Climate Change Risk, and Human Behavior: Theory and Evidence

WORKING PAPER 2306

SEPTEMBER 2023

Sanjit Dhami
Narges Hajimoladarvish
Pavan Mamidi

Climate Change Risk, and Human Behavior: Theory and Evidence

Sanjit Dhami[*] **Narges Hajimoladarvish** [†] **Pavan Mamidi**[‡]

September 21, 2023

Abstract

A group of decision makers simultaneously make contributions towards a green fund that reduces the future probability of a climate catastrophe. We derive the theoretical predictions of the effects on contributions arising from 'behavioral parameters' such as loss aversion and present-bias; 'structural factors' such as variation in the timing of uncertainty; the 'demand for a commitment device'; and 'institutional factors' such as comparing voluntary contributions with mandatory tax financed contributions. We then run experiments to stringently, test our predictions. Loss aversion and present-bias reduce contributions; there is demand for the commitment technology; and voluntary contributions are higher relative to mandatory tax-financed contributions.

Keywords: Climate risk abatement; loss aversion; present-biased preferences; voluntary versus mandatory contribution mechanisms; commitment technology.

JEL Classification: C92 (Laboratory, Group Behavior); D01 (Microeconomic Behavior: Underlying Principles); D02 (Institutions: Design, Formation, Operations, and Impact); D91 (Role and Effects of Psychological, Emotional, Social, and Cognitive Factors on Decision Making).

[*]Corresponding author. Professor of Economics, Department of Economics, Finance and Accounting, School of Business, Room 0.20, Teaching Centre, Brookfield, 266 London Road, University of Leicester, Leicester LE2 1RQ, UK. Phone: +44-116-2522086. Fax: +44-116-2522908. E-mail: sd106@le.ac.uk.

[†]Senior Research Fellow, Centre for Social and Behaviour Change, Ashoka University, Delhi Office No. 222, Second Floor, Okhla Industrial Estate - Phase III, New Delhi, 110020, India. E-mail: narges.hajimoladarvish@ashoka.edu.in.

[‡]Director, Centre for Social and Behaviour Change, Ashoka University, Delhi Office No. 222, Second Floor, Okhla Industrial Estate - Phase III, New Delhi, 110020, India. E-mail: pavan.mamidi@ashoka.edu.in.

1 Introduction

Climate change, and the human response to climate change, pose some of the most challenging and important questions for the social, behavioral, and natural sciences (Stern, 2007, 2008, 2022; IPCC, 2014). We leverage some of the key insights from the core of behavioral economic theory to analyze the human response to climate change, beyond the classical nudge-type interventions.[1] We construct a theoretical model of behavioral choice that makes precise, testable, predictions and then test the predictions stringently using controlled lab experiments. Lab experiments offer the advantage of creating precise environments, where individual-specific behavioral parameters can be satisfactorily estimated, and allowing for stringent tests of the underlying transmission channels in the theoretical model.

Our theoretical model, and the experimental design, which is exclusively informed by our theory, captures the following essential features that should arguably comprise any minimally informed account of the problem.

1. *Temporal dimension*: Investments to mitigate the effects of climate change (green investments) are costly in terms of the current resources foregone, and the benefits materialize in the future.

2. *Risk and uncertainty*: Green investments undertaken now for the abatement of future climate risk, typically map into a distribution of risky future outcomes. Higher green investments reduce the probability of future undesirable environmental outcomes.[2]

3. *Public goods*: Green investment has the nature of a public good, and contributions have the typical characteristics of social dilemmas. The privately optimal decision might be to free-ride on the costly contributions of others that create public environmental benefits in the future, yet the socially optimal decision might be to make positive, possibly high, contributions.[3]

4. *Institutions*: Voluntary contributions to green investment, and mandatory tax financed contributions (formal institutional mechanism) earmarked for green investment, may lead to different outcomes. There are tradeoffs in overcoming potential problems of low expectations of contributions under voluntary contributions and a reduction in human agency and autonomy under mandatory contributions. Ultimately, this is an empirical question.[4]

[1]There is a sizeable and valuable literature on the effects of nudge-type interventions on energy consumption. These interventions include providing information on social comparisons of energy usage (Allcott, 2011b); effects of observation by the experimenter of energy usage (Schwartz et al., 2013); and providing information on energy usage (Allcott, 2011a; Jessoe and Rapson, 2014). There is also an experimental literature on priming individuals for other-regarding or moral/empathetic preferences and studying their behavioral responses in terms of pro-environment behavior; for surveys, see Schultz and Zelezny (1999), Dietz et al., (2005), and Heinz and Koessler (2021).

[2]Stern (2008) identifies 5 channels through which climate change is caused by greenhouse gases, or GHGs (see also IPCC, 2014). In each case, a greater stock of GHGs creates a higher probability of climate change and in at least three of these channels (the absorption-stock accumulation, climate-sensitivity, and warming-climate change channels) the effects are experienced with a time delay. Thus, any investments undertaking now to reduce the stock of GHGs are likely to reduce the probability of the adverse temporal consequences of climate change.

[3]We are, thus, particularly interested in what Miliniski et al. (2008) refer to as a *collective risk social dilemma*. There is a conflict between individual and social interests, and current actions lead to delayed and risky outcomes (Nordhaus and Boyer, 2000; Stern, 2007, 2008; Barrett and Dannenberg, 2012).

[4]Indeed, a central theme of the work by Ostrom (1990) was to show that in many societies, humans can cooperate

We are interested in the behavior of individual decision makers, e.g., consumers, households, firms, regions, or even countries, in groups. We abstract from several issues or types of analyses that include the following. (i) Analyses based on classical price incentives and regulation, as there is already a rich literature on that subject, (ii) issues arising from complex system dynamics under true uncertainty and bounded rationality, particularly in terms of technology choice by firms (Stern, 2022; Dhami, 2023), (iii) issues of long run climate change through an explicit consideration of the stock rather than the flow of GHGs (Stern, 2008; Dannenberg et al., 2015), (iv) we also do not consider a repeated game (although we consider a multi-stage game)[5], and (v) we do not consider threshold public goods games (although we do consider a public goods game).[6]

1.1 The framework

We have three distinct time periods, $t_1 \equiv 0 < t_2 < t_3$; n distinct decision makers who have identical income endowments, but potentially different behavioral preference parameters; and 4 treatments T1–T4. In all treatments, individuals simultaneously choose contributions (or green contributions) towards a green fund. The sum of all individual contributions comprises the 'green fund' and it reduces the probability of future climatic disasters. All decisions are made in the current period, at time $t_1 = 0$. The time t_3 events are identical in all treatments. Treatments T1 and T2 require voluntary green contributions towards the green fund. However, in treatments T3 and T4, green contributions are financed through mandatory income taxes that are earmarked for the green fund; the median tax rate across the most preferred tax rates chosen by the decision makers is implemented. Time t_2 only plays a role in treatments T2 and T4 in order to create a commitment device, but plays no role in in treatments T1 and T3, as we now explain.

In our baseline treatment, T1, the endowments for the three time periods t_1, t_2, t_3 are, respectively, $Y, 0, Z$, where $Y > 0, Z > 0$. At time $t_1 = 0$, all n decision makers simultaneously allocate their current endowment of Y towards current consumption and green investments for the future. The sum of green contributions across all n decision makers determines the stock of green fund, G. At time t_3, which is common to all 4 treatments, the endowment, Z, of each decision maker is received with probability $p(G) \in [0, 1]$ that is increasing in G; this is the 'good' environmental state. However, with probability $1 - p(G)$, a 'bad' environmental state occurs such that each decision maker receives nothing, i.e., loses their entire endowment Z due to a potential environmental catastrophe. How much green contributions should the decision makers engage in, at time $t_1 = 0$? This setup encapsulates the first 3 of our 4 features listed above (temporal dimension, risk and

well on a voluntary basis, even in the absence of mandatory institutional solutions. But which of these institutions is more efficacious for green investment? We partly explore this question.

[5] Calzolari et al. (2018) and Ghidoni et al. (2017) consider a repeated game model of emissions and how cooperation might be influenced by persistence in pollution. However, there are fundamental differences from our work. First, we consider an abatement problem while they consider a damage problem. Second, we are interested in uncovering the primitives of temporal environmental choices in terms of time preferences and risk preferences (present-bias and loss aversion), but these factors do not play a fundamental role in their analysis. Third, we are also interested in comparing the voluntary privately optimal solution with mandatory tax-financed contributions chosen under representative democracy, while they are only interested in the privately optimal solution.

[6] Dannenberg et al. (2015) find that ambiguity is detrimental to public goods contribution and preplay communication can restore a degree of cooperation. Uncertainty makes cooperation harder to acheive (Rapoport et al. 1992; Gustafsson et al., 2000). When public goods are used to prevent a loss (rather than to affect a gain), then greater uncertainty in repeated threshold games produces greater cooperation (Milinski et al., 2008).

uncertainty, and public goods).

We assume that decision makers have present-biased time preferences that take the quasi–hyperbolic form and instantaneous preferences are of the Köszegi-Rabin form.[7] Now consider treatment T2, which also has voluntary green contributions, as in treatment T1. However, in order to isolate the effect of present-bias, in treatment T2, we introduce a commitment technology in the following manner. Unlike treatment T1, in treatment T2 the time pattern of endowments over the three time periods, t_1, t_2, t_3, is respectively, $0, Y, Z$. Each of the n decision makers simultaneously makes a voluntary green contribution decision at time t_1 on allocating their time t_2 endowment Y between consumption and green contributions at time t_2. When time t_2 arrives, the experimenter faithfully implements the decision made at time t_1 by the decision maker (commitment technology). The sum of all green contributions by the n individuals at time t_2 constitutes the green fund G. The time t_3 events are identical in all 4 treatments, hence, the green fund G contributes to reducing the probability $1 - p(G)$ of the 'bad' environmental state in which the decision maker loses the time t_3 endowment Z due to climate catastrophe.

Treatment T3 is the strict analogue of T1 and treatment T4 is the strict analogue of T2. Thus, it follows that the endowment patterns in treatments T1 and T3 are $(Y, 0, Z)$ and in treatments T2 and T4 they are $(0, Y, Z)$. The contributions in treatments T3 and T4 are financed through mandatory income taxes, earmarked for the green fund, and are paid by all n decision makers. The income tax rate chosen by the median voter at time t_1 is implemented at (i) time t_1 in treatment T3 and (ii) at time t_2 in treatment T4. The sum of all income tax revenues constitutes the green fund, G, which reduces the probability of the bad environmental state at time t_3.

Our treatments T2 and T4 are similar in spirit to the SMarT savings plan of Thaler and Benartzi (2004), where individuals are asked to make binding commitments about their future consumption/green contributions decisions.[8] We conduct two sets of experiments. In our 'first set of experiments', we use the method outlined above. However, in our 'second set of experiments' conducted with a longer time horizon, we follow the method in Thaler and Benartzi (2004) more closely in the details. Essentially, we ask subjects in treatment T2, if relative to their current green contributions from current income they would like to commit now to contributing from their future incomes (at time t_2) the same amount, 3% higher, 10% higher, 15% higher, or a lower amount.[9] Benchmarking current against future choices as in Thaler and Benartzi (2004), provides a cleaner test of the demand for commitment.

1.2 Predictions of the theoretical model

Our theoretical model makes the following predictions.

1. *Effects of present-bias*: An increase in present bias (i.e., a decrease in β in the (β, δ) model), relatively increases the marginal utility of current consumption. This reduces individual

[7] However, unlike Köszegi-Rabin (2006, 2009), we do not assume that the reference point is stochastic, state dependent, and consistent with the rational expectations (in the sense of Köszegi-Rabin and their three equilibrium concepts). But we do allow for the rational expectations of income to be the reference point. For a model of optimal climate policies when the discount rate deviates from exponential discounting, see Gerlagh and Liski (2018).

[8] For examples of such commitment devices and a survey of their effectiveness, see Dhami (2019, Vol. 3). Such devices have been shown to increase cooperation in common resource extraction problems (Dengler et al., 2018).

[9] All our experiments are incentive compatible.

green contributions, reducing the total green investment fund, G, and the probability, $p(G)$, of the good environmental state in the future.

