.00%	54.55%
Ex Post	60.71%	50.00%	43.64%
Full	53.57%	28.57%	27.27%
Control	28	14	41
60% consistency			
Ex Ante	59.62%	50.00%	66.67%
Ex Post	53.85%	55.56%	72.41%
Full	44.23%	38.89%	51.72%
Control	52	18	87

Notes: The percentages indicate the share of spectators who followed a particular criterion in the control condition and also followed that criteria in one or both of the treatment conditions. Strict inequality aversion means that spectators equalised ultimate payoffs while weak inequality aversion means spectators gave more to the decision maker who had a lower initial payoff.

As table 4 illustrates, on average, the consistency across all three criteria decreases from control to the ex ante and then to the ex post treatment condition. In most cases, less than half of those who follow a particular criterion in the control condition are consistent in following this criterion across all treatment conditions.

Table A5 in the appendix provides an analysis of demographic variables associated with holding particular fairness criteria in the control group and table B3 estimates treatment effects by fairness criteria in the control condition. This second analysis suggests that only spectators who can already be classified as holding a fairness criterion in the control condition react to the treatment by increasing their reward given to the risk taker. This provides further support for a theoretical model in which spectators, on average, trade-off a fixed fairness criterion with a positive reward parameter.

IV.4 Spectator Beliefs

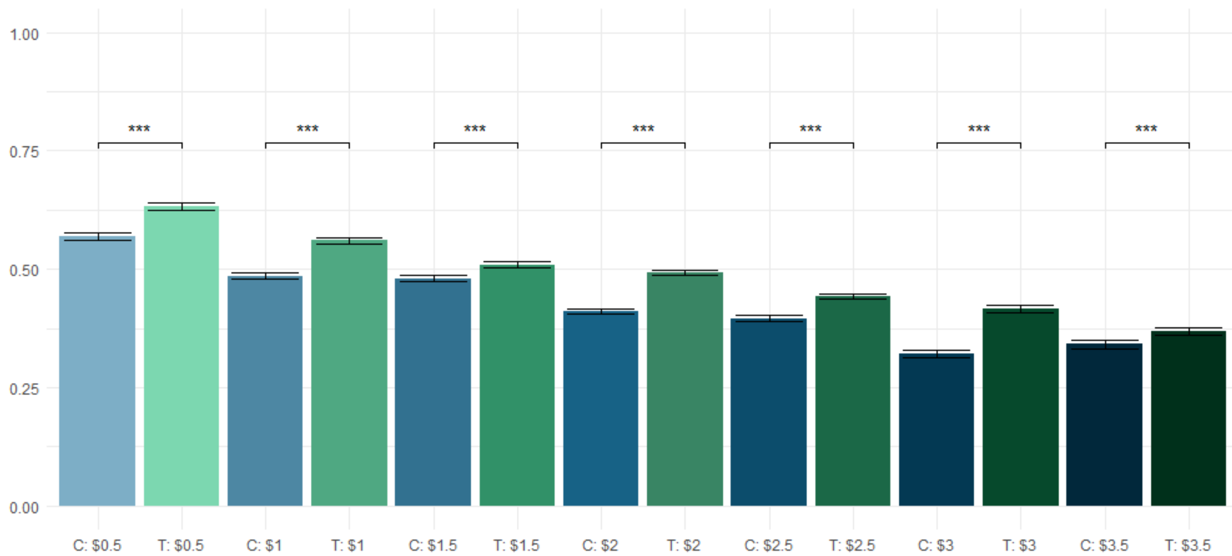
To better understand the main treatment effects, spectators were also asked why they believed decision makers chose the safe or risky option, and what they believe to be the effects of the generated externality for the participant pool. Additionally, they were asked to provide incentivised estimates of the number of decision makers who chose the risky option for each given value of the safe option. Appendix B.3 reports the full results of these belief elicitations.

Figure 6 reports incentivised spectator beliefs about the share of decision makers who chose the risky option for each given level of the safe option by treatment condition. Although spectators believe decision makers were significantly more likely to choose the risky option in the treatment than control condition, these differences are, in fact, much smaller than the actual differences in decision maker choices reported in figure 2. Most strikingly however, when comparing these two figures, is that spectators vastly underestimate how much decision makers respond to the opportunity cost of risk-taking at the different levels of the safe option A. While the difference in the share of decision makers who chose the risky option in the scenario with a safe option of $0.5 compared to the scenario with the safe option of $3.5 is 88.52% in the control and 86.98% in the treatment condition, spectators estimate these same differences to only be equal to 22.83% and 26.54%, respectively. Spectators underestimate the likelihood of decision makers taking risk when the opportunity cost of doing so is low but also overestimate the likelihood of decision makers taking risk when the opportunity cost is high. Importantly, this is the case for the control as well as treatment conditions. On average, spectators estimate the average first-stage treatment effect to be an increase in risk-taking of 5.98 percentage points, while the actual increase in risk-taking is 10.26 percentage points, almost twice as large.

Taken together, these findings suggest that spectators (1) underestimate how strongly decision makers take opportunity costs into account and (2) underestimate how strongly decision makers react to the possibility to generate positive externalities for others. This implies that the main results likely underestimate the treatment effect one would observe if spectators were well-informed about decision makers' choices. Given that spectators do not personally benefit from allocating less of the bonus to risk-takers and receive additional rewards if the provided estimates are correct, it is also unlikely that these beliefs are motivated in some way.

A question naturally following from the previous findings is what spectators believe to be decision-makers' motives in choosing the risky option. Here, incentivising responses is, of

Figure 6: Spectator Beliefs about share of risk-taking decision makers by treatment condition



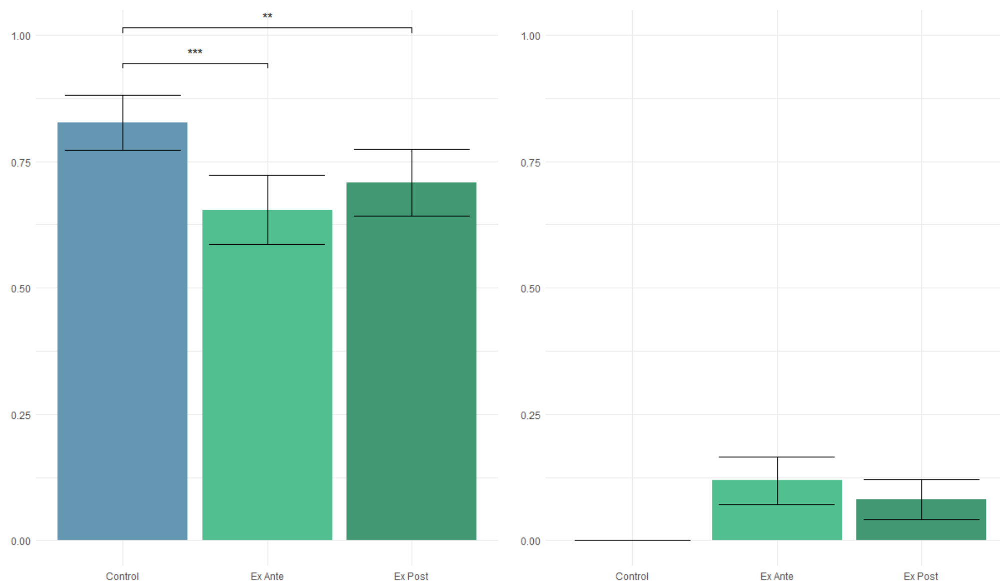
Notes: The figure reports the share of decision makers spectators believe chose the risky option B, given the amount of the safe option and treatment condition. For example, "C: $1" reports the share of decision makers spectators belief chose the risky option if the safe option was equal to $1 and there was no possibility to generate externalities. Significance levels indicate the results of simple t-test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

course, not possible as there is no way to identify decision-makers true motives. Figure 7 therefore provides the results of a non-incentivised belief elicitation. Spectators were asked to state why they believe some decision makers chose the risky option B. For comparability, they were asked to select one of multiple options they most agreed with. The left panel of figure 7 reports the share of spectators who selected the option "To maximise their own payoff" by treatment condition. The right figure does the same for the option "To generate the additional $2 for the other participants". These two statements are the only statements on which spectator beliefs differ significantly between control and treatment conditions.[6]

As figure 7 illustrates, spectators are significantly less likely to believe decision makers want to maximise their own income by taking the risky option in the treatment as opposed to the control condition. This seems to be explained by some spectators believing that the main reason decision makers took the risky option was to generate the externality for others in the treatment conditions. This suggests that the treatment effect, on average, might be driven

[6]The full results of this belief elicitation can be found in Appendix B.3.

Figure 7: Spectator Beliefs about why decision makers took risk



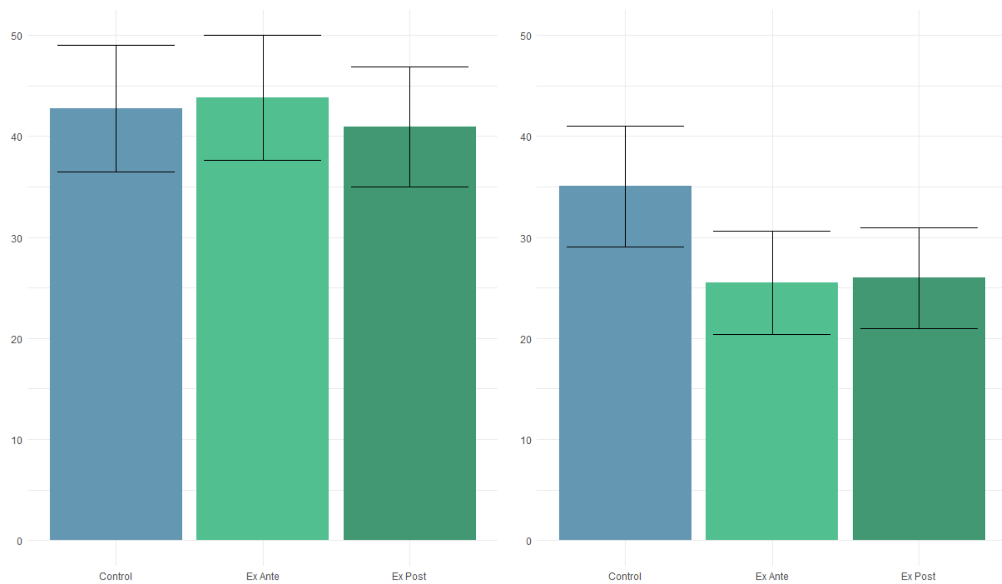
Notes: Spectators were asked why they think some participants chose the risky option B. The left figure reports the share of spectators who selected the option "To maximise their own payoff" by treatment condition. The right figure does the same for the option "To generate the additional $2 for the other participants", which was however only an option in the treatment conditions.

by a change in the motives spectators believe risk-takers to have when choosing the risky option. While not statistically significant, it is also interesting to note that this difference on both measures is larger between the control and ex ante conditions as opposed to the control and ex post conditions.