2. *Effects of loss aversion*: An increase in current green contributions leads to the following two opposing effects on account of loss aversion. (i) Reduction in the current consumption, which reduces current marginal utility on account of loss aversion.[10] (ii) By increasing the green fund, G, it reduces the probability of the bad environmental state at time t_3, which reduces future disutility from loss aversion. For our estimated parameter values, the first effect dominates the second effect, and loss aversion reduces contributions. The intuition is that the future utilities are discounted at the rate $\delta < 1$ and there is only a probability $1 - p < 1$ of the bad environment state in the future, which discounts the second effect at the rate $\delta \times (1 - p) < 1$.

3. *Impatience*: An increase in the terminal date, t_3, by postponing the risk far enough into the future, reduces green contributions because the future is discounted.

4. *Commitment device*: In treatments T2, T4, the decision maker is offered the following commitment device. Make a decision at time t_1, on the green contributions to be made at a future time, t_2, which the experimenter faithfully implements at time t_2. We consider two variants of commitment. In our first set of experiments, we implement exactly as described above. However, in our second set of experiments, we follow the method in Thaler and Benartzi (2004) more closely, as described above, which results in a cleaner test of the demand for commitment. Present-biased individuals are predicted to contribute more under commitment.

   Under both types of commitment (first and second set of experiments), the green fund influences the probability of climate change at time t_3. This leads to another prediction. For a fixed t_3, an increase in t_2, by reducing the gap between the green contributions and the realization of uncertainty in treatments T2 and T4, reduces effective discounting between the two periods, increasing green contributions. Thus, the commitment device is more likely to be efficacious for increasing contributions, if the gap $t_3 - t_2$ is low.

5. *Social versus private optimum*: The optimal contributions under private voluntary contributions are predicted to fall short of the social optimum. This shortfall arises from two sources. (i) The present-bias parameter of the decision makers, which a social planner might not take into account, and (ii) the failure of decision makers to take account of the externality (in terms of a change in the probability of the bad environmental state) that they cause to others by choosing higher green contributions.[11]

6. *Effect of institutions*: The contrast between the magnitude of green contributions under voluntary private contributions (treatments T1, T2) and under representative democracy (treatments T3, T4) is an empirical question that cannot be predicted by our theory.[12]

[10] This effect plays a critical role in the SMarT savings plan of Thaler and Benartzi (2004).

[11] This is a fairly routine exercise in public economics and in behavioral economics, so we do not impose a formal proof of it on the reader in this paper.

[12] The answers depend, in complicated ways, on the specific assumptions that one is willing to make on the shape

1.3 Experiments and findings

We ran two sets of experiments. In our first set of experiments, we used data collected over the period September 2022 to February 2023, from 515 student subjects in 4 Indian Universities. Subjects were randomly assigned to the 4 treatments T1–T4. We then conducted a second set of experiments in July 2023 with 103 students from Ashoka University, but only with treatments T1 and T2, for the main purpose of testing the predictions of our model with longer time horizons. However, we also implemented the specific commitment technology used in Thaler and Benartzi (2004), described above. We report the results of the second set of experiments separately.

We adapt the bisection method to measure loss aversion (Abdellaoui, 2000; Dhami et al., 2023a; Dhami et al., 2023b) and we use the convex time budget constraints method to measure temporal preferences (Andreoni and Springer, 2012, Andreoni et al., 2015). The estimates of the temporal parameters, (β,δ), from the pooled sample, are $(1.0036, 0.9969)$. Across all subjects, the mean of loss aversion parameter is 2.03 with a median value of 1.55. These estimated parameter values are consistent with earlier studies that estimate loss aversion[13] and the temporal parameters[14].

Consistent with the model's predictions, we find that higher loss aversion significantly reduces green contributions. The effect of present-bias is large and negative (as predicted by our model) but attains statistical significance with longer time horizons. Higher time t_3 endowments, Z, that might be lost due to climate change, lead to higher contributions. In the first set of experiments, we find that, as predicted, when the time gap $t_3 - t_2$ increases to 25 weeks from the reference category of 1 weeks, contributions decrease statistically significantly. Tax-financed contributions under the institutional mechanism (median tax rate) are lower relative to the contributions under the voluntary contributions mechanism. Under commitment, contributions are highest in treatment T2. Furthermore, in our second set of experiments, about 50% of our subjects take up the commitment device and make higher contributions. Even those who do not take up a commitment device do not reduce contributions.

1.4 Schematic outline

Section 2 describes the basics of our model. Sections 3 and 4 derive the theoretical predictions and state the comparative static results under, respectively, a voluntary contribution mechanism and an institutional mechanism based on the median tax choice. Section 5 describes the experiments, the data, and gives the descriptive statistics. Sections 6 and 7, respectively, give the regression results from the first and second set of experiments. The Appendix contains all proofs and also describes our methods for measuring the behavioral parameters of loss aversion and present-bias. The supplementary section contains further robustness results; the table of choices used for eliciting time preferences; and the experimental instructions.

of the underlying utility functions, the degree of intertemporal substitution, the joint distribution of the temporal parameters (β, δ), as well as the parameters of prospect theory preferences.

[13]In their meta study, Brown et al. (2023) find that the mean loss aversion coefficient is 1.955. Gachter et al. (2022) find that the mean subject-specific loss aversion for riskless choice is 2.12 and the median is 1.73.

[14]In their meta study, Imai et al. (2021) find that the present-bias parameter, β, is approximately $0.95 - 0.97$ for studies that follow the CTB protocol. However, for monetary-reward studies the estimates of β are close to 1. Andreoni et al. (2015) estimate δ to be 0.9986.

2 Model

We consider three integer time periods $t_1 = 0 < t_2 < t_3$, such that t_1 is the current time period where decisions are made in all treatments; at times t_2, t_3 there are only consequences of the time t_1 decision.[15] We vary t_2, t_3 in our experiments. The description of the game at time t_3 is identical in every treatment.

There is a set of n decision makers, indexed by $i = 1, 2, ..., n$, where n is odd. All decision makers have identical endowments but potentially different underlying behavioral parameters. The decision makers simultaneously choose contributions (voluntary, or mandatory tax-financed contributions) to a green fund that is earmarked to reduce future climate risk and has the nature of a public good. All this is public knowledge.

2.1 Overview of the 4 treatments

There are 4 treatments, {T1,T2,T3,T4}. Treatments T1 and T2 consider 'voluntary contributions' to the green fund, while in Treatments T3 and T4, contributions to the green green fund are financed through 'mandatory income taxes'; the median tax across the most preferred tax rates of the decision makers is implemented. We do not consider the 'savings for consumption smoothing channel' because we wish to isolate the other channels predicted by our theoretical model cleanly, with minimal confounds.[16] Contributions to the green fund mitigate the probability of the adverse effects of climate change.

Table 1: Description of the treatments

Treatment	t_1	t_2	t_3
T1 and T3	Y	0	Z
	c_{t_1}	0	c_{t_3}
T2 and T4	0	Y	Z
	0	c_{t_2}	c_{t_3}

Depiction of one dimension of our 2×2 design along the dimension of the timing of endowments and consumption levels in each treatment. In each cell of the table, the first and second rows give, respectively, the endowment and the consumption, in that time period.

Over the three time periods, t_1, t_2, t_3, we consider two different temporal endowment patterns (identical for each of the n decision makers) in the different treatments. We summarize the situation in Table 1 and describe it in detail below. We denote consumption at time $t = t_1, t_2, t_3$ by c_t. In each cell of the table, the first and second rows give, respectively, the endowment and the consumption, in that time period.

1. Treatments T1, T3: The endowment pattern over the time periods t_1, t_2, t_3 is respectively, $Y, 0, Z$. The decision maker receives no endowment at time t_2. The decision on contributions to the green fund are made and implemented at time t_1; hence current consumption c_{t_1} occurs

[15]We do not use the less cumbersome notation for the time periods, $t = 0, 1, 2$, because this suggests a linear and constant difference between successive time periods, while we allow for any non-linear difference between the time periods. For instance, we are interested in the comparative static effects of the time gap $t_3 - t_2$ on contributions.

[16]For specific theoretical and empirical results on the conventional savings channels, while still allowing for the operation of behavioral factors such as loss aversion and present-bias, see Dhami et al. (2023a).

at time t_1. The sum of all contributions at time t_1 determines the green fund G available for time t_3.

2. Treatments T2, T4: The endowment pattern over the time periods t_1, t_2, t_3 is $0, Y, Z$. Hence, the decision maker receives no endowment at time t_1. The decision maker decides, at time t_1, the split of the time t_2 endowment Y, between consumption at time t_2, c_{t_2}, and contributions to the green fund. The time t_1 decision is faithfully implemented at time t_2 by the experimenter for all decision makers; in this sense, treatments T2 and T4 offer a commitment device. The sum of all contributions determines the green fund G available for time t_3. The purpose of treatments T2, T4 is to explore, in the spirit of the SMarT savings plans (Thaler and Benartzi, 2004), if a binding commitment to a future action, for a present-biased individual, improves the level of contributions.

Thus, in effect, we have a 2×2 design. Along one dimension, we vary the pattern of endowments to offer a commitment device; this is shown in Table 1. Along the second dimension, we vary the institutions that determine the contributions to the green fund (voluntary contributions in treatments T1, T2 versus income tax financed mandatory contributions in treatments T3, T4). We now explain the treatments in detail, separating treatments with and without commitment.

2.2 Treatments T1 and T3 (no commitment)

In treatments T1 (voluntary contributions) and T3 (mandatory tax-financed contributions), there is no commitment device, and the sequence of moves is as follows.

1. Time t_1: In treatment T1, each of the n decision makers receives an endowment Y. The decision makers simultaneously decide on allocating Y between current consumption, c_{t_1}, and their contributions, g_i, $i = 1, 2, ..., n$, towards a green fund, $G = \sum_{i=1}^{n} g_i$, and this is common knowledge. The budget constraint of decision maker i at time t_1 in treatment T1 is

$$c_{t_1} = Y - g_i. \tag{2.1}$$

   In Treatment T3, each of the n decision makers first simultaneously choose their most preferred tax rate τ to pay on their endowment Y. It is common knowledge that the tax revenues are earmarked for a green fund. Each decision maker then pays an income tax τY on their endowment. The chosen tax rate $\tau \in [0, 1]$ is the median tax rate among the most preferred tax rates stated by all n decision makers.[17] The total tax paid at time t_1 equals $n\tau Y$ and this constitutes the green fund, $G = n\tau Y$. The budget constraint of the consumer is

$$c_{t_1} = (1 - \tau) Y. \tag{2.2}$$

2. Time t_2: Nothing occurs at time t_2 in treatments T1 and T3.

3. Time t_3: Each of the n decision makers receives an endowment Z. There are two climatic states of the world, $s \in \{g, b\}$. The good state, $s = g$, arises with an endogenous probability

[17] We show that under certain conditions, the median tax rate is also the Condorcet winner.

$p \in [0, 1]$, and the bad state, $s = b$ arises with probability $1 - p$. In state $s = g$, each decision maker retains the time t_3 endowment Z, but in state $s = b$, each decision maker loses the endowment due to an environmental catastrophe.[18]

The green fund, G, raised at time t_1, increases the probability, p, of the good state, $s = g$, at time t_3.[19] The probability of the good state, $s = g$, satisfies

$$p(G) : [0, nY] \to [0, 1] \, ; p' > 0, p'' < 0. \tag{2.3}$$

Thus, an increase in G increases the probability of the good state at time t_3, but there are diminishing returns to the underlying risk abatement technology, as reflected in the concavity of p. A special case of (2.3) that we use in our experiments is

$$p(G) = \left(\frac{G}{nY}\right)^{\gamma} \in [0, 1] \, ; \gamma \in (0, 1) \, . \tag{2.4}$$

From (2.4), if there are no contributions ($G = 0$) then $p = 0$, and if $G = nY$ (i.e., everyone contributes their time t_1 endowment Y fully towards the green fund) then $p = 1$. For intermediate values of G we have $p \in (0, 1)$.

In effect, at time t_3, decision makers face the following risky lottery

$$(0, 1 - p(G); Z, p(G)) \, . \tag{2.5}$$

The probability p takes the following form in Treatment T3. Using (2.4) and $G = n\tau Y$, the probability of the good state $s = g$ at time t_3 is

$$p(G) = \tau^{\gamma} \in [0, 1] \, ; \gamma \in (0, 1) \, . \tag{2.6}$$

2.3 Treatments T2 and T4 (commitment)

In treatments T2 (voluntary contributions) and T4 (mandatory tax-financed contributions), there is a commitment device, and the sequence of moves is as follows.

1. Time t_1: In treatment T2, at time each decision maker knows at time t_2 that they will receive an endowment $Y > 0$ at time t_2. Decision maker $i = 1, ..., n$ is asked, at time t_1, to make a binding commitment to allocate income Y to be received at time t_2 between consumption at time t_2, denoted by c_{t_2}, and contributions, g_i, towards a green fund. All n decision makers choose simultaneously.

   In treatment T4, the n decision makers, at time $t_1 = 0$, simultaneously state their most preferred tax rate τ for the income tax τY to be paid at time t_2 on their time t_2 endowment Y. The median tax rate among all decision makers is implemented at time t_2. It is common knowledge that all tax revenues are earmarked towards a green fund.

[18] This is not a restrictive assumption. Our insights also go through if we assumed that in state $s = b$, the decision maker loses only a fraction of the endowment, Z.

[19] For instance, the green fund might have been invested in flood defenses; reducing harmful greenhouse gas emissions; and developing new technologies to clean the environment, thereby reducing the risk of environmental damage and the harm caused to people.

2. Time t_2: Each of the n decision makers receives an endowment Y. Their time t_1 decisions are faithfully implemented by the experimenter (commitment device).

   For the voluntary contribution mechanism (treatment T2), the experimenter faithfully implements the time t_1 contributions choice of each decision maker (implementation of commitment). Hence, the experimenter deducts an amount g_i from the endowment Y of decision maker $i = 1, ..., n$ resulting in the green fund $G = \sum_{i=1}^{n} g_i$ being raised at time t_2. Thus, at time t_2, decision maker $i = 1, ..., n$ consumes an amount

$$c_{t_2} = Y - g_i. \tag{2.7}$$

   In treatment T4, the implemented tax at time t_2 is the median tax rate across all most preferred tax rates chosen at time t_1. At time t_2, the experimenter deducts an amount τY from the endowment Y of each decision maker (implementation of commitment for time t_1 decisions) and earmarks it as the green investment fund $G = n\tau Y$. Thus, the consumption of decision maker i is

$$c_{t_2} = (1 - \tau) Y. \tag{2.8}$$

3. Time t_3: Identical to treatments T1 and T3, as explained above in Section 2.2.

2.4 Intertemporal preferences

Since $t_1 = 0$, we have $t_3 - t_1 = t_3$ and $t_2 - t_1 = t_2$. Individuals have (β, δ) or quasi-hyperbolic preferences to capture present-bias. The intertemporal preferences of the decision maker at time t_1 are given by

$$U = \begin{cases} v\left(c_{t_1}; r_{t_1}\right) + \beta\delta^{t_3} Ev(c_{t_3}; r_{t_3}) & if \ (Y, 0, Z);\ T1, T3 \\ \beta\delta^{t_2}\left[v\left(c_{t_2}; r_{t_2}\right) + \delta^{t_3 - t_2} Ev(c_{t_3}; r_{t_3})\right] & if \ (0, Y, Z);\ T2, T4 \end{cases} ; \beta \in (0, 1], \delta \in (0, 1]. \tag{2.9}$$

In (2.9), the instantaneous utility function, $v(c_t; r_t)$, is the Köszegi-Rabin (2006, 2009) prospect theory utility function at time $t = t_1, t_2, t_3$ and r_t is the time t reference point. We describe instantaneous preferences and the reference points in Section 3.1 below in detail for the case of voluntary contributions (T1, T2); for mandatory tax-financed contributions (T3, T4), we describe these preferences in Section 4 below.

The first row of (2.9) captures preferences in treatments T1 and T3; while the second rows captures preferences in treatments T2 T4. In the first row of (2.9), if $\beta \in (0, 1)$, then the *present-bias parameter* β shrinks future utility relative to current utility in the initial time period, t_1. The *extent of the present-bias* is given by $1 - \beta$; thus, an increase in β reduces present-bias. If $\beta = 1$, 'present-bias' disappears completely, leaving only the 'impatience' embedded in the classical discount factor $\delta \in (0, 1]$; this special case is the exponential discounted utility model. This clarifies the sense in which the terms 'present-bias' and 'impatience' are used.

When the endowment pattern is $(Y, 0, Z)$ (treatments T1, T3; first row of (2.9)), consumption occurs at the current time t_1 when the consumption-contributions decision is made. Hence, the time t_3 utility is discounted at the rate $\beta\delta^{t_3}$. However, when the endowment pattern is $(0, Y, Z)$ (treatments T2, T4; second row of (2.9)), all consumption occurs in the future at dates t_2, t_3 (recall that all discounting is from the perspective of the current time period, t_1). In this case, the present-bias parameter, β, does not play any role in influencing the 'relative weights' assigned to the two consumption levels at the dates t_2, t_3. Indeed, the optimal choices in the second row of (2.9) can be shown to be independent of the common term $\beta\delta^{t_2} > 0$. Thus, when interested in deriving the optimal choices in what follows, we omit the common term $\beta\delta^{t_2}$ in the second row of (3.1), without altering the first order condition or the optimal choices, and rewrite it as

$$U = \begin{cases} v\left(c_{t_1}; r_{t_1}\right) + \beta\delta^{t_3} Ev(c_{t_3}; r_{t_3}) & if\ (Y, 0, Z);\ T1, T3 \\ v\left(c_{t_2}; r_{t_2}\right) + \delta^{t_3 - t_2} Ev(c_{t_3}; r_{t_3}) & if\ (0, Y, Z);\ T2, T4 \end{cases}; \beta \in (0, 1], \delta \in (0, 1]. \tag{2.10}$$

The expectation operator conditional on the information set at time t_1 is denoted by E.[20]

Using (2.10), let us define the discount factor θ_T, which captures the effects of discounting in treatments T1, T2, T3, T4.

$$\theta_T = \begin{cases} \beta\delta^{t_3} & if\ (Y, 0, Z),\ T1, T3 \\ \delta^{t_3 - t_2} & if\ (0, Y, Z),\ T2, T4 \end{cases}; \beta \in (0, 1], \delta \in (0, 1]. \tag{2.11}$$

Remark 1 *From (2.11) and comparing the two rows, we have that $\beta\delta^{t_3} \leq \delta^{t_3 - t_2}$ (and with strict inequality if $\delta < 1$). Thus, from the perspective of time t_1, the weight placed on the time t_3 payoff is lower in treatment T1 relative to treatment T2, where all implemented decisions are in the future at dates t_2, t_3. This arises partly on account of the present-bias parameter, β; the sharp increase in impatience as a choice is brought back towards the present at time $t_1 = 0$ in T1, and this is a distinguishing feature of the quasi-hyperbolic discounting model.*

We now describe the details of the individual treatments, discussing separately the treatments T1, T2 (voluntary contributions to green investment) and treatments T3, T4 (tax-financed green investments).

3 Voluntary contribution mechanisms: Treatments T1, T2

In this section, we consider the solution under the voluntary contributions mechanism. In the next section, Section 4, we consider mandatory tax financed contributions.

3.1 Instantaneous preferences under voluntary contributions

We define Köszegi-Rabin preferences in their standard form at time t as

$$v(c_t; r_t) = u(c_t) + \mu\phi(c_t - r_t), \mu \in (0, 1], \tag{3.1}$$

where $u : \Re \longrightarrow \Re$ is the instantaneous utility that one receives from the 'absolute level' of consumption. We assume that $u(0) = 0$, and u is increasing and concave

$$u' > 0;\ u'' < 0;\ u(0) = 0. \tag{3.2}$$

[20] Thus, in full notation, the last term in each of the two rows in (2.10), is $E\left[v(c_3; r_{t_3}) \mid I_0\right]$, where I_0 is the information set at time $t_1 = 0$.

The second term on the RHS in (3.1) with relative weight, $\mu \in (0, 1]$, is gain-loss utility relative to the reference point, $\phi : \Re^2 \longrightarrow \Re$, given by

$$\phi(c_t - r_t) = \begin{cases} (c_t - r_t) & if\ c_t \geq r_t \\ -\lambda(r_t - c_t) & if\ c_t < r_t \end{cases}, \tag{3.3}$$

where λ is the parameter of loss aversion in prospect theory. The linear form of gain-loss utility in (3.3) follows the suggestion of 'linearity over small stakes' in Köszegi-Rabin (2006, 2009), and this has good empirical support.

There is good evidence that under certainty, the "status-quo" provides a satisfactory reference point (Kahneman and Tversky, 2000; Dhami, 2019, Vol. 1). Recall that the initial endowment Y is received at time $t = t_1$ in treatments T1, T3 and at time $t = t_2$ in treatments T2, T4. Hence, we take the time $t = t_1$ reference point, r_{t_1}, in treatments T1, T3; and the time $t = t_2$ reference point, r_{t_2}, in treatments T2, T4 as the status quo income, Y, thus[21]

$$r_{t_1} = Y;\ r_{t_2} = Y. \tag{3.4}$$

It is less clear how reference points are formed in time periods where there is risk; in all treatments, this is the case in time period t_3. Proposals include using the rational expectations of future incomes, the expected value of the future income, or a fraction of the expected value (Kahneman and Tversky, 2000; Dhami, 2019, Vol. 1); we allow for these possibilities in our paper.[22] Denoting the time $t = 3$ reference point by r_{t_3}, we allow for all reference points that satisfy

$$0 < r_{t_3} < Z. \tag{3.5}$$

For instance, if the reference point is the time t_3 expected income, i.e., $r_{t_3} = E[Z] = pZ$, then (3.5) is satisfied. The condition in (3.5) is also satisfied by any convex combination of the incomes in the two states (0 and Z) at time t_3. It follows from (3.5) that at time t_3, in the 'good state' $s = g$, the decision maker is in the domain of gains, and in the 'bad state' $s = b$, the decision maker is in the domain of losses, which is reasonable, particularly because in our experiments $Z \geq Y$.

Using the budget constraints in (2.1) and (2.7), we have $c_t = Y - g_i$, $t = t_1, t_2$. Hence, using (2.10) –(3.5) we can write the time $t = t_1, t_2$ Köszegi-Rabin utility as[23]

$$v\,(c_t; r_t) = u\,(Y - g_i) - \lambda\mu g_i;\ t = t_1, t_2. \tag{3.6}$$

For time $t = t_3$, we can write the Köszegi-Rabin utility as

$$Ev(c_{t_3}; r_{t_3}) = Eu(Z) + \mu E\phi(c_{t_3} - r_{t_3}), \tag{3.7}$$

[21] In the context of lifecycle models, this is also the assumption made in Thaler and Benartzi (2004).

[22] Köszegi-Rabin (2006, 2009) preferences allow for stochastic, state-dependent, reference points that are consistent with the rational expectations via three types of equilibrium concepts. However, the rationality and cognitive requirements for such a reference point are incredibly strigent and in our view, unlikely to be met in one shot lab experiments that do not allow for any learning opportunities. Furthermore, in our view there is no definitive empirical evidence in support of such reference points, at least for one shot experimental games (Dhami, 2019 Vol. 1; Dhami and Sunstein, 2022). Such reference points might be more suitable in other contexts with a large number of repetitions and learning opportunities.