Finally, spectators were also asked which attributes of the scenarios they took into account when making their allocation decisions. Here, they were asked to allocate a total of 100 points across multiple options. The full results of this elicitation can be found in appendix B.3.

Figure 8 reports the average points allocated to two options: The left panel of the figure reports the results for the attribute "Inequality between participants" while the right panel does the same for the attribute "The choices participants made". These two options were specifically designed to test for inequality aversion and choice egalitarianism in spectator reasoning. While there appears to be a drop in the points allocated to "The choices participants made" between the control and treatment conditions, this difference is not statistically

Figure 8: What mattered to your allocation decision? - 100 points



Notes: Spectators were asked to allocate 100 points between a selection of attributes depending on the significance of the attributes for the spectator's decision maker in each treatment condition. The left figure reports the average points allocated to the attribute "Inequality between participants" by treatment condition while the right figure does the same for the attribute "The choices participants made".

significant. The points allocated to "Inequality between participants" are strikingly consistent across treatments.

In line with the discussion in section IV.3., these findings provide further support for the theoretical assumption that spectators trade-off their fairness criterion with their desire to reward prosocial risk-taking.

V Conclusion

Rewarding prosocial risk-taking may be a crucial strategy to address societal challenges such as the climate emergency. The aim of this paper has been to provide a theoretical framework and first experimental test of how prosocial risk-taking, and potential externalities more broadly, affect distributive decision making. The results of this paper suggest that individuals do, in fact, reward prosocial risk-taking. On average, spectators increase the share of the bonus allocated to the risk-taker from 49% to 53% if risk-taking potentially generates positive externalities for others. The treatment effect also increases as the opportunity cost of taking risk increases.

This main treatment effect however masks outcome bias and substantial heterogeneity in distributive decision-making. While individuals reward prosocial risk-taking if outcomes are unknown; they only reward successful prosocial risk-takers ex post and do not compensate unlucky risk-takers who took on prosocial risk. That is, despite the fact that individuals are more likely to believe that risk-takers have altruistic motives when there are potential positive externalities resulting from taking the risk. At least five distinct reward types of spectators can also be identified, the largest group consisting of spectators who only reward prosocial risk-taking if successful.

These results have potential implications for both, policy making and economic theory. If prosocial risk-taking is a behaviour societies wish to reward, then the results of this paper suggest that support for such policies will likely be unstable. While a large share of people may have a preference to reward prosocial risk-taking ex ante, on average, a policy in line with this preference will likely lose support ex post, if the potential positive externalities did not realise. Important questions for further research are therefore whether this instability in policy support has effects on the prevalence of prosocial risk-taking itself and whether there are ways to reduce the outcome bias in people's preferences over rewarding prosocial risk-takers.

Finally, the results of this paper demonstrate the importance of a dimension of distributive decision-making which has, so far, not received much attention: Externalities. This paper has looked at prosocial risk-taking, which focuses on stochastic choices and positive externalities. Often, externalities however also have negative implications for overall welfare and can affect distributive decision-making in deterministic settings. The size and type of externalities might also be important to distributive choices and their relevance might vary across different reward types. There is therefore a lot of scope for future research on the role of externalities in distributive decision making. The aim of this paper has been to provide a theoretical framework and an experimental design which can be adapted to study many more aspects of externalities in distributive decision-making.

References

Acs, Z. J., P. Braunerhjelm, D. B. Audretsch, and B. Carlsson (2009): "The knowledge spillover theory of entrepreneurship," *Small business economics*, 32, 15–30.

Akcigit, U., D. Hanley, and S. Stantcheva (2022): "Optimal taxation and R&D policies," *Econometrica*, 90, 645–684.

Akcigit, U. and S. Stantcheva (2020): "Taxation and Innovation: What Do We Know?" Tech. rep., National Bureau of Economic Research.

Alesina, A. and G.-M. Angeletos (2005): "Fairness and redistribution," *American economic review*, 95, 960–980.

Alesina, A. and P. Giuliano (2011): "Preferences for Redistribution," in *Handbook of Social Economics*, Elsevier, vol. 1, 93–131.

Alesina, A., A. Miano, and S. Stantcheva (2018a): "Immigration and redistribution," Tech. rep., National Bureau of Economic Research.

Alesina, A., S. Stantcheva, and E. Teso (2018b): "Intergenerational mobility and preferences for redistribution," *American Economic Review*, 108, 521–54.

Almås, I., A. W. Cappelen, and B. Tungodden (2020): "Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians?" *Journal of Political Economy*, 128, 1753–1788.

Andre, P. (2021): "Shallow Meritocracy: An Experiment on Fairness Views," *Available at SSRN 3916303*.

Brownback, A. and M. A. Kuhn (2019): "Understanding outcome bias," *Games and Economic Behavior*, 117, 342–360.

Cappelen, A. W., S. Fest, E. Ø. Sørensen, and B. Tungodden (2016): "Choice and personal responsibility: What is a morally relevant choice?" *The Review of Economics and Statistics*, 1–35.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007): "The pluralism of fairness ideals: An experimental approach," *American Economic Review*, 97, 818–827.

Cappelen, A. W., J. Konow, E. Ø. Sørensen, and B. Tungodden (2013a): "Just luck: An experimental study of risk-taking and fairness," *American Economic Review*, 103, 1398–1413.

Cappelen, A. W., K. O. Moene, E. Ø. Sørensen, and B. Tungodden (2013b): "Needs versus entitlements—an international fairness experiment," *Journal of the European Economic Association*, 11, 574–598.

Cappelen, A. W., E. Ø. Sørensen, and B. Tungodden (2010): "Responsibility for what? Fairness and individual responsibility," *European Economic Review*, 54, 429–441.

Cavatorta, E. and D. Schröder (2019): "Measuring ambiguity preferences: A new ambiguity preference survey module," *Journal of Risk and Uncertainty*, 58, 71–100.

Charness, G., U. Gneezy, and B. Halladay (2016): "Experimental methods: Pay one or pay all," *Journal of Economic Behavior & Organization*, 131, 141–150.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002): "Hardnose the dictator," *American Economic Review*, 92, 1218–1221.

Cherry, T. L. and J. F. Shogren (2008): "Self-interest, sympathy and the origin of endowments," *Economics Letters*, 101, 69–72.

Chowdhury, S. M. and J. Y. Jeon (2014): "Impure altruism or inequality aversion?: An experimental investigation based on income effects," *Journal of Public Economics*, 118, 143–150.

Corneo, G. and H. P. Grüner (2002): "Individual preferences for political redistribution," *Journal of Public Economics*, 83, 83–107.

Da Rin, M., M. Di Giacomo, and A. Sembenelli (2011): "Entrepreneurship, firm entry, and the taxation of corporate income: Evidence from Europe," *Journal of public economics*, 95, 1048–1066.

de Oliveira, A. (2021): "When risky decisions generate externalities," *Journal of Risk and Uncertainty*, 63, 59–79.

de Oliveira, A. C., A. Smith, and J. Spraggon (2017): "Reward the lucky? An experimental investigation of the impact of agency and luck on bonuses," *Journal of Economic Psychology*, 62, 87–97.

Djankov, S., T. Ganser, C. McLiesh, R. Ramalho, and A. Shleifer (2010): "The effect of corporate taxes on investment and entrepreneurship," *American Economic Journal: Macroeconomics*, 2, 31–64.

Do, K. T., J. F. G. Moreira, and E. H. Telzer (2017): "But is helping you worth the risk? Defining prosocial risk taking in adolescence," *Developmental cognitive neuroscience*, 25, 260–271.

Durante, R., L. Putterman, and J. Van der Weele (2014): "Preferences for redistribution and perception of fairness: An experimental study," *Journal of the European Economic Association*, 12, 1059–1086.

Embry, E., J. Jones, and J. G. York (2019): "Climate change and entrepreneurship," in *Handbook of Inclusive Innovation*, Edward Elgar Publishing, 377–393.

Falk, A., E. Fehr, and U. Fischbacher (2003): "On the nature of fair behavior," *Economic inquiry*, 41, 20–26.

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

Gee, L. K., M. Migueis, and S. Parsa (2017): "Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness," *Experimental Economics*, 20, 894–923.

Gurdal, M. Y., J. B. Miller, and A. Rustichini (2013): "Why blame?" *Journal of Political Economy*, 121, 1205–1247.

Heap, S. P. H., E. Levi, and A. Ramalingam (2021): "Does group identification promote giving with respect to windfall gains in a dictator game?" .

Klor, E. F. and M. Shayo (2010): "Social identity and preferences over redistribution," *Journal of Public Economics*, 94, 269–278.

König-Kersting, C., M. Pollmann, J. Potters, and S. T. Trautmann (2021): "Good decision vs. good results: Outcome bias in the evaluation of financial agents," *Theory and Decision*, 90, 31–61.

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

Kuziemko, I., M. I. Norton, E. Saez, and S. Stantcheva (2015): "How elastic are preferences for redistribution? Evidence from randomized survey experiments," *American Economic Review*, 105, 1478–1508.

Lefgren, L. J., D. P. Sims, and O. B. Stoddard (2016): "Effort, luck, and voting for redistribution," *Journal of Public Economics*, 143, 89–97.

Luttmer, E. F. and M. Singhal (2011): "Culture, context, and the taste for redistribution," *American Economic Journal: Economic Policy*, 3, 157–79.

Mollerstrom, J., B.-A. Reme, and E. Ø. Sørensen (2015): "Luck, choice and responsibility—An experimental study of fairness views," *Journal of Public Economics*, 131, 33–40.

Rey-Biel, P., R. Sheremeta, and N. Uler (2018): *When income depends on performance and luck: The effects of culture and information on giving*, Emerald Publishing Limited.

Rodriguez-Lara, I. and L. Moreno-Garrido (2012): "Self-interest and fairness: self-serving choices of justice principles," *Experimental Economics*, 15, 158–175.

Ruffle, B. J. (1998): "More is better, but fair is fair: Tipping in dictator and ultimatum games," *Games and Economic Behavior*, 23, 247–265.

Stantcheva, S. (2021a): "Perceptions and preferences for redistribution," Tech. rep., National Bureau of Economic Research.

——— (2021b): "Understanding tax policy: How do people reason?" *The Quarterly Journal of Economics*, 136, 2309–2369.

A Descriptive Analysis of Experimental Data

A.1 Balance Test of First Stage Treatment Assignment

Although decision makers are not the primary subjects of the experiment, table A1 reports a balance test for decision makers in the first stage of the experiment by random treatment assignment. Detailed descriptions of each variable can be found in appendix C. Interestingly, quite a few variables are not properly balanced in the decision maker sample despite the random assignment. Specifically, decision makers in the treatment group are more likely to be female, less likely to have studied economics at university, somewhat less likely to be black or African-American, and display slightly less ambiguity aversion.