[23] For treatments T1 and T2, when the contributions are made at respectively, time $t = t_1$ and $t = t_2$, we have $c_t = Y - g_i$ and $r_t = Y$ (see (3.4)), so $c_t - r_t = -g_i \leq 0$. Hence, the second row of (3.3) applies, so $\phi = -\lambda g_i$.

where (recalling that income in the bad state, $s = b$, equals zero and $u(0) = 0$), we have

$$Eu(Z) = p(G)u(Z), \tag{3.8}$$

and, recalling the restriction on r_{t_3} in (3.5), we have

$$E\phi(c_{t_3} - r_{t_3}) = p(G)\left(Z - r_{t_3}\right) - \left(1 - p(G)\right)\lambda\left(r_{t_3} - 0\right). \tag{3.9}$$

Substituting (3.8), (3.9) in (3.7), we get the time t_3 Köszegi-Rabin utility as

$$Ev(c_{t_3}; r_{t_3}) = p(G)\left(u(Z) + \mu Z\right) - \mu r_{t_3}\left(p(G) + \lambda\left(1 - p(G)\right)\right). \tag{3.10}$$

3.2 Optimization problem under voluntary contributions

Using (2.10), (3.1), (3.3), (3.6), (3.7) (2.11) we get the unconstrained maximization problem of decision maker i, in treatments T1, T2, given the choices of the other players captured in the contributions vector, $\mathbf{g}_{-i}$:

$$g_i^* \in argmax\ U = \left[u\left(Y - g_i\right) - \lambda\mu g_i\right] + \theta_T\left[p(G)\left(u(Z) + \mu Z - \mu r_{t_3}\right) - \mu r_{t_3}\lambda\left(1 - p(G)\right)\right], \tag{3.11}$$

given

$$\mu > 0, g_i \in [0, Y], \mathbf{g}_{-i},$$

and θ_T is defined in (2.11). For treatment T1, θ_T is defined in the first row of (2.11), and for treatment T2 it is defined in the second row of (2.11).

Remark 2 *In treatments T1, T2, we allow for heterogeneity between the decision makers, potentially with respect to the behavioral parameters, $\lambda, \beta, \delta, \mu$. However, in order to minimize notation, we omit subscripts for the decision makers on these parameters, such as $\lambda_i, \beta_i, \delta_i, \mu_i$; $i = 1, ..., n$.*

Since U in (3.11) is continuously differentiable, we have by direct differentiation

$$\frac{\partial U}{\partial g_i} = \left(-u'\left(Y - g_i\right) - \lambda\mu\right) + \theta_T p'(G)\left[\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda\right]. \tag{3.12}$$

The two terms on the RHS of (3.12) give the marginal effects of an increase in the time $t = 1$ contributions, g_i, of the i^{th}decision maker by a unit. The first term captures the following two kinds of current marginal costs. (a) A reduction in current marginal utility at time t_1 in treatment T1 and at time t_2 in treatment T2, and (b) loss aversion from parting with some of the current endowment, Y, in the form of contributions, g_i, as in Thaler and Benartzi (2004).[24] The second term captures the future marginal benefits at time t_3 that arise from two sources. (a) An increase in g_i increases the probability of the good state $s = g$ at time t_3, hence, increasing the expected absolute utility at time t_3. (b) It also reduces expected losses in the future gain-loss utility term via changes in $p(G)$, the size of which depends on the size of the loss aversion parameter, λ.

[24] Just as the first unit of savings creates a current cost in terms of a fall in marginal utility, $-u'\left(Y - g_i\right)$, it also creates a current cost in terms of loss aversion, which is not present in the traditional model. Thaler and Benartzi (2004) are explicit about this channel and they write (p. S169-70): "Loss aversion affects savings because once households get used to a particular level of disposable income, they tend to view reductions in that level as a loss. Thus, households may be reluctant to increase their contributions to the savings plan because they do not want to experience this cut in take-home pay."

Differentiating (3.12) again, we get

$$\frac{\partial^2 U}{\partial g_i^2} = u''\left(Y - g_i\right) + \theta_T p''(G)\left[\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda\right] < 0. \tag{3.13}$$

From (3.11), (3.13), for any vector of contributions of the other players, $\mathbf{g}_{-i}$, the objective function of decision maker i is twice continuously differentiable, strictly concave in g_i, and defined over a closed and bounded interval. Hence, a unique maximum value, g_i^*, exists.

The expression in (3.12), with the RHS set equal to zero, gives the optimal solution in treatments T1 and T2, respectively, for the two cases of θ_T defined in the two rows of (2.11). We have avoided introducing an additional subscript on the optimal solution g_i^* to differentiate the two solutions (e.g., g_{im}^*, where $m = 1$ for treatment T1 and $m = 2$ for treatment T2).

We now outline an important example for the comparative static effects of loss aversion.

Example 1 *: A key determinant of contributions in our model is loss aversion. We formally study the comparative static effects in Proposition 1. But here we give an illustrative example. From (3.12), an increase in the parameter of loss aversion λ has the following net marginal effect*

$$-\mu\left[1 - \theta_T r_{t_3} p'(G)\right] \gtreqless 0. \tag{3.14}$$

From (3.14), the net marginal effect of loss aversion on contributions is an empirical question. On the one hand, current loss aversion reduces marginal contributions towards the public good by μ units, but on the other hand, the future reduction in the probability of the bad state increases the incentive to contribute by $\mu\theta_T r_{t_3} p'(G)$ units. However, for all possible parameter estimates and simulations, our data overwhelmingly shows that

$$1 > \theta_T r_{t_3} p'(G), \tag{3.15}$$

so that the net effect of loss aversion is to reduce contributions. This is, we believe, the first demonstration of such a result in the literature. To get a feel for the numbers involved, we ran a Monte Carlo simulation of 1000 random samples of the $n = 7$ subjects' contributions, and then calculated the group contributions, G. Using the parameters used in our experiments (e.g., we used $\gamma = 0.5$ in (2.4) in our experiments), we get $p'(G) = 0.0013$. The sample estimates of β and δ, respectively, for our data, are 1.0036 and 0.9969. The time unit for the measured value of δ is in days. Hence, this is a daily discount factor.

We would like to show that the inequality in (3.15) holds even when we make the term $\theta_T r_{t_3} p'(G)$, as large as possible. From the first row of (2.11), $\theta_T = \beta\delta^{t_3}$ in treatment T1. Since $\beta = 1.0036$ and $\delta = 0.9969 < 1$, let us take t_3 to be the conservative value of 5 weeks or 35 days (increasing t_3 reduces the size of δ^{t_3} making θ_T smaller so makes it even more likely that (3.15) holds).[25] The highest possible value of $r_{t_3} = Z$. A representative value of Z is $Z = 200$ in our experiments. Thus, we can check that

$$\theta_T r_{t_3} p'(G) = (1.0036)(0.9969)^{35}(200)(0.0013) = 0.2341 < 1,$$

[25] In our first set of experiments, for treatments T1 and T3, the terminal date, t_3, is 5 weeks and 25 weeks (for three questions, it is 5 weeks, and in one question, it is 25 weeks). In treatments T2 and T4, the intermediate date, t_2, is 5 weeks in 4 questions; and, in the extra two questions, it is 1 week and 9 weeks. The terminal date t_3 is 10 weeks (in 5 questions) and 30 weeks in one question for treatments T2 and T4. In our second set of experiments, with a longer time horizon, we also consider t_3 to be 52 weeks.

which comfortably holds. If we picked an even more conservative value of t_3 *equals to 1 week, or 7 days, then we still have* $\theta_T r_{t_3} p'(G) = 0.2553 < 1$*. If we had picked the highest possible value of* $Z = 400$ *in our experiments, then too we have* $\theta_T r_{t_3} p'(G) = 0.468 < 1$*. In every possible simulation that we tried (including using the second row of (2.11)), the inequality in (3.15) comfortably holds. Thus, we expect an increase in loss aversion to reduce contributions.*

3.3 Solution and predictions under voluntary contributions

Our assumptions guarantee the existence of a Nash equilibrium. These are: non-empty compact strategy spaces that are subsets of a convex Euclidean space and objective functions which are continuous and strictly concave in the contribution choices. As our solution concept we take the *symmetric Nash equilibrium* (SNE).[26] In a SNE all players choose identical contributions, $g_i^* = g^*$. This can be found by setting $g_i = g$ in (3.12) and setting the RHS equal to zero. Thus, a SNE solves

$$\frac{\partial U}{\partial g_i} = \left(-u'\left(Y - g^*\right) - \lambda\mu\right) + \theta_T p'(ng^*)\left[\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda\right] = 0. \tag{3.16}$$

From (3.16), we cannot rule out corner solutions $g^* = 0$ and $g^* = Y$, unless we impose further technical restrictions.[27] In our analysis below, we assume an interior solution. Since the first term on the RHS in (3.16) is strictly negative, an interior solution requires that the second term on the RHS in (3.16) must be strictly positive.

The comparative static results that can be tested with our data are summarized in the next proposition.

Proposition 1 *Assume an interior solution to the SNE,* $g^* \in (0, Y)$*. Then the comparative static results are as follows:*

(a) (Loss aversion) g^* *is decreasing (resp. increasing) in the parameter of loss aversion,* λ*, if* $\theta_T r_{t_3} p'(G)$ *is less than (resp. greater than) 1.*[28]

(b) (Present-bias) In treatment T1, g^* *is decreasing in the magnitude of present-bias,* $1 - \beta$*, but in treatment T2, there is no effect of* β *on optimal contributions* g^**.*

(c) (Size of time t_3 *endowment, Z)* g^* *is increasing in the size of the time* t_3 *endowment,* Z*.*

(d) (Effect of time delays) (i) The greater is the gap between time $t_1 \equiv 0$ *and* t_3 *(size of* t_3*), the lower is* g^**; and strictly lower if* $\delta < 1$*. (ii) The smaller is the time gap between time* t_2 *and* t_3 *(higher* t_2*, for a fixed* t_3*) the greater is* g^* *in treatment T2 (and strictly greater if* $\delta < 1$*), but there is no effect in treatment T1.*

(e) (Treatment contrasts between T1 and T2) Contributions are predicted to be higher (and strictly higher if $\delta < 1$*) under treatment T2 as compared to treatment T1.*

---

[26] We can also consider other solution concepts, such as a *best response to beliefs*, which is a particularly empirically relevant concept to use for experimental games where repetitions and learning are limited; see, for instance the discussion and references in Dhami et al. (2023c). Suppose that player $i = 1, ..., n$ has beliefs that the expected contributions of other players is $\mathbf{g}_{-i}^e$ and player i plays a best response to such beliefs. Our central comparative static results, that we test in our experiments, continue to hold in this case as well.

[27] We can rule out $g^* = 0$ by requiring $\left(-u'\left(Y\right) - \lambda\mu\right) + \theta_T\left[p'(0)\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda p'(0)\right] > 0$ and rule out $g^* = Y$ by requiring $\left(-u'\left(0\right) - \lambda\mu\right) + \theta_T\left[p'(nY)u(Z)\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda p'(nY)\right] < 0$.

[28] See Example 1, above, for the empirical plausibility of these two alternative cases.

Discussion of Proposition 1: Proposition 1 lists the comparative static results that we can directly test with our data. An increase in g^* by decision maker $i = 1, ..., n$ (i) decreases current utility, but (ii) increases future utility by decreasing the probability of the bad state, $s = b$, conditional on the contributions of others, g^*_{-i}. This leads to an ambiguous effect of loss aversion, λ, on contributions (Proposition 1a). Example 1 showed that for our estimated behavioral parameters, an increase in loss aversion reduces contributions, and this is confirmed by our data.

From Proposition 1b, in treatment T1, g^* is decreasing in the magnitude of present-bias, $1-\beta$, which reduces the weight placed on future marginal utility (through a fall in β). This reduces the future marginal benefit of a reduction in the probability of the bad state, hence, reducing optimal contributions. In treatment T2, since both relevant dates (t_2 and t_3) are in the future at the time of making the decision at time t_1, the parameter β has no effect on contributions (comparing the two rows in (2.11), β is missing from the second row).

From Proposition 1c, an increase in the endowment Z, at time t_3, increases the size of the loss in the bad state of the world in the future (where the entire endowment Z is lost). This increases the marginal costs of making low contributions. Decision makers then contribute more to reduce the likelihood of the bad state.

From Proposition 1d(i), if the future is more distant (higher t_3) then it is discounted more relative to the current loss in marginal utility from making extra contributions, hence, contributions optimally fall. From Proposition 1d(ii), the smaller is the gap between time t_2 and t_3, the less the future marginal benefits are discounted in treatment T2 (recall that $\theta_T = \delta^{t_3-t_2}$ in T2 in the second row of (2.11)), hence, contributions optimally increase in treatment T2. But there is no effect of this gap in treatment T1 where the contributions are made at time $t_1 = 0$ and $\theta_T = \beta\delta^{t_3}$, so time t_2 plays no role. One implication, and potential policy insight, is that to ensure higher levels of contributions towards green investments, individuals should be asked at time $t_1 = 0$ to make a contributions precommitment for a date t_2 that is as close as possible to the fixed date t_3. However, the effectiveness of this channel, and the precise form that commitment ought to take (as evidenced by the contrasts between our two sets of experiments; see Section 7) is an empirical question.

From Proposition 1e, in the treatment contrast T1 vs T2, we expect contributions to be relatively higher in T2. The intuition is that in treatment T1, the present-bias parameter β induces a relatively larger weight on the current loss in marginal utility from making higher contributions, reducing optimal contributions.

4 Mandatory tax financed contributions: Treatments T3, T4

We now consider a formal institutional mechanism for the provision of green investment that requires mandatory contributions through the tax system. On the one hand, mandatory taxes might reduce the free rider problem under voluntary contributions, but on the other hand, mandatory taxes may interfere with human autonomy and agency. Ultimately, relative contributions under mandatory tax-financed contributions and voluntary contributions is an empirical question that we examine.

Recall from Remark 2, that our model allows for multidimensional heterogeneity between the decision makers with respect to the parameters $\lambda, \beta, \delta, \mu$. It is important to stress that in our theoretical model and in our experiments, we are interested in implementing the tax rate chosen by the median voter, and this is common knowledge. Whether a median voter equilibrium does, or does not, exist is an issue that is orthogonal to our implementation of this institution. However, multidimensional heterogeneity violates the median voter theorem unless further strong restrictions are imposed. In this section, we show that when heterogeneity is unidimensional, then our chosen institution also has the property that a Condorcet winner exists such that the median tax rate is also the outcome of majority voting. Hence, assume there is heterogeneity with respect to loss aversion, but in all other respects, the decision makers are identical.[29] Denote the loss aversion parameter of decision maker $i = 1, ..., n$ by λ_i.

4.1 Optimization problem in treatments T3 and T4

We have already described the details of treatments T3 (in Section 2.2) and T4 (in Section 2.3). The probability of the bad environmental state in both treatments is given in (2.6). Using (2.6), and proceeding as in the derivation of (3.11), the most preferred tax rate of decision maker $i = 1, ..., n$, in Treatments T3, T4, can be found by solving the following unconstrained optimization problem[30]

$$\begin{aligned} \tau_i^* \in argmax\ U = & \left[u\left(Y\left(1-\tau_i\right)\right) - \lambda_i \mu\left(\tau_i Y\right)\right] + \\ & \theta_T \left[\tau_i^{\gamma}\left(u(Z) + \mu Z - \mu r_{t_3}\right) - \mu r_{t_3} \lambda_i \left(1 - \tau_i^{\gamma}\right)\right], \mu > 0, \tau_i \in [0,1], \end{aligned} \tag{4.1}$$

where θ_T is given in (2.11) and the two rows in (2.11) capture, respectively, the two cases of treatment T3 and T4. These two rows create different optimal values of τ_i^*, the most preferred tax rate of decision maker i, in treatments T3 and T4, respectively.[31]

4.2 Solution and predictions under voluntary contributions

Below we shall focus only on interior solutions. We first find the most preferred tax rate of decision maker i.[32] The first order condition is found by differentiating (4.1) with respect to τ_i and setting it equal to zero.

$$\frac{\partial U}{\partial \tau_i} = \left[-u'\left(Y\left(1-\tau_i\right)\right)Y - \lambda_i \mu Y\right] + \theta_T \gamma \tau_i^{\gamma-1}\left[\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda_i\right] = 0. \tag{4.2}$$

At any interior solution, the second term on the RHS of (4.2) is strictly positive. Hence, and using $\gamma \in (0,1)$, it follows that.

$$\frac{\partial^2 U}{\partial \tau_i^2} = u''\left(Y\left(1-\tau_i\right)\right)Y^2 - \theta_T \gamma\left(1-\gamma\right)\tau_i^{\gamma-2}\left[\left(u(Z) + \mu Z - \mu r_{t_3}\right) + \mu r_{t_3}\lambda_i\right] < 0. \tag{4.3}$$

[29] We could similarly have chosen heterogeneity with respect to any of the other parameters, β, δ, μ, one at a time.

[30] Notice that under mandatory tax financed contributions, we have $g_i = \tau_i Y$ because tax-financed contributions are earmarked for the green fund. From (2.2), the budget constraint in treatment T3 at time t_1 is given by $c_{t_1} = (1-\tau_i)Y$. Using $g_i = \tau_i Y$, this constraint can be written as $c_{t_1} = Y - g_i$ which is identical to the budget constraint in (2.1) for treatment T1 (voluntary contributions). A similar equivalence holds for the time t_2 budget constraints for treatment T4 and T2 (i.e., $c_{t_2} = (1-\tau_i)Y$ in T4 is equivalent to $c_{t_2} = Y - g_i$ in treatment T2). These equivalences are exploited in writing down the expression in (4.1).

[31] We do not introduce separate subscripts or superscripts on τ_i^* to distinguish between the optimal values in treatments T3 and T4; the context makes clear the treatment that we are referring to.

[32] In the experiments, each of the n voters is asked to state their most preferred tax rate and the median tax rate is implemented.

We summarize these observations in the next Lemma.

Lemma 1 *The objective function in (4.1) is strictly concave in the tax rate, τ_i. The most preferred tax rate of decision maker i, τ_i^*, exists and is unique. At an interior solution, τ_i^* is the solution to*

$$\left[-u'\left(Y\left(1-\tau_i^*\right)\right)Y-\lambda_i\mu Y\right]+\theta_T\gamma\tau_i^{*\gamma-1}\left[\left(u(Z)+\mu Z-\mu r_{t_3}\right)+\mu r_{t_3}\lambda_i\right]=0, \tag{4.4}$$

and the second term on the RHS of (4.4) is strictly positive.

We now describe the equilibrium under majority voting.

Lemma 2 *Suppose that heterogeneity across voters is unidimensional and, in particular, it is with respect to the loss aversion parameter only, λ_i, $i=1,...,n$. In a majority vote, where each voter votes sincerely and has the optimization problem in (4.1), in any pairwise comparison of the most preferred tax rates in a majority vote, the chosen tax rate is the tax rate most preferred by the median voter, τ_M^*.*

We now describe the comparative static results on the optimal choice of the tax rate for decision maker $i=1,...,n$; these effects are identical whether we choose τ_i^* or τ_M^* in the first order condition (4.4). Since (4.4) is expressed in terms of τ_i^*, in Proposition 2 below we study the comparative static effects with respect to τ_i^*.[33]

Proposition 2 *Assume an interior solution to the optimization problem in (4.1), $\tau_i^*\in(0,1)$. The comparative static effects are as follows:*
(a) (Loss aversion) τ_i^* is decreasing (resp. increasing) in the parameter of loss aversion, λ, if Y is greater (resp. less) than $\gamma\tau^{\gamma-1}\theta_T r_{t_3}$.
(b) (Present-bias) In treatment T3, τ_i^* is decreasing in the magnitude of present-bias, $1-\beta$, but in treatment T4, there is no effect of β on τ_i^*.
(c) (Size of time t_3 endowment, Z) τ_i^* is increasing in the size of the time t_3 endowment, Z.
(d) (Effect of time delays) (i) The greater is the gap between time $t_1\equiv 0$ and t_3 (size of t_3), the lower is τ_i^* (and strictly lower if $\delta<1$). (ii) The smaller is the time gap between time t_2 and t_3, the greater is τ_i^* in treatment T4 (and strictly greater if $\delta<1$), but there is no effect in treatment T3.
(e) (Treatment contrasts between T3 and T4) The optimal tax rate τ_i^* is predicted to be relatively higher in treatment T4 as compared to treatment T3 and strictly higher if $\delta<1$.

Discussion of Proposition 2: An increase (resp. decrease) in the chosen tax rate corresponds to higher (resp. lower) contributions/green fund. The comparative static effects for the choice of the optimal tax rate to finance contributions in treatments T3 and T4 (Proposition 2) are very similar to the comparative static effects for the private contributions mechanisms in treatments T1 and T2 (Proposition 1); and the same intuition applies.

[33] Suppose that we allowed for multidimensional heterogeneity such the the median voter theorem does not hold. In that case, the tax rate chosen by the median voter, say, τ_m, is still implemented in our theoretical model and in our experiments. All the results in Proposition 2 continue to hold, but they need to be stated with respect to τ_m instead of τ_i^*.

An increase in present-bias increases the relative marginal utility from current consumption, hence, reducing the desired tax rate earmarked for contributions towards the future (Proposition 2b). An increase in Z increases the size of the loss in the bad state of the world in the future, and loss averse decision makers then optimally choose a higher tax rate to make more contributions, in order to reduce the likelihood, $1-p(G)$, of the bad state (Proposition 2c). The more distant are the consequences of climate change (higher t_3) the more is the future discounted, reducing the marginal utility from choosing a higher tax rate to offset the probability of future climate damages (Proposition 2d(i)). The smaller is the gap t_3-t_2, the less the future marginal benefits are discounted in treatment T4 from the perspective of time t_1 (recall from (2.11) that $\theta_T=\delta^{t_3-t_2}$ from the perspective of time t_1). Hence, the optimally chosen tax rate increases in treatment T4 and contributions increase; but there is no effect in treatment T3 (Proposition 2d(ii)) where time t_2 plays no role. From Proposition 2e, since the present-bias parameter does not diminish future marginal utility in treatment T4, relative to treatment T3, tax-financed contributions in treatment T4 are predicted to be relatively higher.

As in Proposition 1(a), one of the key comparative static results is with respect to loss aversion, given in Proposition 2(a). Whether higher loss aversion increases or decreases the optimal tax rate is an empirical question. However, for our parameter estimates, all possible numerical estimates show that loss aversion reduces the optimal tax rate earmarked for green contributions. Hence, loss aversion is predicted to reduce contributions towards the green fund. We discuss this in the next example, which follows a parallel discussion in Example 1 above for the case of voluntary contributions.

Example 2 *From Proposition 2(a), we show that τ_i^* is decreasing in the parameter of loss aversion, λ, if $Y>\gamma\tau^{\gamma-1}\theta_T r_{t_3}$. This is true, even for the largest possible value of r_{t_3}, given the parameters that we used in our experiment ($\gamma=0.5$) and the estimated parameters from our data, as explained in Example 1, $\beta=1.0036$ and $\delta=0.9969<1$. From the first row of (2.11), $\theta_T=\beta\delta^{t_3}$ in treatment T3. Let us take t_3 to equal the conservative value of 5 weeks or 35 days in one of the cases in our experiments (increasing t_3 reduces the size of δ^{t_3} and reduces θ_T). The highest possible value of $r_{t_3}=Z$. A representative value is $Z=200$ in our experiments, and it is always the case that $Y=100$. Let us assume an income tax rate of 30% which is representative of most western democracies. Thus, we can check that for treatment T3*

$$Y=100>\gamma\tau^{\gamma-1}\theta_T(200)=(0.5)\,(0.3)^{0.5}\,(1.0036)(0.9969)^{35}(200)=49.3091,$$

which is comfortably satisfied. If we picked an even more conservative value of t_3 equals to 1 week, or 7 days, then we have $\gamma\tau^{\gamma-1}\theta_T(200)=53.7876<Y=100$, which also holds comfortably. In every possible simulation that we tried (including using the second row of (2.11)) for treatment T4, the inequality in (3.15) holds. Thus, we expect an increase in loss aversion to reduce contributions.

5 Experiments, data, and summary statistics

5.1 Brief summary of testable predictions

We briefly summarize our testable predictions. Green contributions are predicted to be:

1. Decreasing in the loss aversion parameter, λ (Propositions 1a, 2a, and Examples 1, 2).
2. Decreasing in the magnitude of the present-bias parameter, $1 - \beta$ (Propositions 1b, 2b).
3. Increasing in the size of the time t_3 endowment, Z (Propositions 1c, 2c).
4. Increasing when $t_3 - t_2$ is small but decreasing when $t_3 - t_1$ is large (Propositions 1d, 2d).
5. Higher in T2 relative to T1; and higher in T4 relative to T3 (Propositions 1e, 2e).

5.2 Experimental design

The experiment has 3 tasks in a within-subjects design. Task 1 and task 2 respectively elicit, for each subject i, their behavioral parameters of present-bias, β_i, and loss aversion, λ_i. In task 3, individuals decide their green contributions that are either voluntary or tax financed. We now provide more details.