To test whether prosocial risk-taking in the first stage is driven by any of these underlying demographic differences between the samples, table A2 reports the results of simple and logit regressions with the switch point of decision makers as the outcome variables. The switchpoint is hereby defined as the level of the safe option at which decision makers moved from the risky to the safe option in the first two models and as a binary variable dependent on whether the decision maker switched before or after the safe option is equal to $2. This binary specification of the switchpoint aims to capture risk-seeking behaviour within the personal lottery decision.

The results of these regressions show that the main first stage treatment effect is robust to adding the required control variables based on the balance test in table A1. Decision makers in the treatment group are significantly more likely to switch to the safe option at a higher level of the safe option A, and are also substantially more likely to display risk neutral or risk seeking behaviour by switching at a level of the safe option above $1.50. On average, the highly significant coefficient of 0.862 in the full model suggests that decision makers in the treatment group switch almost one level later than decision makers in the control group, even when accounting for the unbalanced variables in the sample.

When looking at the treatment effects interacted with each of the unbalanced variables, some interesting patterns emerge. Although male decision makers react more strongly than female decision makers to the treatment by switching later in the individual interactions model, this is not robust to any of the other specifications. Decision makers who studied economics at university, however, are more likely to switch at a later point in the treatment group in both interaction models. They are no more likely to become risk neutral or risk seeking in the treatment group though. The opposite pattern can be observed for black and African-American decision makers. They are substantially more likely to become risk neutral or risk-seeking when risk-taking leads to potential positive externalities for others,

Table A1: Balance Test by treatment assignment of decision makers

	Control	Treatment	t-statistic
Demographics			
Age	40.37	40.37	-0.009
	(13.78)	(14.09)	(2,518)
Gender	0.61	0.45	7.83***
	(0.49)	(0.50)	(2,469)
Education	4.40	4.47	-1.07
	(1.47)	(1.47)	(2,497)
Income	54,960	58,665	-1.81*
	(50,040)	(47,334)	(2,469)
Employment	0.70	0.66	2.09**
	(0.46)	(0.47)	(2,518)
Economics	0.47	0.35	5.59***
	(0.02)	(0.01)	(2,518)
Ethnicity			
American Indian or Alaska Native	0.00	0.00	-1.90*
	(0.00)	(0.00)	(2,518)
Asian	0.10	0.12	-1.76*
	(0.30)	(0.33)	(2,518)
Black or African-American	0.16	0.11	3.35***
	(0.36)	(0.31)	(2,518)
Native Hawaiian or Pacific Islander	0.00	0.00	-1.90*
	(0.00)	(0.06)	(2,518)
Spanish, Hispanic or Latino	0.07	0.05	1.90*
	(0.26)	(0.23)	(2,518)
White	0.72	0.73	-0.52
	(0.45)	(0.44)	(2,518)
Preferences			
Risk seeking	4.78	4.72	0.60
	(2.18)	(2.48)	(2,518)
Ambiguity aversion	16.63	15.66	4.10***
	(5.72)	(5.56)	(2,518)
Left-Right Placement	3.65	3.55	0.78
	(2.76)	(2.83)	(2,462)
Party affiliation			
Democrats	0.61	0.63	-0.76
	(0.49)	(0.48)	(2,518)
Republicans	0.18	0.20	-1.28
	(0.38)	(0.40)	(2,518)
Other	0.11	0.12	-0.21
	(0.32)	(0.32)	(2,518)
2020 Vote			
Joe Biden	0.57	0.54	1.13
	(0.50)	(0.50)	(2,518)
Donald Trump	0.17	0.16	0.80
	(0.38)	(0.37)	(2,518)
Other	0.02	0.03	-1.42
	(0.13)	(0.16)	(2,518)
Didn't vote	0.21	0.22	-0.31
	(0.41)	(0.41)	(2,518)

Notes: Table reports the mean values for decision makers based on treatment group assignment. Detailed descriptions of each variable can be found in appendix section C. Asterisks indicate significant differences in mean values between samples from a simple test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

compared to other ethnicities. Finally, ambiguity aversion does not seem to have a substantial impact on the treatment effect.

Given that decision makers in the treatment group are more likely to be female, less likely to have studied economics, and less likely to be black or African-American, these results actually indicate that the first stage treatment effect reported in the main text is most likely an underestimate of the true treatment effect. That is, as each of these demographic groups more present in the treatment than control group also react less strongly to the treatment.

Table A2: Prosocial Risk-taking of Decision Makers with Controls

	Full Results			Above $1.50		
	Indv. Interacted	Full Interacted	Controls	Indv. Interacted	Full Interacted	Controls
Treatment	-	-	0.862***	-	-	0.659***
	-	-	(0.185)	-	-	(0.249)
Male	0.516***	0.230	0.336*	0.255	0.029	0.064
	(0.195)	(0.229)	(0.187)	(0.237)	(0.268)	(0.226)
Economics	0.777***	0.670***	0.297	0.438*	0.351	0.074
	(0.221)	(0.248)	(0.191)	(0.253)	(0.288)	(0.233)
Black	0.774**	0.520*	0.131	1.201***	1.103**	0.580*
	(0.312)	(0.309)	(0.250)	(0.428)	(0.438)	(0.320)
Ambiguity	0.027***	0.008	-0.025	0.022*	0.010	-0.009
	(0.010)	(0.011)	(0.016)	(0.012)	(0.014)	(0.020)
Constant	-	3.616***	3.481***	-	-0.746***	-0.893**
	-	(0.136)	(0.332)	-	(0.174)	(0.414)
Observations	326	326	322	358	358	353

Notes: Estimates come from linear regressions. Models 1 and 4 report results from individual regressions with interactions between the specific control variable and the treatment. Models 2 and 5 report results from one model in which the treatment is interacted with each control variable separately. Models 3 and 6 report a simple model in which the variables are added as control variables. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

A.2 Balance Test of Second Stage Treatment Assignment

As the second stage of the experiment consists of a within subject design, whereby spectators see all three treatment conditions but in randomised order, a balance test by treatment assignment cannot be conducted. However, to test whether the randomised viewing order is correlated with any demographic variables, table A3 reports the results of a balance test by the first treatment condition spectators were randomly assigned to. This can be conducted as spectators made allocation decisions in randomly ordered treatment blocks. In other words, spectators, for example, first made five allocation decisions in the control scenario, then five in the ex ante, and then ten in the ex post treatment condition. The table therefore reports the results of a balance test based on the first randomly shown treatment block. The significance tests are always conducted between the control and the respective treatment

Table A3: Balance Test by first treatment condition of spectators

	Control	Ex Ante	t-statistic	Ex Post	t-statistic
Demographics					
Age	39.27	40.20	-1.52	41.67	-4.20***
	(14.20)	(15.59)	(2,358)	(14.89)	(2,598)
Gender	0.46	0.32	6.91***	0.37	4.52***
	(0.50)	(0.47)	(2,318)	(0.48)	(2,598)
Education	4.06	4.62	-9.74***	4.52	-7.40***
	(1.50)	(1.23)	(2,358)	(1.65)	(2,598)
Income	48,246	49,020	-0.53	57,232	-5.94***
	(32,248)	(35,732)	(2,158)	(41,173)	(2,438)
Employment	0.67	0.65	0.62	0.67	-0.27
	(0.47)	(0.48)	(2,358)	(0.47)	(2,598)
Economics	0.38	0.51	-6.30***	0.39	-0.37
	(0.49)	(0.50)	(2,358)	(0.49)	(2,598)
Ethnicity					
American Indian or Alaska Native	-	-	-	0.03	-6.22***
	-	-	-	(0.17)	(2,598)
Asian	0.08	0.11	-2.48**	0.12	-3.41***
	(0.27)	(0.31)	(2,358)	(0.32)	(2,598)
Black or African-American	0.11	0.13	-1.21	0.10	0.55
	(0.31)	(0.33)	(2,358)	(0.31)	(2,598)
Native Hawaiian or Pacific Islander	-	-	-	-	-
	-	-	-	-	-
Spanish, Hispanic or Latino	0.08	0.20	-8.67***	0.18	-7.62***
	(0.27)	(0.40)	(2,358)	(0.38)	(2,598)
White	0.73	0.65	3.99***	0.66	4.07***
	(0.44)	(0.48)	(2,358)	(0.47)	(2,598)
Preferences					
Risk seeking	5.11	5.49	-3.89***	5.02	1.09
	(2.30)	(2.44)	(2,358)	(2.13)	(2,558)
Ambiguity aversion	15.52	15.82	-1.28	15.34	0.81
	(6.03)	(5.00)	(2,358)	(5.28)	(2,598)
Left-Right Placement	3.56	3.85	-2.78***	3.82	-2.37**
	(2.69)	(2.42)	(2,338)	(2.85)	(2,558)
Party affiliation					
Democrats	0.52	0.47	2.48**	0.61	-4.55***
	(0.50)	(0.50)	(2,358)	(0.49)	(2,598)
Republicans	0.22	0.20	1.32	0.19	1.77*
	(0.42)	(0.40)	(2,358)	(0.40)	(2,598)
Other	0.22	0.27	-2.85***	0.16	3.76***
	(0.42)	(0.45)	(2,358)	(0.37)	(2,598)
2020 Vote					
Joe Biden	0.60	0.47	6.40***	0.54	3.40***
	(0.49)	(0.50)	(2,358)	(0.50)	(2,598)
Donald Trump	0.13	0.20	-4.84***	0.15	-1.64
	(0.33)	(0.40)	(2,358)	(0.36)	(2,598)
Other	0.05	0.04	1.35	0.01	4.84***
	(0.21)	(0.19)	(2,358)	(0.12)	(2,598)
Didn't vote	0.14	0.25	-6.90***	0.27	-8.00***
	(0.35)	(0.44)	(2,358)	(0.44)	(2,598)

Notes: Table reports the mean values for spectators based on the first treatment condition block they were randomly allocated to make distribution decisions for. Asterisks indicate significant differences in mean values between samples from a simple test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

conditions.

Similarly to the random treatment assignment of decision makers, there are multiple demographic variables which are correlated with the randomly assigned first treatment condition spectators saw. However, only some are significantly different in both, the ex ante and ex post conditions, compared to the control. Given that the primary concern about the viewing order would be that being informed of the potential treatment externalities before completing the control condition might bias results, this aggregate comparison between seeing either of the treatment as opposed to the control condition first seems most important. Specifically, spectators who first completed either of the treatment conditions were more likely to be female, more highly educated, very slightly more likely to be asian and spanish, hispanic, or latino, somewhat more likely to place themselves further towards the right of a political left-right scale, somewhat less likely to have voted for Joe Biden, and more likely to not have voted at all in the 2020 election. Simiarly to table A2, table A4 reports the results of simple regressions with the absolute share of the bonus allocated to the risk taking subject as the outcome variable.