1. In task 1, we elicit subject-specific time preferences using the Convex Time Budgets (CTB) method (Andreoni and Springer, 2012, Andreoni et al., 2015), which we explain in detail in Section 9.2.1 in the Appendix.

2. In task 2, we use the bisection method (Abdellaoui, 2000), to estimate the subject-specific loss aversion parameter for each individual.[34] Section 9.2.2 in the Appendix explains the details.

3. In task 3, subjects choose their green contributions and they are randomly assigned to one of the 4 treatments, T1, T2, T3, T4. In treatments T1 and T2, subjects were asked to declare their voluntary contributions to the green fund. In treatments T3 and T4, subjects were asked to declare their most preferred tax rate that finances mandatory contributions towards the green fund. The median tax rate was then implemented, and this was common knowledge. The experimental instructions (see the supplementary section) closely followed the sequence of moves described in the theoretical model for all 4 treatments.

We randomize the order of these tasks in the following way. In one block we have decisions that are made over time (task 1 and task 3) and in the second block, we have the elicitation of loss aversion, which is a static task (task 2). We randomize between the two blocks such that subjects always face task 1 before task 3 in the first block because we would like to elicit their deep underlying present-bias parameter prior to any context that is offered by the experiment.

The unit of currency, throughout our experiments, was Indian Rupees (INR). The average amount of money earned during the experiment was 734 INR and the participation fee was 100 INR.[35] The sessions lasted 37 minutes, on average, inclusive of the time for the instructions.

[34] Similar methods are used in Dhami et al. (2023a) and Dhami et al. (2023b) although they estimate the utility parameter differently.

[35] The exchange rate between the US dollar and the Indian rupee fluctuated over the time that the experiments were held, we may take it as approximately \$1 = 80 INR. Given that our sessions lasted only 37 minutes on average, total subject earnings were more than twice the hourly wage rate.

Subjects were assured of complete anonymity of their responses.

In experiments on temporal choices, it is absolutely critical that the future payments are made at the promised future date, and in a credible manner. We calculated the payments owed to each subject after the experiment and created an Excel file for payments. The payment was made through RazorPay, where CSBC has an institutional account.[36] CSBS pays subjects by sending them a link where subjects need to provide their UPI details[37] and then the payment goes through anonymously at the correct future date, promised in the experiment.

In order to test the comparative static effects of t_2, t_3, Z on contributions (see Section 5.1), we use the strategy method in task 3 within each treatment. We ask subjects to declare their voluntary contributions (treatments T1, T2) or choose their most preferred tax rates that finance mandatory contributions (treatments T3, T4) in a series of questions while we vary the parameters t_2, t_3, Z; see Table 2. In all treatments, and questions, the initial endowment was chosen as $Y = 100$ INR.

Table 2: Parameters used for the task 3 questions

Treatment	Q1	Q2	Q3	Q4	Q5	Q6
T1 and T3	$Z = 200$ $t_2 = -$ $t_3 = 5$ weeks	$Z = 200$ $t_2 = -$ $t_3 = 25$ weeks	$Z = 100$ $t_2 = -$ $t_3 = 5$ weeks	$Z = 400$ $t_2 = -$ $t_3 = 5$ weeks		
T2 and T4	$Z = 200$ $t_2 = 5$ weeks $t_3 = 10$ weeks	$Z = 200$ $t_2 = 5$ weeks $t_3 = 30$ weeks	$Z = 100$ $t_2 = 5$ weeks $t_3 = 10$ weeks	$Z = 400$ $t_2 = 5$ weeks $t_3 = 10$ weeks	$Z = 200$ $t_2 = 1$ weeks $t_3 = 10$ weeks	$Z = 200$ $t_2 = 9$ weeks $t_3 = 10$ weeks

Alternative values of t_2, t_3, Z used in 6 different questions, using the strategy method, to elicit contributions. In all cases, the initial, time t_1 endowment is $Y = 100$ INR.

Table 2 shows the six questions in our first set of experiments. The time t_3 endowment Z varied between 100–400; time t_2 varied between 1–9 weeks (and by definition, t_2 plays no role in treatments T1 and T3, so Q5 and Q6 are not relevant to these treatments); and time t_3 varied between 5–30 weeks.[38] We can now exploit the differences in contributions in the various questions in order to test some of our predictions. We give two examples.

1. For Q1 and Q2, we have $Z = 200$ in both questions. As we go from Q1 to Q2 for T1, we compare the effects of voluntary contributions for $t_3 = 5$ weeks and $t_3 = 25$ weeks. Similarly, as we go from Q1 to Q2 for T3, we compare the effects of mandatory tax-financed contributions for $t_3 = 10$ weeks and $t_3 = 30$ weeks. This allows us to test Proposition 1d(i) and Proposition 2d(i) on the effects of the time gap $t_3 - t_1$ (recall $t_1 \equiv 0$). Thus, we expect contributions to be lower in Q2 relative to Q1.

2. The contrast between Q3 and Q4 gives the effects of Z on voluntary contributions (and on tax-financed mandatory contributions for T3), as Z increases from 100 to 400, keeping fixed the respective dates, t_1, t_2 and t_3. This allows us to test Proposition 1c for the predicted increase in voluntary contributions for T1, and tax-financed contributions in T3.

[36] CSBC is the acronym for "Center For Social and Behavioral sciences" at Ashoka University in India, which conducted the experiments.

[37] The Unified Payments Interface (UPI) is a mobile-based, fast, payment system invented in India.

[38] The numbers were chosen in order to enable sensible comparisons. For instance, the time gap between the two consumption dates (t_1 and t_3 in T1 and T3; and t_2 and t_3 in T2 and T4) is 5 weeks for Q1 and 25 weeks for Q2 (compare both rows in Table 2) with Z held fixed at 200.

5.3 Data

The experiments were conducted with 515 students from 4 Indian Universities, over the period September 2022 to February 2023.[39] The sessions were conducted in classrooms at these universities. The experiment was conducted using Qualtrics and the recruitment took place via the SONA system. 421/515 subjects passed our comprehension test and we only included these 421 subjects in our analyses. Of these 421 subjects, 194 are from Ashoka University, 29 from the University of Lucknow, 108 from Indian Institute of Technology Madras and 90 from Lady Sriram College. These 421 subjects are randomly allocated across the four treatments. There are 105 subjects in treatments T1, T2 and T4 and 106 in treatment T3. Subjects are predominantly undergraduates (408/421), 59% are female and the average age is 21 years.

5.4 Descriptive statistics

Across all subjects, the mean of the loss aversion parameter is 2.03 with a median value of 1.55, which is consistent with earlier estimates (see discussion in the introduction).[40]

Table 3: Parameter estimates of time preferences for the pooled sample

Parameters	OLS	NLS
$\hat{\gamma}$	0.9452***	0.8225***
	(0.0011)	(0.0041)
$\hat{\delta}$	0.9958***	0.9969***
	(0.0001)	(0.0000)
$\hat{\beta}$	1.0055***	1.0036***
	(0.0091)	(0.0059)

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are in parentheses.

Using our estimation method for time preferences based on convex budget constraints, outlined in Section 9.2.2, Table 3 gives the parameter estimates of the following three parameters: $\hat{\gamma}$ (parameter of CRRA utility), $\hat{\delta}$ (discount factor in the exponential discounted utility model), and $\hat{\beta}$ (β parameter of the quasi-hyperbolic discount function). We find that $\hat{\beta} \approx 1$, however, we are interested in the variation of subject-specific β value around this estimated mean value, and the effect on contributions.

We now report the 'unconditional' statistical results. A conditional analysis, that controls for the effects of other variables, is carried out in Section 6 (and in Section 7 for the second set of experiments). Figure 1 shows a box and whiskers plot of contributions in each of the 4 treatments (T1–T4) and for the four questions (Q1–Q4; see Table 2). The median contributions are shown by a solid horizontal line and the mean by a dotted red line (the numeric values within each box are

[39] We have described this as the 'first set of experiments' in the introduction. Data were collected for a second set of experiments in order to test the effects of longer time horizons and to enable a cleaner test of the commitment device. The results from this second set of experiments are described separately in Section 7 below.

[40] There is no statistical difference in our two measures of loss aversion for two different values of income that we used in our experiments; see section 9.2.2 in the Appendix for the estimation details. Hence, loss aversion is independent of income levels, as assumed in the theory. This is rarely demonstrated in experimental results but it contributes towards validating our assumptions.

the mean values).

1. *Effect of a commitment device*: Recall that under voluntary contributions, treatment T2 offers a commitment device relative to treatment T1. Similarly, under mandatory tax financed contributions, treatment T4 offers a commitment device relative to treatment T3. Thus, we expect contributions to be higher in T2 relative to T1 (Proposition 1e) and in T4 relative to T3 (Proposition 2e). We introduce a slightly different, and direct, form of commitment in our second experiments, in Section 7, that is closer to the proposal in Thaler and Benartzi (2004).
   Figure 1 shows contributions made for each of these four questions across all treatments. The highest contributions are in treatment T2 for all questions, although none of the pairwise differences is statistically significant. Hence, it would appear the commitment device, of the form used in our first set of experiments, is more useful under a voluntary contributions mechanism as opposed to a tax-financed contributions mechanism.

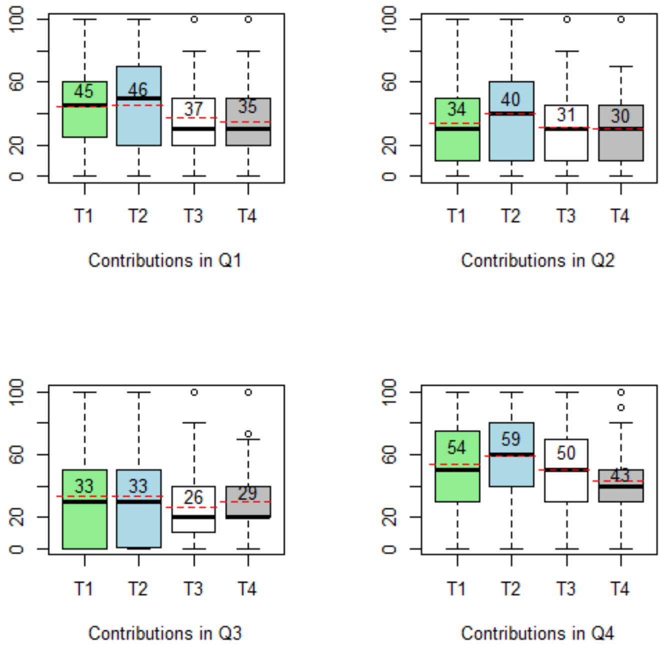

Figure 1: Contributions to the green fund across treatments and questions

2. *Institutional effects*: The contrasts T1 vs T3, and T2 vs T4 show the effect of voluntary versus tax-financed contributions on green contributions, holding fixed the pattern of endowments and discounting. In both comparisons, green investment is higher under voluntary contributions relative to tax financed contributions. Several of these contrasts, for the individual questions that compare mean contributions, reveal significantly higher contributions in the voluntary contributions mechanism. For instance: (i) Mean contributions are higher in T2 as compared to T4 for Q1 ($p = 0.0030$); Q2 ($p = 0.0057$); and Q4 ($p = 0.000$). (ii) Mean contributions are higher in T1 as compared to T3 for Q1 ($p = 0.0565$); and Q3 ($p = 0.0632$).

3. *Temporal effects-I* (size of $t_3 - t_1$): Our theory predicts that the greater is the time gap $t_3 - t_1$, the lower are contributions (Proposition 1d(i); Proposition 2d(i)). Hence, across all treatments, we expect to see higher contributions in Q1 as compared to Q2 (see Table 2). A two-sample Kolmogorov-Smirnov test indicates that there is a significant difference between contributions in Q1 and Q2 across all treatments ($p = 0.0025$). Figure 2 shows contributions in Q1 and Q2 in each treatment. Mean contributions are higher is Q1 as compared to Q2 within each treatment with all p-values less than 0.0000.

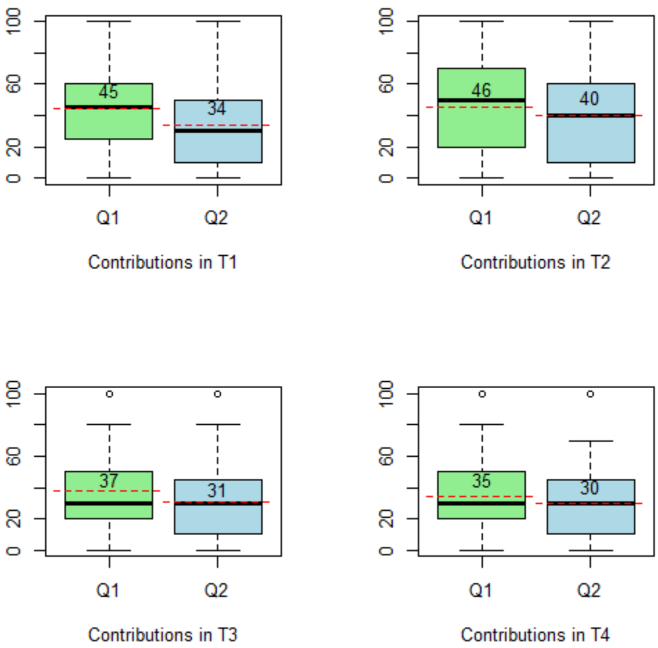

Figure 2: Contributions to the green fund in Q1 and Q2

4. *Temporal effects-II* (size of $t_3 - t_2$): Our theoretical model predicts that the smaller is the time gap $t_3 - t_2$, the greater are the contributions (Proposition 1d(ii); Proposition 2d(ii)). Time t_2 plays no role in treatments T1 and T3. Hence, in treatments T2 and T4, we expect to see higher contributions in Q6 where the time gap is smaller, as compared to Q5 (see Table 2). Figure 3 shows contributions in each treatment, indicating an almost null effect of increasing the time gap, $t_3 - t_2$, from 1 weeks to 9 weeks. However, one potential explanation is that the time horizon is not large enough to result in the predicted differences in contributions.

5. *Size of losses and contributions*: Our theoretical model predicts that contributions are increasing in Z, the size of time t_3 endowment, which is lost with a probability $1 - p$ at time t_3. We can test this prediction within each treatment by comparing contributions in Q3 and Q4 (see Table 2). From Figure 4, there are significantly higher contributions in Q4 where Z is higher, as compared to Q3, across all treatments. A two-sample Kolmogorov-Smirnov test indicates that there is a significant difference between Q4 and Q3 across all treatments ($p = 0.0000$). Hence, increasing the time t_3 endowment from 100 to 400 significantly increases

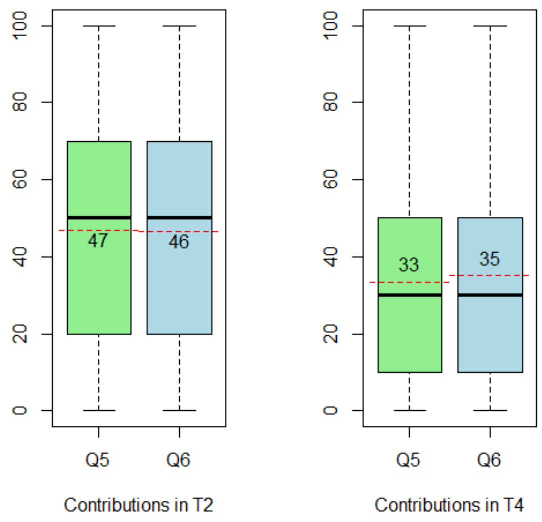

Figure 3: Contributions to the green fund in Q5 and Q6

contributions; the p-values from pairwise comparisons in each treatment are $p < 0.000$.

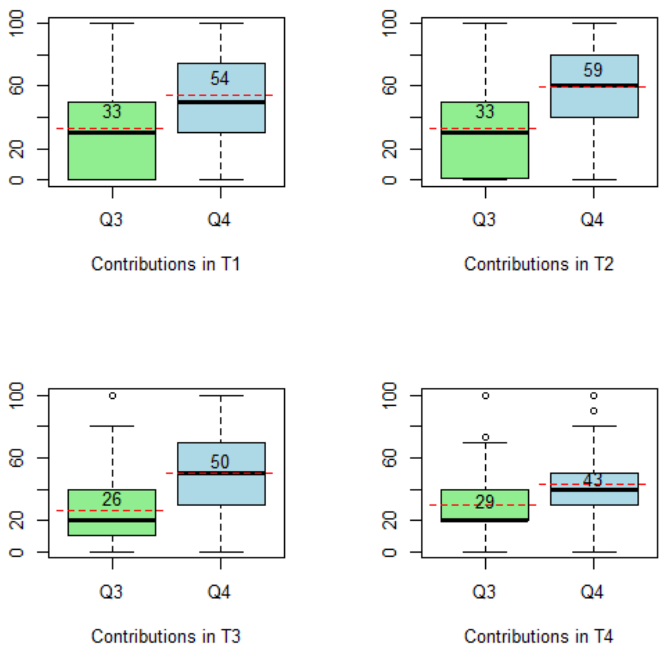

Figure 4: Contributions to the green fund in Q3 and Q4

6 Regression Results

In order to test our predictions (see Section 5.1 for a summary), we run OLS and Tobit regressions to examine the effect of behavioral (loss aversion, present-bias), structural (variation in endowments and time periods), and demographic variables (e.g., age, gender, marital status) on contributions.[41] The dependent variable is contributions towards green investment, in Indian Rupees, under either of two regimes– voluntary contribution mechanism (T1, T2) and tax-financed mandatory contributions (T3, T4). The details of the independent variables are as follows.

- 'Loss aversion': Mean of the two elicited measures of loss aversion corresponding to two different values of income $x = 100$ and $x = 400$ used in the lotteries in the elicitation procedure (see Section 9.2.2 for the details). 355/405 (88%) of subjects are loss averse.

- 'Present bias Mag': Magnitude of present-bias, $(1 - \beta)$. β is estimated through the CBT method (see Section 9.2.1 for the details). 205/405 (51%) of subjects are present biased, i.e., have $\beta < 1$.

- 'T' is a categorical variable for treatments. It takes the value 0 for the baseline treatment T1; 1 for T2; 2 for T3; and 3 for T4.

- 'Z' is a categorical variable for the time t_3 endowment. It takes the value 0 for the reference category when the endowment is 100; 1 for an endowment of 200; and 2 for an endowment of 400.

- 'Time' indicates the length of time taken for the completion of the experiment.

- 'House ownership' is a dummy variable that takes the value 0 if the subject's house is rented and 1 when the house is owned by the subject or other household members. 81/405 (20%) of subjects live in rented houses. This variable is a proxy for income and social status because, as expected, many subjects did not reveal their income in the SONA system.

- 'Gender' is a dummy variable for gender that takes the value 0 for male and 1 for female and others. 165/405 subjects (41%) are males and 229/405 subjects (57%) are females. 6 subjects identify themselves as non-binary; 1 as transgender; and 4 people prefer not to say.

- 'Religion' is a dummy variable for religion and takes the value 0 for non-Hindu subjects and 1 for Hindu subjects. 270/405 subjects (67%) identify with the Hindu religion.

- 'Marital' is a dummy variable for marital status and takes the value 1 for married status and 0 for others. 53/405 subjects (15%) are married.

- 'Age' gives the self-reported age of subjects. The mean and median age is, respectively, 20.66 and 20.16 years.

Table 4 reports the regression results from OLS regression and a Tobit specification, by pooling data from all questions (Q1–Q6) and all treatments (T1–T4). We could not estimate the behavioral parameters for 16 subjects as they made inconsistent choices in the time preferences exercise (task 2). Hence, we are left with data on 405 subjects. In Table 4, column 1 reports the results from an OLS regression and Column 2 reports results from a Tobit specification. Columns 3 and 4 replicate the same specifications as columns 1 and 2, respectively, while adding several control variables, such as gender, religion, and age.

In all four specifications in Table 4, the coefficient of loss aversion is negative and significant at the 5% significance level. Thus, contributions are decreasing in loss aversion, as predicted (Propositions 1a, 2a, and Examples 1, 2). For instance, from column 3, on average, for each unit increase in loss aversion, contribution to the green fund decreases by 1.536 units.

[41] In our experiments, for 12% of the choices we have zero contributions across all treatments and questions, while for 7% of the choices the entire endowment is contributed, which necessitates reporting Tobit regressions as well.

As predicted, contributions are increasing in the size of Z (Propositions 1c, 2c). As compared to the reference category of $Z = 100$, when the time t_3 endowment increases to 200, on average, contributions increase by 7.662 units. Similarly, when Z increase to 400, contributions increase by 20.744 units.

As compared to T1, on average, contributions are higher in T2 by 4.030 units; the result is consistent with our theoretical model (Proposition 1e). Hence, the commitment device under the voluntary mechanism is effective, although the coefficient is not statistically significant. We demonstrate a significant role for commitment in enhancing contributions in our new experiments reported in Section 7 with a longer time horizon, and a more direct form of commitment, as in Thaler and Benartzi (2004). Tax-financed contributions under the institutional mechanism are lower relative to the contributions under the voluntary contributions mechanism, when we take treatment T1 to be the baseline. This can be seen from the coefficients of T3 and T4 which are negative and significant at the 1% significance level. On average, the contributions are 5.495 units less in T3 as compared to T1.

The coefficient of the magnitude of present-bias, $1 - \beta$, is large, and negative as predicted (Proposition 1b, Proposition 2b), but it is not statistically significant. However, when we make the time horizon longer in our second set of experiments (see Section 7), then the effect of the present-bias parameter is significant, in addition being large.

Propositions 1d(i) and 2d(i) predict the effect of contributions as the time gap $t_3 - t_2$ varies. We have four different levels of the gap $t_3 - t_2$ (see Table 2), i.e., 1, 5, 9, 25 weeks. Due to multicollinearity issues we could not include the effect of time gaps in our combined regression in Table 4. Hence, in Table 5 we use data from treatments T2 and T4, the only treatments where t_2 is relevant, to explore the effect of variations in the time gap $t_3 - t_2$. We introduce a new categorical variable $t_3 - t_2$ where $t_3 - t_2 = 1$ week is the baseline and the other cases, $t_3 - t_2 = 5, 9, 25$ weeks are introduced as independent categories in Table 5. We find that only when the time gap $t_3 - t_2$ increases to 25 weeks relative to the reference category of 1 week, do contributions decrease statistically significantly, as predicted by our theory.

The other independent variables in Table 5 have similar signs and magnitudes to the ones obtained from regressions using data from all treatments in Table 4. Tax financed contributions produce a significantly lower level of the green fund relative to voluntary contributions (see coefficient on T4), and, on average, 12.558 units lower.

7 Results with a longer time horizon

In this section, we wish to explore the consequences of giving subjects choices over longer time horizons; this is our second set of experiments. We ran these new experiments in July-August 2023 with 103 subjects from Ashoka University. Only 79 subjects passed the first attention test and the rest were excused at the beginning of the experiment. Of these 79 subjects, 12 subjects did not pass the comprehension test, so they were dropped from the analyses. Another subject was dropped because of no variation in responses to the time preferences question. Hence, we have a final sample of 66 subjects.