Table A4: Main Treatment Effects with Controls

	Ex Ante			Ex Post		
	Indv. Interacted	Full Interacted	Controls	Indv. Interacted	Full Interacted	Controls
Treatment	-	-	0.142***	-	-	0.149***
	-	-	(0.049)	-	-	(0.043)
Male	0.157**	0.103	0.066	0.066	-0.013	0.005
	(0.066)	(0.073)	(0.053)	(0.049)	(0.052)	(0.043)
Education	0.027***	0.011	-0.018	0.038***	0.048***	0.021
	(0.010)	(0.018)	(0.016)	(0.008)	(0.012)	(0.013)
Asian	0.223**	0.171	0.053	0.176**	0.183*	0.103
	(0.112)	(0.126)	(0.092)	(0.086)	(0.095)	(0.077)
Spanish, Hispanic, Latino	-0.089	-0.142*	-0.132**	0.051	0.053	-0.009
	(0.081)	(0.086)	(0.065)	(0.063)	(0.067)	(0.056)
Left-Right Placement	0.017*	0.006	-0.019	0.018**	0.001	-0.010
	(0.010)	(0.013)	(0.012)	(0.007)	(0.009)	(0.010)
Joe Biden Vote	0.122**	0.058	-0.156**	0.091**	-0.088	-0.161**
	(0.052)	(0.079)	(0.079)	(0.041)	(0.056)	(0.063)
No Vote	-0.022	-0.068	-0.136	-0.106*	-0.217***	-0.202***
	(0.075)	(0.095)	(0.083)	(0.057)	(0.068)	(0.068)
Constant	-	1.962***	2.208***	-	1.958***	2.019***
	-	(0.034)	(0.127)	-	(0.034)	(0.107)
Observations	1,820	1,820	1,800	2,730	2,730	2,700

Notes: Estimates come from linear regressions. Models 1 and 4 report results from individual regressions with interactions between the specific control variable and the treatment. Models 2 and 5 report results from one model in which the treatment is interacted with each control variable separately. Models 3 and 6 report a simple model in which the variables are added as control variables. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

The first result to note is that the main treatment effects are robust to including all the variables as controls that are unbalanced in the sample. Additionally, none of the variables show a consistent treatment effect interaction across all the reported models. The only variable showing a consistent and significant effect across all models is the "no vote" variable in the ex post treatment condition. Those spectators who self-reported to not have voted in the 2020 election, on average, are less likely to reward the risk taker in the ex post treatment condition than in the control. They are also, however, less likely to reward the risk taker across the control and ex post conditions, as the negative and significant coefficient in the control model shows.

A.3 Demographics by Fairness Criteria in Control Condition

Table A5 reports differences in demographics between spectators who were identified as having either a choice egalitarian or weakly inequality averse fairness criterion. For the purpose of this comparison, the 60% consistency requirement was used, meaning that spectators had to make at least three of the five control allocation decisions in line with the criterion. As fairness criteria are endogenous, the results in table A5 are not relevant for the robustness of any of the main treatment effects.

Spectators who have a choice egalitarian criterion in the control condition are somewhat older, more likely to be male, less educated, have a lower household income, are more likely to be employed, and are less likely to have studied economics. There are also more likely to be white and less likely to be asian and, although only weakly significanly so, less likely to be spanish, hispanic, or latino. Arguably, the sample size for American Indian or Alaska Native is too small to justify an interpretation of the significance test for this variable. Choice egalitarians are also less likely to be ambiguity averse and less likely to be Republicans than inequality averse spectators.

B Additional Results

B.1 Main effects with different specifications

B.1.1 Treatment effects by confidence

Table B1 reports interactions of treatment and levels of confidence in the allocation decisions. Spectators were asked about their level of confidence in their decision on a ten-point scale after each allocation scenario.

The results show that the treatment effect is stronger for spectators who are more confident in their allocation decisions. This result is comparable for both treatment conditions and also increases somewhat when focusing only on those scenarios where the safe option is above

Table A5: Demographics by revealed fairness criteria of spectators

	Choice Egalitarian	Inequality Aversion	t-statistic
Demographics			
Age	40.96	39.29	-2.91***
	(15.19)	(14.35)	(2,778)
Gender	0.33	0.38	3.01***
	(0.47)	(0.49)	(2,758)
Education	4.15	4.45	4.95***
	(1.62)	(1.45)	(2,778)
Income	45,401	52,840	5.06***
	(31,206)	(39,492)	(2,618)
Employment	0.69	0.64	-2.63***
	(0.46)	(0.48)	(2,778)
Economics	0.38	0.43	2.11**
	(0.49)	(0.49)	(2,778)
Ethnicity			
American Indian or Alaska Native	0.02	0.00	-5.84***
	(0.14)	(0.00)	(2,778)
Asian	0.02	0.15	11.28***
	(0.00)	(0.01)	(2,778)
Black or African-American	0.13	0.13	-0.62
	(0.34)	(0.33)	(2,778)
Native Hawaiian or Pacific Islander	-	-	-
	-	-	-
Spanish, Hispanic or Latino	0.13	0.16	1.87*
	(0.34)	(0.37)	(2,778)
White	0.71	0.66	-3.08***
	(0.45)	(0.48)	(2,778)
Preferences			
Risk seeking	5.17	5.13	-0.53
	(2.38)	(2.28)	(2,738)
Ambiguity aversion	14.67	15.47	3.68***
	(6.03)	(5.21)	(2,778)
Left-Right Placement	3.50	3.57	0.69
	(2.40)	(2.63)	(2,718)
Party affiliation			
Democrats	0.60	0.56	-1.70*
	(0.49)	(0.50)	(2,778)
Republicans	0.13	0.18	3.39***
	(0.34)	(0.39)	(2,778)
Other	0.21	0.23	1.12
	(0.41)	(0.42)	(2,778)
2020 Vote			
Joe Biden	0.58	0.57	-0.11
	(0.49)	(0.49)	(2,778)
Donald Trump	0.12	0.14	1.71*
	(0.32)	(0.34)	(2,778)
Other	0.00	0.03	6.09***
	(0.00)	(0.18)	(2,778)
Didn't vote	0.23	0.23	-0.05
	(0.42)	(0.42)	(2,778)

Notes: Table reports the mean values for spectators based on their revealed fairness criteria in the control condition assuming 60% consistency and using the weak definition of inequality aversion. Asterisks indicate significant differences in mean values between samples from a simple test of significance (with degrees of freedom in parentheses). Standard deviations are below the means, in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

$1. This result is encouraging as it suggests that the treatment effect is due to a conscious choice by spectators.

Table B1: Treatment interactions with level of confidence

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment x Confidence	0.016***	0.015***	0.023***	0.022***
	(0.006)	(0.005)	(0.007)	(0.006)
Constant	1.963***	1.975***	2.000***	2.005***
	(0.033)	(0.033)	(0.038)	(0.037)
Observations	1,833	2,731	1,309	1,949

Notes: Estimates come from linear regressions. Confidence is self-reported and ranges from 1 to 10 with higher values indicating more confidence in the allocation decision. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

B.1.2 Treatment effects by understanding

Table B2 reports interactions of treatment assignment and the level of understanding of spectators. Prior to making their first allocation decisions, spectators were asked four questions about the scenarios they would face. After answering each question, they were provided with the correct answer to each question. The specific questions can be found in appendix D.

Table B2: Treatment interactions with degree of understanding

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment x Understanding	0.061***	0.068***	0.075***	0.079***
	(0.014)	(0.012)	(0.016)	(0.014)
Constant	1.928***	1.910***	1.971***	1.957***
	(0.033)	(0.032)	(0.038)	(0.036)
Observations	1,850	2,775	1,322	1,982

Notes: Estimates come from linear regressions. Understanding ranges from 0 to 4 depending on the number of questions correctly answered. A higher value indicates a better understanding of the decision problem. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

In table B2, the understanding variable is coded on a five-point scale with 0 for spectators who got all questions wrong and 4 for spectators who correctly answered all questions. As the table reports, a higher value on this understanding scale increases the treatment effect for both treatments. Similarly to the confidence models in table B1, this effect is even stronger when focusing only on the scenarios where the safe option is above $1. Again, these results are encouraging as they suggest that a better understanding of the decision problem led to spectators increasing their reward for prosocial risk-taking.

B.1.3 Treatment effects by fairness criteria

The experimental design allows for spectators to be classified by their fairness criteria in the control condition. Table B3 therefore reports interactions between control condition fairness criteria and treatment effects. In other words, the models reported in the table test whether certain types of spectators react more strongly to the treatment than others.

Both, choice egalitarians and weakly inequality averse spectators have a robust and highly significant positive treatment effect across both, the ex ante and the ex post treatment conditions. Interestingly, spectators who cannot be classified as either choice egalitarians or as inequality averse only have a robustly positive and significant treatment effect in the models when using a strict definition of inequality aversion. Under the 60% consistency requirement for weak inequality aversion the treatment effect for this subject group disappears entirely. This suggests that the treatment effect is entirely driven by spectators whose allocation decisions in the control group already follow a version of the two fairness criteria. This result is especially surprising considering that even under the 60% consistency requirement and the weak definition of inequality aversion, still 46 out of 185 spectators cannot be classified as following a particular fairness criterion so given that each spectator contributes a total of 20 allocation decisions, the lack of a significant treatment effect for this group is unlikely to be an issue of power. Additionally, the coefficients are substantially smaller and at times negative for this group, again suggesting that there really is no treatment effect in either the ex post or ex ante conditions for this group of spectators.

This finding also provides additional support for the proposed theoretical model assuming that spectators have a given fairness criterion in the control condition which they trade-off against a reward parameter in the treatment conditions.

B.1.4 Treatment effects with demographic controls

Although table A4 already reports main treatment effects with controls based on the unbalanced variables reported in table A3, table B4 below reports main treatment effects for the full sample and for scenarios with the safe option above $1 including all available control variables in the dataset. The first result to note is again that the main treatment effects

Table B3: Treatment interactions with fairness criteria

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
80% consistency				
Treatment x CE	0.137***	0.158***	0.165***	0.201***
	(0.050)	(0.049)	(0.061)	(0.060)
Treatment x strict IA	-0.048	0.119	0.091	0.290***
	(0.120)	(0.097)	(0.130)	(0.104)
Treatment x NF	0.174***	0.161***	0.227***	0.192***
	(0.054)	(0.046)	(0.062)	(0.052)
Treatment x CE	0.137***	0.158***	0.165***	0.201***
	(0.050)	(0.049)	(0.061)	(0.060)
Treatment x weak IA	0.269***	0.199***	0.413***	0.354***
	(0.070)	(0.059)	(0.076)	(0.066)
Treatment x NF	0.092	0.134***	0.100	0.119**
	(0.061)	(0.049)	(0.071)	(0.057)
60% consistency				
Treatment x CE	0.131**	0.130***	0.199***	0.123**
	(0.052)	(0.047)	(0.059)	(0.056)
Treatment x strict IA	0.052	0.170*	0.190	0.357***
	(0.109)	(0.090)	(0.118)	(0.097)
Treatment x NF	0.176***	0.167***	0.214***	0.212***
	(0.060)	(0.049)	(0.068)	(0.056)
Treatment x CE	0.131**	0.130***	0.199***	0.123**
	(0.052)	(0.047)	(0.059)	(0.056)
Treatment x weak IA	0.268***	0.211***	0.359***	0.328***
	(0.060)	(0.051)	(0.067)	(0.057)
Treatment x NF	-0.046	0.085	-0.086	0.041
	(0.093)	(0.068)	(0.109)	(0.080)
Observations	1,850	2,775	1,322	1,982

Notes: Estimates come from linear regressions. Strict inequality aversion means that spectators equalised ultimate payoffs while weak inequality aversion means spectators gave more to the decision maker who had a lower initial payoff. Consistency relates to the number of decisions spectators made in line with the particular criterion. NF refers to spectators who cannot be classified by either of the fairness criteria in the control condition. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

Table B4: Treatment effects with complete demographic controls

	Full Results		Above $1	
	Ex Ante	Ex Post	Ex Ante	Ex Post
Treatment	0.127**	0.136***	0.196***	0196***
	(0.050)	(0.044)	(0.057)	(0.050)
Demographics				
Age	-0.002	-0.003*	-0.001	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)
Gender	0.025	-0.052	-0.036	-0.060
	(0.055)	(0.045)	(0.064)	(0.051)
Education	-0.034*	0.005	-0.044**	0.012
	(0.018)	(0.014)	(0.020)	(0.017)
Income	0.007	0.012*	0.014	0.012
	(0.008)	(0.007)	(0.010)	(0.008)
Employment	-0.045	-0.140***	-0.029	-0.193***
	(0.062)	(0.051)	(0.071)	(0.058)
Economics	0.094*	0.019	0.093	-0.007
	(0.056)	(0.046)	(0.064)	(0.052)
Ethnicity				
American Indian or Alaska Native	0.210	0.384**	0.156	0.427**
	(0.268)	(0.178)	(0.313)	(0.186)
Asian	0.079	0.119	0.066	0.207**
	(0.109)	(0.087)	(0.126)	(0.097)
Black or African-American	0.042	-0.092	0.044	-0.102
	(0.110)	(0.088)	(0.131)	(0.099)
Native Hawaiian or Pacific Islander	-	-	-	-
	-	-	-	-
Spanish, Hispanic or Latino	-0.111	-0.014	-0.050	0.038
	(0.085)	(0.069)	(0.095)	(0.076)
White	0.070	0.088	0.078	0.100
	(0.090)	(0.073)	(0.101)	(0.081)
Preferences				
Risk seeking	0.031**	0.058***	0.019	0.039***
	(0.012)	(0.010)	(0.014)	(0.012)
Ambiguity aversion	-0.008	-0.012***	-0.015***	-0.017***
	(0.005)	(0.004)	(0.006)	(0.005)
Left-Right	-0.015	0.008	-0.006	0.004
	(0.015)	(0.013)	(0.017)	(0.014)
Party affiliation				
Democrats	0.491***	0.677***	0.342	0.583***
	(0.157)	(0.124)	(0.210)	(0.160)
Republicans	0.302*	0.380***	0.160	0.379**
	(0.165)	(0.128)	(0.214)	(0.161)
Other	0.292*	0.374***	0.091	0.321**
	(0.159)	(0.125)	(0.211)	(0.159)
2020 Vote				
Joe Biden	-0.396***	-0.293***	-0.247*	-0.194*
	(0.117)	(0.093)	(0.141)	(0.105)
Donald Trump	-0.165	-0.115	0.050	-0.040
	(0.134)	(0.105)	(0.159)	(0.113)
Other	-0.080	0.393**	-0.003	0.305*
	(0.209)	(0.157)	(0.246)	(0.183)
Didn't vote	-0.259**	-0.184*	-0.165	-0.168
	(0.120)	(0.096)	(0.147)	(0.107)
Constant	2.014***	1.557***	2.154***	1.782***
	(0.237)	(0.197)	(0.286)	(0.232)
Observations	1,680	2,520	1,201	1,807

Notes: Estimates come from linear regressions. Detailed descriptions of each variable can be found in appendix section C. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

are robust to the inclusion of this battery of control variables. The coefficients are also still remarkably similar to the effect observed in the simple treatment comparison in figure 3 in the main text. None of the control variables have a similarly robust and significant effect on the share allocated to the risk taker as the treatment assignment. Being employed decreases the share to a similar extent as the treatment assignment in the ex post condition. A higher degree of risk seeking preferences is also somewhat robustly associated with a higher reward for the risk taker while higher ambiguity aversion is correlated with a lower reward for the risk taker. These two findings suggest that spectators judge decision makers based on their own preferences which is not surprising. Interestingly, in the ex post condition, any type of party affiliation is associated with a higher reward for the risk taker. On the other hand, a vote for Joe Biden in 2020 is associated with a lower reward.

B.2 Pre-registered interaction effects

As outlined in the pre-analysis plan, table B5 reports estimates of the increasing treatment effect reported in table 1 in the main text, interacted with outcomes over the personal lottery and externalities. The coefficients can be read as the increase in the amount allocated to the risk-taker for a given outcome and treatment group, if the safe option increases by 50ct.

Table B5: Increasing treatment effect interactions with outcomes

	Amount allocated to Risk-Taker			
	Externality		Lottery	
	✓	x	✓	x
Control x Level x			-0.025	0.195***
			(0.023)	(0.029)
Ex Ante x Level x			0.050**	0.268***
			(0.023)	(0.028)
Ex Post x Level x	0.181***	0.127***	0.031	0.279***
	(0.022)	(0.022)	(0.020)	(0.023)
Observations	3,700	3,700	3,700	3,700

Notes: Estimates come from linear regressions. Coefficients report the effect of a particular treatment assignment interacted with the level of the safe option and interacted with a particular outcome. Robust standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

The reported results indicate, similarly to table 2, that spectators show outcome bias over the externality in the ex post treatment group. The coefficient for the scenarios in which the externality realised is somewhat larger than if the externality did not realise. Both are highly significant. This suggests that spectators increase the reward they allocate to risk-takers more as the value of the safe option rises, if the externality realised.

Figure B1: Full Spectator Beliefs about why decision makers took risk

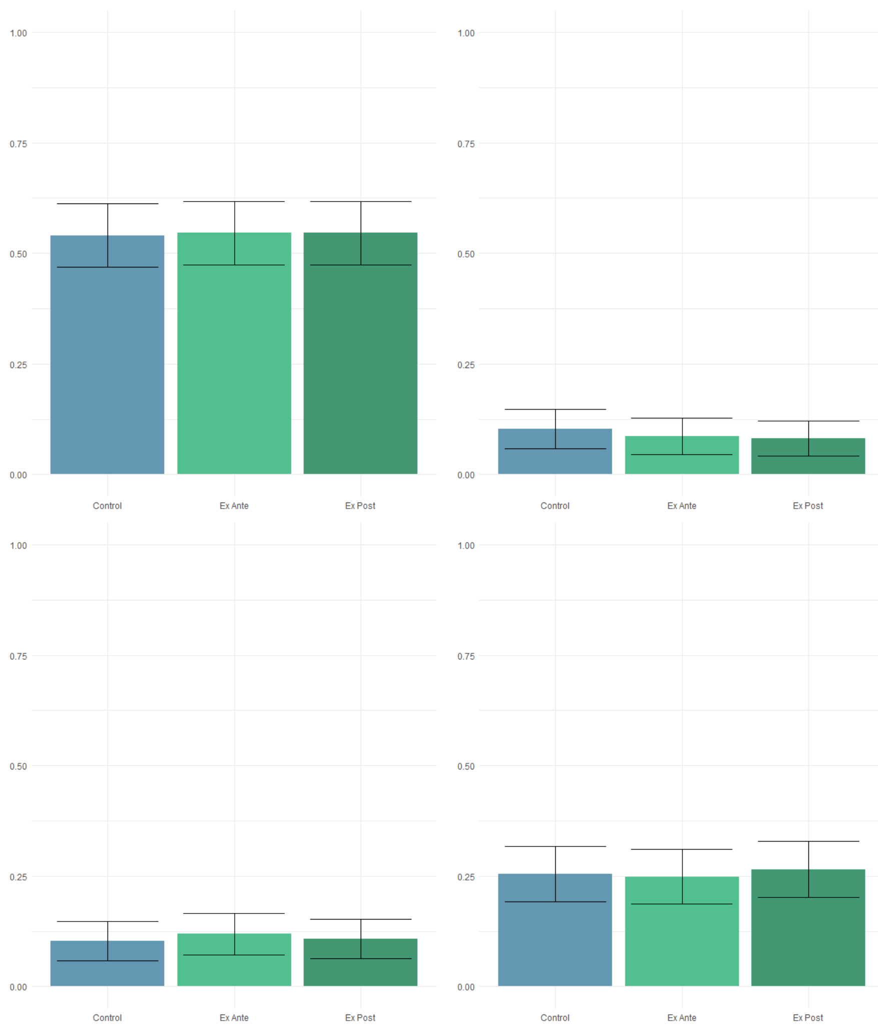
Notes: Spectators were asked why they think some participants chose the risky option B. From top left to bottom right, the panels report the share of spectators by treatment conditions who selected the option "To maximise their own payoff", "To maximize the pair's payoff", "To influence your allocation decision of the additional \$4", "To generate the additional \$2 for the other participants", "Other". The only significant difference between treatments can be observed for the first statement and the fourth, which was however only an option in the treatment conditions.

Turning to the results of the interaction effects with outcomes over the risk-takers personal lottery, a similar pattern across all treatment groups can be observed. While spectators increase the reward of the risk-taking decision maker as the value of the safe option rises if the personal lottery was unsuccessful, there is no significant increase if the personal lottery was successful. Importantly, this increase in the reward of the risk-taking decision maker is somewhat larger in the treatment groups than in the control group, if the personal lottery was unsuccessful. While, on average, spectators do not reward risk-takers more if the personal lottery was unsuccessful in the treatment compared to the control group, as reported in table 2, they seem to nonetheless respond more to the opportunity cost of choosing risk in the treatment group.

B.3 Full Results of Belief elicitation

Figure B1 illustrates spectators' responses to the question "Why do you think some participants chose the risky option B?". Each panel reports the share of spectators by treatment condition who selected a particular option as the most important in answering that question.

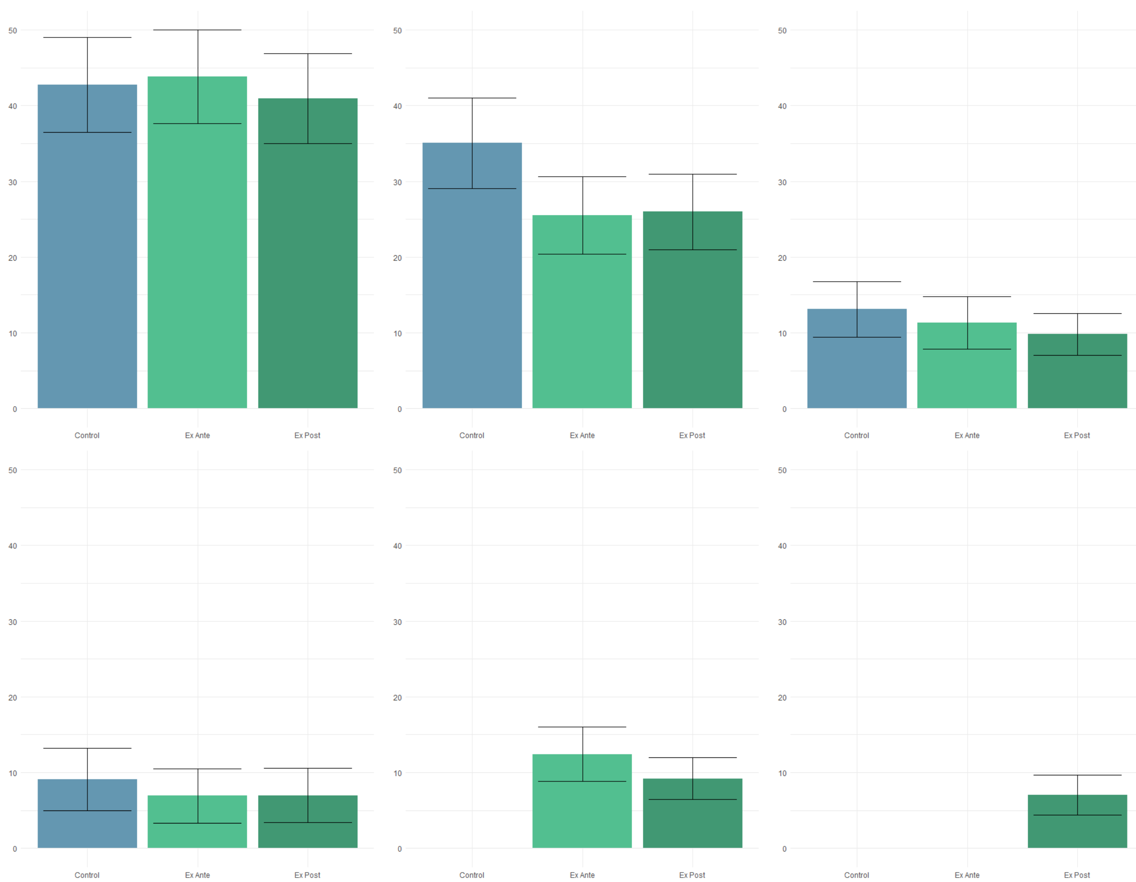
Figure B2: Full Spectator Beliefs about why decision makers took safe option



Notes: Spectators were asked why they think some participants chose the safe option A. From top left to bottom right, the panels report the share of spectators by treatment conditions who selected the option "To maximise their own payoff", "To maximize the pair's payoff", "To influence your allocation decision of the additional $4", "Other". There are no significant differences between treatments for any of the statements.

Spectators were asked this question after each treatment or control block which allows for this comparison of answers across conditions. As already reported in the main text, the only significant treatment difference can be observed for the option "To maximize their own payoff". Spectators are less likely to select this option as the most important one in answering the question in either treatment condition compared to the control conditions. This difference can be explained by the option "To generate the additional $2 for the other participants", which is only an option in the treatment conditions but selected by a small share of spectators as the most important option in explaining decision makers' risk-taking.

Figure B3: Full Spectator Beliefs about allocation decisions



Notes: Spectators were asked which attributes they considered when making their allocation decisions. From top left to bottom right, the panels report the share of spectators by treatment conditions who selected the option "Inequality in earnings between participants", "The choices participants made", "The outcome of the personal lottery", "Other", "The potential benefit for the participant pool of choosing option B", "Whether the additional $2 generated for the participant pool". The fifth option was added for both treatment conditions and the sixth only for the ex post treatment condition. There are no significant differences between treatments for any of the statements.

After each treatment or control block of allocation decisions, spectators were also asked why they believe some decision makers took the safe option A. Figure B2 reports in each panel the share of spectators by treatment condition who selected a particular option as the most important in answering that question. Compared to figure B1, it is striking how consistent spectator responses are across treatment conditions. There are no significant differences but the means are also almost identical across the conditions. This result suggests that there are no spillovers in beliefs about the decision maker who did not take the risk when externalities become part of the decision problem.

Spectators were also asked after each treatment or control condition block, which attributes they considered when making their allocation decisions. Figure B3 reports the share of spectators who selected a particular attribute as most important by treatment condition.

The most selected option across all treatment conditions is "Inequality in earnings between participants". Between 40-45% of spectators select this option as the most important in each condition. The second most chosen option is "The choices participants made". Here, although not statistically significant, there is a substantial drop of almost 10% in the share of spectators who select this option as most important between the control and treatment conditions. This difference can be explained by the small share of spectators who opt for the options "The potential benefit for the participant pool of choosing option B" or "Whether the additional $2 generated for the participant pool" in the treatment conditions.

These beliefs are interestingly in line with the results reported in table B3 as spectators who are choice egalitarians in revealed and stated preferences seemingly trade-off their fairness criterion with a consideration for rewarding prosocial risk-taking in the treatment conditions.

C Description of Variables

Share b. Continuous variable indicating the absolute share of the $4 bonus spectators allocated to the decision maker who chose the risky option B.

Age. Continuous variable indicating the self-reported age of subject i.

Gender. Binary variable coded as 1 if subject i indicated to be male, 0 if subject i indicated to be female. Subjects who indicated "other" or "prefer not to answer" were coded as missing values.

Education. Categorical variable capturing subject i's highest level of education.
0: Primary education or less

1: Some high school
2: High school degree/GED
3: Some college
4: 2-year college degree
5: 4-year college degree
6: Master's degree
7: Professional degree (JD, MD, MBA)
8: Doctoral degree

Income. Categorical variable capturing the income bracket of subject i. Values are stated in US Dollar ($).
1: Under $10,000
2: $10,000 to $20,000
3: $20,001 to $30,000
4: $30,001 to $40,000
5: $40,001 to $50,000
6: $50,001 to $60,000
7: $60,001 to $80,000
8: $80,001 to $100,000
9: $100,001 to $150,000
10: $150,001 to $200,000
11: $200,001 to $350,000
12: $350,001 to $500,000
13: Above $500,000

Employment. Binary variable coded as 1 if subject i indicated that they are in either full or part time employment or a business owner and 0 if subject i indicated to be unemployed, retired or a student.

Economics. Binary variable coded as 1 if subject i indicated that they have taken a module in economics or a related subject at University. A value of 0 indicates that subject i either has not taken a module in economics or has never attended higher education.

Left-Right Placement. Continuous variable capturing subject i's self-reported placement on a political left-right scale with 0 indicating "Left" and 10 indicating "Right".

Party affiliation. Categorical variable capturing subject i's self-reported political party affiliation. Those who responded with "Don't know" were coded as missing variables.
1: Democratic Party
2: Republican Party
3: Other

2020 Vote. Categorical variable capturing subject i's self-reported vote in the 2020 Presidential election. Those who responded with "Don't remember" or "Prefer not to say" were coded as missing variables.
1: Joe Biden
2: Donald Trump
3: Other candidate
4: Didn't vote

Risk seeking. Variable capturing subject i's willingness to take risks on a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means "very willing to take risks".

Ambiguity aversion. Variable capturing subject i's degree of ambiguity aversion on a scale from 4 to 28. This variable combines subjects' responses to the four questions of item D12 which can be found in appendix D and is taken from the module on measuring ambiguity aversion in (Cavatorta and Schröder 2019). The possible responses to each question range from 1 indicating "strongly agree" to 7 indicating "strongly disagree". A higher value on this scale therefore indicates a lower degree of ambiguity aversion.

D Experimental Instrument

D.1 Stage I: Decision Makers

Decisions 1-7 are presented in randomized order during the experiment and vary in the amount of the safe option (from $0.5 in Decision 1 to $3.5 in Decision 7). Below, only decision 3 (with a safe option of $1.50) is shown as an example. Text which is only shown in the treatment condition is highlighted.

Introduction:

Thank you for participating in this study. In the following, you will be asked to make a number of decisions that will influence the bonus payment you can receive for this study. Specifically, you will be asked seven times to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. If you choose option B and irrespective of the outcome of the lottery, there is also a 50% chance that an additional $2 will generate. If this is the case, we will use these $2 to reward two randomly chosen participants from other studies. Below is an example decision you might face:

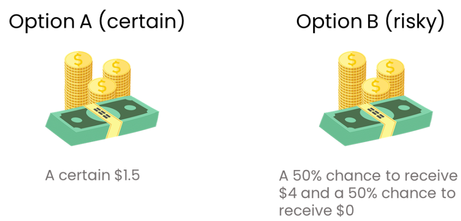


Figure B1a: Control Screen

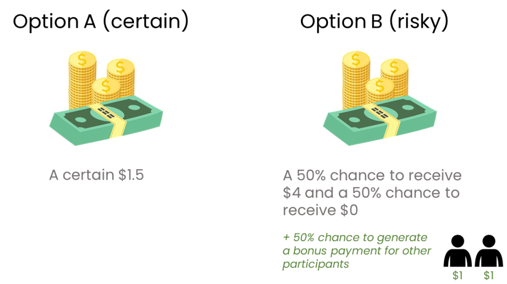


Figure B1b: Treatment Screen

While option B will remain the same throughout the seven decisions, the value of option A, which is a certain payment, will vary.

After you made your decisions, you will be paired with another participant and **one of the decisions** you made will be randomly selected to determine your bonus payment. All of the seven decisions you make have an equal chance of being selected for payment.

You will then be asked a set of questions about yourself and about the choices you just made.
Before you receive your bonus payment, a third participant will have the option to allocate an **additional $4** between you and the other participant as they wish. Depending on the choices you make and the decision of the third participant, you can therefore receive **a total of $8 in bonus payments** in this study. To receive a bonus payment, you need to complete the full study. This should take about 10 minutes.

Understanding:
Before you make your first decision, please answer the following questions. Your final payment will not depend on your answers to these questions. However, please answer to the

best of your ability as your answers will impact the quality of our research.[7]

U1: How many decisions are you asked to make?

- 5
- 7
- 10
- Don't know

U2: Which one of the two options remains the same throughout the decisions?

- The certain option A
- The risky option B
- Don't know

U3: What will determine your final bonus payment? Please select all that apply.

- The sum of all the choices you make
- One randomly selected choice
- The allocation decision of $4 by a third participant between you and the other participant in your pair
- The sum of all the choices made by the other participant in your pair
- Don't know

U4: If you choose the risky option B, how high is the chance that $2 will generate that we will use to pay additional participants?

- 0%
- 25%
- 50%
- 75%

[7] After subjects have submitted their answers to U1-U4, the correct answers will be displayed before they can proceed to the next page.

- 100%
- Don't know

Example Decision:
Please carefully consider the below two options. The option you choose has a chance of 1 in 7 to determine your bonus payment for this study. Once you leave this screen, you cannot change the decision you made. Please remember that if you choose option B and irrespective of the outcome of the lottery, there is also a 50% chance that an additional $2 will generate that will be used to reward two randomly chosen participants from other studies. Which option would you like to choose?

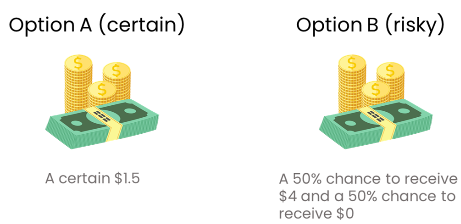


Figure B2a: Control Screen

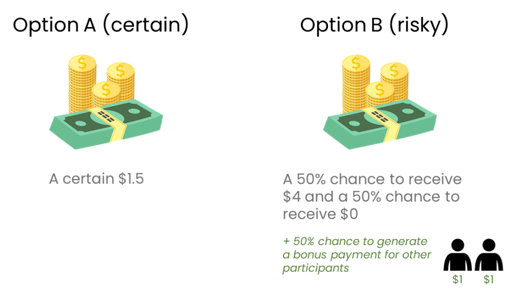


Figure B2b: Treatment Screen

Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Reasoning Questions:

J1: Please explain your reasoning for the decisions you just made:

J2: Did you choose option B in any of the decisions:

- Yes
- No

J3: *[if previous question Yes]* Why did you choose option B in those cases? Please select all that apply.

- To maximise my own payoff
- To generate the additional $2 for the other participants
- To influence the allocation decision of the $4 by the third participant
- Other

J4: Which of the reasons you selected for choosing option B was most important for your decision making?

- To maximise my own payoff
- To generate the additional $2 for the other participants
- To influence the allocation decision of the $4 by the third participant
- Other
- All were of equal importance

Belief Questions:

B1: Please assume for now that the below decision will get selected for payment and the other participant in your pair did not choose the same option as you. Remind yourself of the choice you made. How much of the additional $4 do you expect the third participant would allocate to you in this case if they were fully informed about your and the other participant's choices?

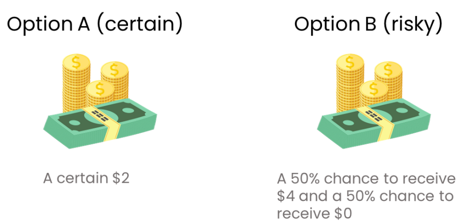


Figure B3a: Control Screen

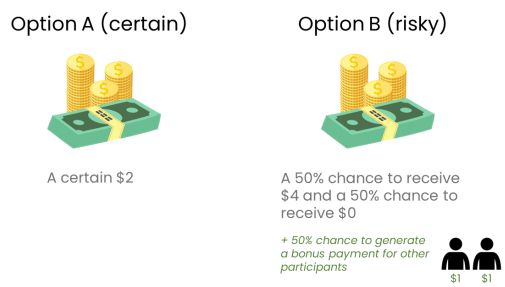


Figure B2b: Treatment Screen

Confidence: On a scale from 0 to 10, how confident are you in the estimate you just provided?

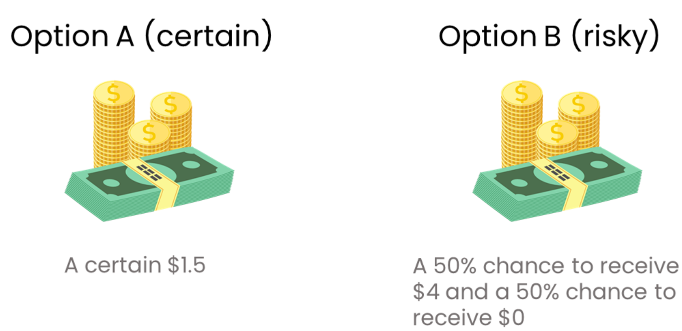
B2: On a scale from 0 to 10, how much autonomy do you feel you have over your final earnings?

D.2 Stage II: Spectators

Decisions 1-20 are presented in randomized order during the experiment. Below, one decision from each treatment and control condition is shown as an example.

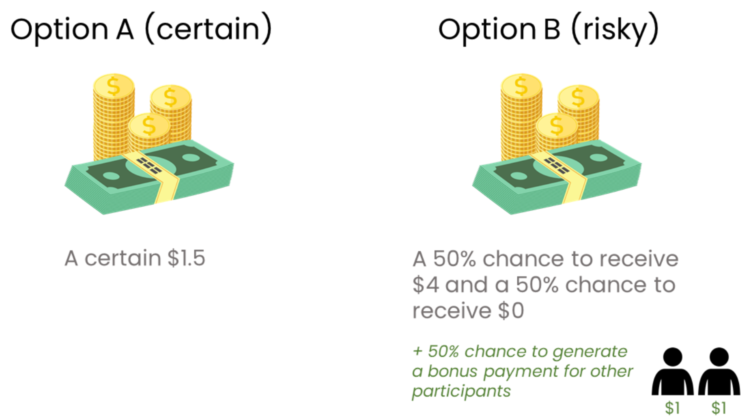
Introduction:
Thank you for participating in this study. In the following, you will be asked to decide on an allocation of money between two participants. You will be asked to make a total of 20 allocation decisions. While 19 of these decisions are hypothetical, 1 will determine the actual payment for a participant pair. Please consider each decision carefully as we will not inform you about which of the 19 decisions is the one that determines actual payoffs for the two participants. After making the 20 allocation decisions you will be asked a set of questions about yourself and about the choices you just made. You will also have the opportunity to earn an additional bonus payment.

Control Condition:
You will now be asked to make your first/next five allocation decisions. All of the participants within the pairs already had the chance to receive an income based on a choice they were asked to make. Specifically, they were asked to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. Below is an example of a decision they might have faced:



While option B always remains the same across all potential scenarios you will face, the value of option A, which is a certain payment, may vary. You will be informed about **the decision the participants faced, the choice they made, and their resulting current earnings**. You then have the option to allocate an additional $4 between the pair.

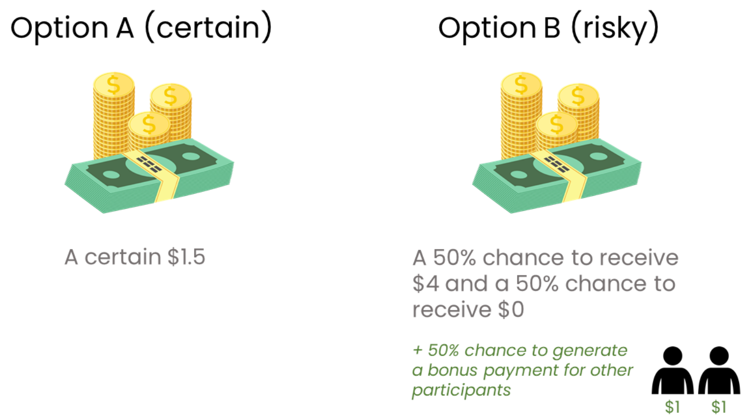
Ex ante Treatment Condition:
You will now be asked to make your first/next five allocation decisions. All of the participants within the pairs already had the chance to receive an income based on a choice they were asked to make. Specifically, they were asked to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. If a participant chose option B and irrespective of the outcome of the lottery, there was also a 50% chance that an additional $2 generated. In these cases, the $2 have been used to reward two randomly chosen participants from other studies. Participants were aware of this possibility when making their decisions. Below is an example of a decision they might have faced:



While option B always remains the same across all potential scenarios you will face, the value of option A, which is a certain payment, may vary. You will be informed about **the decision the participants faced, the choice they made, and their resulting current earnings** but **not** about whether the **additional $2** generated. You then have the option to allocate an additional $4 between the pair.

Ex post Treatment Condition:
You will now be asked to make your first/next ten allocation decisions. All of the participants within the pairs already had the chance to receive an income based on a choice they were asked to make. Specifically, they were asked to decide between two options. Option A is always a certain payment while option B is a lottery with a potentially higher payoff. If a participant chose option B and irrespective of the outcome of the lottery, there was also a 50% chance that an additional $2 generated. In these cases, the $2 have been used to reward two randomly chosen participants from other studies. Participants were aware of

this possibility when making their decisions. Below is an example of a decision they might have faced:



While option B always remains the same across all potential scenarios you will face, the value of option A, which is a certain payment, may vary. You will be informed about **the decision the decision the participants faced, the choice they made, their resulting current earnings, and whether the additional $2 generated**. You then have the option to allocate an additional $4 between the pair.

Understanding:
Before you make your first decision, please answer the following questions. Your final payment will not depend on your answers to these questions. However, please answer to the best of your ability as your answers will impact the quality of our research.[8]

U1: How many decisions are you asked to make in total?

- 10
- 20
- 30
- Don't know

U2: Which one of the two options remains the same throughout the decisions?

[8] After subjects have submitted their answers to U1-U4, the correct answers will be displayed before they can proceed to the next page.

- The certain option A
- The risky option B
- Don't know

U3: How many of your decisions will result in an actual payment for the two participants?

- 0
- 1
- 5
- 10
- Don't know

U4: If a participant chose the risky option B, how high is the chance that $2 will generate that we will use to pay additional participants?[9]

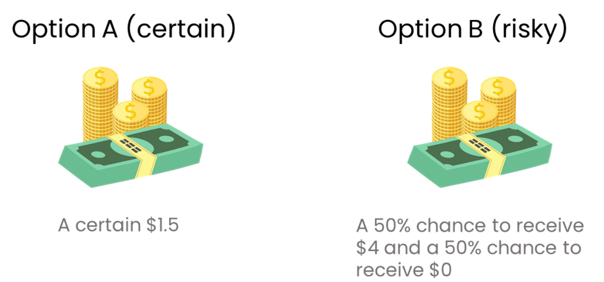
- 0%
- 25%
- 50%
- 75%
- 100%
- Don't know

Reminder:
You will now make five/ten allocation decisions. Please remember that one of your decisions will determine the real payment of two individuals who participated in this study.

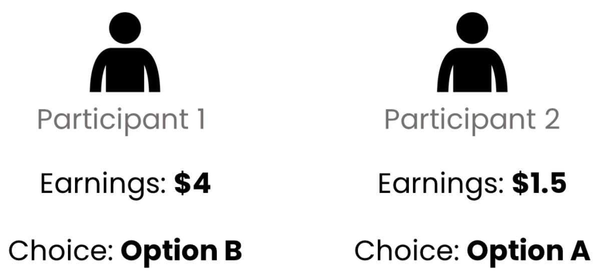
[9] This question is displayed together with U1-U3 if spectators are randomly allocated to see one of the treatment conditions first. If spectators see the control conditions first, this question is displayed on its own after the introduction of the second condition.

Control Decision Example:

Please carefully consider the below scenario. Participant 1 and 2 both faced the following decision:



The outcomes for both participants and the choices they made are given below:



You are now asked to allocate an additional amount of money between the pair. You can allocate this money as you wish. Please note, you have to allocate the total amount of $4.

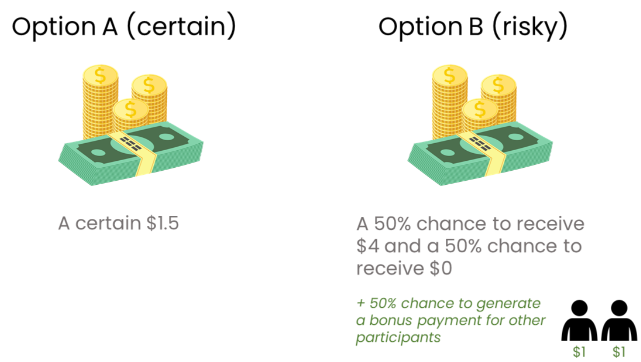
Please specify the amount you would like to allocate to each participant. You may use up to two decimal points when specifying each amount. Please ensure the two values add up to $4 before proceeding.

Participant 1 (in $):
Participant 2 (in $):

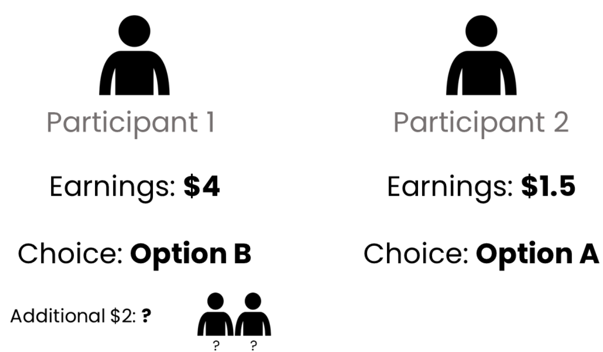
Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Ex Ante Treatment Decision Example:

Please carefully consider the below scenario. Participant 1 and 2 both faced the following decision:



The outcomes for both participants and the choices they made are given below. Whether the additional $2 generated or not will be revealed after your allocation decision:



You are now asked to allocate an additional amount of money between the pair. You can allocate this money as you wish. Please note, you have to allocate the total amount of $4.

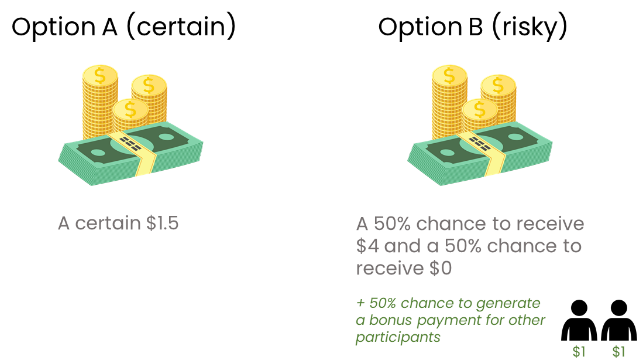
Please specify the amount you would like to allocate to each participant. You may use up to two decimal points when specifying each amount. Please ensure the two values add up to $4 before proceeding.

Participant 1 (in $):
Participant 2 (in $):

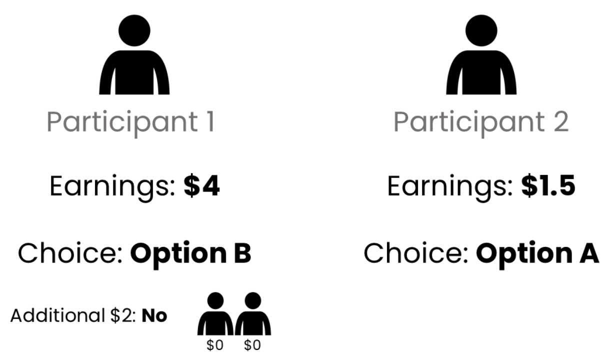
Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Ex Post Treatment Decision Example:

Please carefully consider the below scenario. Participant 1 and 2 both faced the following decision:



The outcomes for both participants and the choices they made are given below. Because participant 2 chose the risky option B, an additional $2 generated and two randomly chosen participants from other studies **have been given** a $1 bonus payment ($2 in total) / Despite participant 2 having chosen option B, the additional $2 did not generate and the two randomly chosen participants from other studies **have not been given** a $1 bonus payment ($2 in total):



You are now asked to allocate an additional amount of money between the pair. You can allocate this money as you wish. Please note, you have to allocate the total amount of $4.

Please specify the amount you would like to allocate to each participant. You may use up to two decimal points when specifying each amount. Please ensure the two values add up to $4 before proceeding.

Participant 1 (in $):

Participant 2 (in $):

Confidence: On a scale from 0 to 10, please indicate how confident you are in the decision you just made.

Beliefs and Preferences[10]

B1: How did you decide on the allocation of income within the participant pairs?

B2a: Which of the following attributes did you consider when making your allocation decisions? Please select all that apply.

- Inequality in earnings between participants
- The choices participants made
- The outcome of the personal lottery in option B
- Whether the additional $2 generated for the participant pool[11]
- The potential benefit for the participant pool of choosing option B
- Other

B2b: How important were each of the attributes you just selected for your allocation decisions? Please allocate a total of 100 points across the attributes you selected. Please ensure that the more important an attribute was to your decision making, the more points you allocate to it.

B3a: Why do you think some participants chose the risky option B? Please select all that apply.

- To maximise their own payoff
- To maximise the pair's payoff
- To generate the additional $2 for the other participants
- To influence your allocation decision of the additional $4

[10]B1-B4 are asked after each condition. Highlighted options are only displayed in the two treatment conditions. B5-B9 are asked after all three conditions were completed.

[11]This option is only included in the ex ante treatment condition.

- Other

B3b: Which of the reasons you just selected do you think was the main reason why some participants chose the risky option B?

B3conf: On a scale from 0 to 10, please indicate how confident you are that this was the main reason why some participants chose the risky option B.

B4a: Why do you think some participants chose the certain option A? Please select all that apply.

- To maximise their own payoff
- To maximise the pair's payoff
- To influence your allocation decision of the additional $4
- Other

B4b: Which of the reasons you just selected do you think was the main reason why some participants chose the certain option A?

B4conf: On a scale from 0 to 10, please indicate how confident you are that this was the main reason why some participants chose the certain option A.

B5: What percentage of decision makers do you believe chose the risky option B in the below scenarios? You will receive a bonus payment of 10ct for each estimate that is within +/-5 percentage points of the correct answer. If all of your estimates are correct, you will therefore be able to earn an additional bonus payment of $1.[12]

I think the percentage of decision makers who chose the risky option B in the scenario where the certain option A was [**x**] and it was **not** possible to generate the additional $2 is:

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

I think the percentage of decision makers who chose the risky option B in the scenario where the certain option A was [**x**] and it **was** possible to generate the additional $2 is:

[12] Each spectator is asked to answer B5 for five randomly selected values of A in the control and treatment condition.

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

B6a: Which of the following do you believe apply? Please choose all options that you agree with. The $2 for the two other randomly chosen participants:

- Will decrease income inequality between all participants on Prolific Academic
- Will increase income inequality between all participants on Prolific Academic
- Will benefit the two participants who receive it
- Will not matter to the two participants who receive it
- Is unfair because they did nothing to receive it
- Will increase the total amount of money all participants on Prolific Academic have combined

B6b: Which of the statements you just selected do you agree with most? The $2 for the two other randomly chosen participants:

B7: On a scale from 0 (lower) to 10 (higher), what do you believe is the effect of the $2 the two randomly chosen participants might receive, if a decision maker chose option B, on income inequality between all participants on Prolific Academic?

B8: Do you believe risk-taking should be rewarded?

- Yes
- No
- I don't know
- I don't have an opinion on this

B9: Do you believe risk-taking for the benefit of society should be rewarded?

- Yes
- No
- I don't know
- I don't have an opinion on this

D.3 Demographics

In this final part of the study, we will ask you a few of questions about yourself. Please read the questions carefully and answer honesty. This part should take only 2-3 minutes.

D1: How old are you?

D2: What is your gender?

- Female
- Male
- Other
- Prefer not to answer

D3: To which of these groups do you consider you belong? You can choose more than one group.

- American Indian or Alaska Native
- Asian
- Black or African-American
- Native Hawaiian or other Pacific Islander
- Spanish, Hispanic or Latino
- White
- Other group
- Prefer not to answer

D4: Which category best describes your highest level of education?

- Primary education or less
- Some high school
- High school degree/GED
- Some college
- 2-year college degree

- 4-year college degree
- Master's degree
- Doctoral degree
- Professional degree (JD, MD, MBA)
- Prefer not to answer

D5: What is your total (annual) household income before tax?

- Under $10,000
- $10,000 - $20,000
- $20,001 - $30,000
- $30,001 - $40,000
- $40,001 - $50,000
- $50,001 - $60,000
- $60,001 - $80,000
- $80,001 - $100,000
- $100,001 - $150,000
- $150,001 - $200,000
- $200,001 - $350,000
- $350,001 - $500,000
- Above $500,000
- Don't know
- Prefer not to answer

D6: What is your current employment status?

- Full-time employee
- Part-time employee

- Self-employed or small business owner
- Medium or large business owner
- Unemployed and looking for work
- Student
- Not currently working and not looking for work (e.g. full-time parent)
- Retiree
- Prefer not to answer

D7: Have you ever taken a module on economics or a related subject area at university?

- Yes
- No
- I have never attended higher education

D8: In politics people sometimes talk of left and right. Where would you place yourself on the following scale?

D9: Which party do you feel closest to?

- Democratic Party
- Republican Party
- Other
- Don't know

D10: Who did you vote for in the recent 2020 Presidential Election?

- Joe Biden
- Donald Trump
- Other candidate
- Didn't vote
- Don't remember

- Prefer not to say

D11: Please tell us, in general, how willing or unwilling you are to take risks. Please use a scale from 0 to 10, where 0 means "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale.

D12: Please respond to the following statements by indicating the extent to which you agree or disagree with them on a scale from 1 (I strongly agree) to 7 (I strongly disagree).

- There is a right way and a wrong way to do almost everything
- Practically every problem has a solution
- I feel relieved when an ambiguous situation suddenly becomes clear
- I find it hard to make a choice when the outcome is uncertain

D13: Do you have any feedback or impressions regarding this study?