Table 4: Full Regression Results - All Treatments

	Dependent variable: Contributions to the Green Fund			
	Without controls		With controls	
	(1)	(2)	(3)	(4)
Loss aversion	-1.612^{**}	-1.967^{**}	-1.536^{**}	-1.856^{**}
	(0.743)	(0.864)	(0.763)	(0.893)
Present Bias Mag	-7.447	-6.261	-6.930	-6.006
	(10.286)	(12.134)	(10.489)	(12.489)
T2	4.898	5.570	4.030	4.257
	(3.306)	(4.083)	(3.420)	(4.231)
T3	-5.231^{*}	-4.903	-5.495^{*}	-5.333
	(3.110)	(3.841)	(3.128)	(3.860)
T4	-7.521^{**}	-7.101^{*}	-8.273^{***}	-8.207^{**}
	(3.084)	(3.820)	(3.112)	(3.858)
$Z = 200$	7.662^{***}	9.611^{***}	7.662^{***}	9.631^{***}
	(0.949)	(1.250)	(0.951)	(1.255)
$Z = 400$	20.744^{***}	25.132^{***}	20.744^{***}	25.157^{***}
	(1.549)	(2.062)	(1.551)	(2.068)
Time			0.078	0.099
			(0.147)	(0.178)
House ownership			-0.169	0.922
			(2.488)	(3.098)
Gender			2.661	4.229
			(2.444)	(3.027)
Religion			1.885	2.595
			(2.311)	(2.859)
Marital			-0.197	-0.959
			(3.251)	(4.075)
Age			0.151	0.164
			(0.478)	(0.579)
logSigma		3.498^{***}		3.496^{***}
		(0.042)		(0.042)
Constant	36.323^{***}	33.474^{***}	29.656^{***}	24.200^{*}
	(2.933)	(3.682)	(11.309)	(13.776)

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the individual level (the level of randomization) and are reported in parentheses. Column 1 reports output from OLS regressions and Column 2 reports results using a Tobit specification. Columns 3 and 4 replicate the same specifications as columns 1 and 2, respectively, while incorporating control variables.

Table 5: Regression results using data from T2 and T4

	Dependent variable: *Contributions to the Green Fund*			
	(1)	(2)	(3)	(4)
Loss aversion	-2.612^{**}	-2.920^{**}	-2.424^{*}	-2.652^{*}
	(1.256)	(1.425)	(1.344)	(1.550)
Present Bias Mag	-8.810	-7.699	-8.078	-7.225
	(13.749)	(16.152)	(14.527)	(17.255)
T4	-12.742^{***}	-12.990^{***}	-12.558^{***}	-12.693^{***}
	(3.059)	(3.686)	(3.106)	(3.750)
$Z = 200$	7.230^{***}	8.623^{***}	8.090^{***}	9.456^{***}
	(1.335)	(1.685)	(1.556)	(1.913)
$Z = 400$	18.860^{***}	22.398^{***}	18.860^{***}	22.438^{***}
	(2.075)	(2.662)	(2.083)	(2.676)
$t_3 - t_2 = 5$			-0.066	-0.591
			(1.749)	(2.149)
$t_3 - t_2 = 9$			1.190	0.790
			(2.372)	(2.917)
$t_3 - t_2 = 25$			-4.829^{***}	-5.742^{***}
			(1.774)	(2.201)
Time			0.081	0.113
			(0.185)	(0.218)
House ownership			-1.862	-1.370
			(3.436)	(4.142)
Gender			4.148	6.499
			(3.576)	(4.414)
Religion			2.327	3.387
			(3.218)	(3.949)
Marital			-0.916	-2.375
			(4.245)	(5.262)
Age			-0.119	-0.113
			(0.739)	(0.890)
logSigma		3.475^{***}		3.467^{***}
		(0.058)		(0.057)
Constant	43.989^{***}	42.273^{***}	41.949^{**}	37.356^{*}
	(3.691)	(4.469)	(17.503)	(21.053)

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the individual level (the level of randomization) and are reported in parentheses. Column 1 reports output from OLS regressions and Column 2 reports results using a Tobit specification. Columns 3 and 4 replicate the same specifications as columns 1 and 2, respectively, while incorporating control variables.

In the new experiments, we are interested in the effects of a longer time horizon but not in the question of institutional vs voluntary contributions. Hence, we focus only on the voluntary contributions mechanism. We ran the baseline treatment, T1, and varied t_2, t_3 to accommodate longer time horizons in the first four questions in Table 2. We also added two extra questions to examine the effect of a commitment device, as in treatment T2. However, commitment takes a slightly different form to the first set of experiments and is closer in spirit to the proposal in Thaler and Benartzi (2004), as we explain below.

Table 6 summarizes the values of parameters for each question in the new experiments. Q1–Q4 are relevant for treatment T1, where t_2 plays no role, and Q5–Q6 are relevant for treatment T2, where $t_2 > 0$. In these questions, we picked t_3 to be between 6 months to 12 months; by contrast, the time horizon in the first set of experiments was shorter at 5–10 weeks for most questions and for just one question it was 25–30 weeks. All 66 subjects answered all six questions.

In order to study the effects of commitment, we constructed the new questions Q5 and Q6 (see Table 6) along similar lines to the SMarT pension plan of Thaler and Benartzi (2004). This offers a slight modification to our first set of experiments in the following way. In Q5, we informed subjects at time t_1 that they will receive 100 + 50 Rupees in 4 months (this is time t_2). Subjects were told that at time t_1 their contribution to the green fund from the first sum of 100 Rupees to be received in 4 months from now, is the same amount that they declared in Q2 (where t_2 was not relevant). Suppose that a subject had chosen to invest $x\%$ from 100 in Q2 (the value of $x\%$ would vary from subject to subject).

We then asked subjects to decide how much they would like to commit today at time t_1 to contribute from the extra 50 Rupees that they receive in 4 months time from now, towards the green investment, when t_3 is 1 year from present. Subjects had 5 options to choose from in contributing from this additional 50 Rupees.

(a) $x\%$ of 50 Rupees.

(b) $x\% + 3\%$ of 50 Rupees.

(c) $x\% + 10\%$ of 50 Rupees.

(d) $x\% + 15\%$ of 50 Rupees.

(e) less than $x\%$ of 50 Rupees.

The advantage of this design is that it cleanly gives us a within-subjects comparison of the voluntary contributions decisions without commitment and with commitment. Clearly, if the commitment device has no value, then subjects should choose option (a) and if the commitment device is valuable, as in Thaler and Benartzi (2004), then we would expect subjects to choose one of the options (b), (c), or (d). In Q5, the gap $t_3 - t_2$ is 8 months, while in Q6, this gap is 2 months, keeping fixed t_3 equal to 1 year; otherwise Q6 gives the same options as Q5. As such, this gives us a clear method of studying the demand for commitment devices.

Out of 66 subjects, 30 subjects (45%) declared that they would contribute the same percentage from the extra 50 Rupees in Q5 (option (a)), while 22 subjects (36%) were willing to contribute 15% more (option (d)), 10 subjects were willing to contribute 10% more (option (c)), 2 subjects chose 3% more (option (b)), and 2 subjects chose to contribute less (option (e)). Similar figures are obtained in Q6, where 30 subjects declared they would contribute the same percentage from the

Table 6: Questions in the second set of experiments

Q1	Q2	Q3	Q4	Q5	Q6
$Z = 200$	$Z = 200$	$Z = 100$	$Z = 400$	$Z = 200$	$Z = 200$
$Y = 100$	$Y = 100$	$Y = 100$	$Y = 100$	$Y = 100 + 50$	$Y = 100 + 50$
$t_2 = -$	$t_2 = -$	$t_2 = -$	$t_2 = -$	$t_2 = 4$ months	$t_2 = 10$ months
$t_3 = 6$ months	$t_3 = 1$ year	$t_3 = 6$ months	$t_3 = 6$ months	$t_3 = 1$ year	$t_3 = 1$ year

extra 50 Rupees, 19 people chose 15% more, 6 and 5 subjects chose 10% and 3% more contribution, respectively, and 6 subjects chose to declare less. A two-sample Kolmogorov-Smirnov test indicates that there is not a significant difference between the contributions distributions in Q5 and Q6 ($p = 0.9997$). Thus, in a nutshell, about 50% of the subjects choose to contribute more when they had access to a commitment technology. These results point qualitatively in the same direction as those in Thaler and Benartzi (2004), albeit in a different context.

Figure 5 shows contributions to the green fund across the first 4 questions, in the new experiments. There is a significant difference between contributions in Q1 and Q2 ($p = 0.0063$). Consistent with our theory, as the gap between the terminal date (t_3) and the decision date (t_1) increases, contributions to the green fund decrease (Propositions 1d(i) and 2d(i)). Also, as predicted by our theory, when the size of endowment, Z, at time t_3 increases, we observe higher contributions ($p = 0.0000$) (Propositions 1c and 2c).

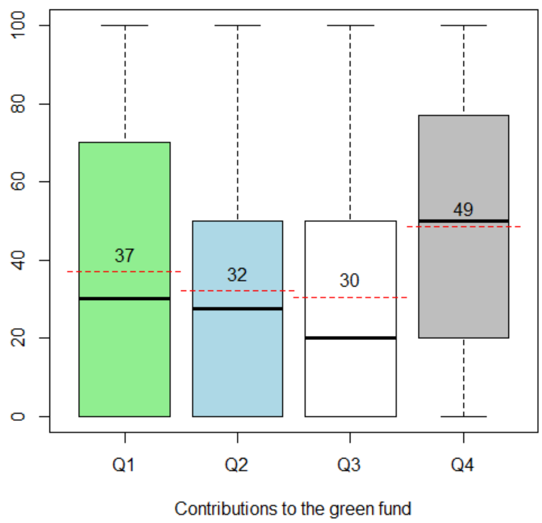

Figure 5: Contributions to the Green Fund in the new, second set of, experiments

Table 7 reports the regression results for these longer-horizon experiments; we report OLS and Tobit results in two different columns. Contributions to the green fund are decreasing in both behavioural parameters, loss aversion and the magnitude of present-bias, as predicted by our theory (Propositions 1a,b and 2a,b). Both coefficients are significant at the 5% significance level. The effect of present-bias is more pronounced compared to the impact of loss aversion. Recall that in the first set of experiments (Section 6) with a shorter time horizon, the effect of present-bias was not statistically significant; but with a longer horizon the effect is large, negative (as predicted by our model) and significant. In the OLS regression, a 1 unit increase in loss aversion, decreases contributions by 3.591 units, but a 1 unit increase in the magnitude of the present-bias parameter decreases contributions by 25.507 units. The effects are even larger in the Tobit regression.

As compared to the reference category of $Z = 100$, when the endowment at time t_3 increases to $Z = 400$, contributions increase by 18.076 units and this is significant at the 1% level. Of the control variables, only time is significant; subjects who took a longer time to deliberate their actions, contributed significantly higher amounts. The variable $t_3 = 1$ year is a dummy variable that takes a value 0 when t_3 is 6 months and a value 1 when it is 1 year. An increase in t_3 reduces contributions as predicted (Propositions 1c, 2c) and the effect is statistically significant. Thus, overall, the empirical results are in good conformity with our theoretical predictions, although some of the results attain statistical significance with a longer horizon.

8 Conclusions

The challenges posed by climate change require a multi-disciplinary approach. Traditional economic theory has already demonstrated the power of economic incentives and regulation in influencing the behavior of consumers. In order to address an important gap in the literature, we focus on how some of the core components of behavioral economics can be leveraged to analyze the underlying determinants of green contributions. The decision makers in our paper can be consumers, households, firms, region, or countries. Our model incorporates the following key components: The temporal and risk dimensions of the problem, which rely on time and risk preferences; the public goods nature of green contributions; the probabilistic nature of climate change abatement; and a comparison of alternative institutions such as voluntary versus mandatory contributions towards the green fund.

We first construct a rigorous theoretical model that incorporates these key components to derive the relevant predictions, and then we stringently test them with the data. The experiments closely implement all the details of the theoretical model; the predictions of the model came first, followed by the experiments, i.e., we are not interested in a just-so theoretical model.

We find that loss aversion and present-bias, which are both key behavioral attributes of human and primate preferences, reduce green contributions significantly. Commitment devices are valuable in increasing contributions and when available in the form suggested in Thaler and Benartzi (2004), such devices are chosen by about half of all subjects. We also demonstrate the role played by structural factors such as how immediate is the threat of climate change and the time gap between exercising commitment and climate change. Voluntary contributions elicit higher contributions than mandatory tax-financed contributions, echoing the findings of earlier work by Ostrom (1990) on the superiority of private solutions to manage the commons relative to formal incentives and regulation. We test for the effects of long and short time horizons. Some of our predictions attain statistical significance when the time horizon is large enough, say, a year.

We demonstrate the potentially rich insights offered by behavioral economics through several key and novel insights in terms of risk and time preferences, as well as in the domain of public

Table 7: Regression results in the new, second set of, experiments

	Dependent variable:	
	Contributions to the Green Fund	
	OLS	*Tobit*
	(1)	(2)
Loss aversion	-3.591^{**}	-5.747^{**}
	(1.607)	(2.750)
Present Bias Mag	-25.507^{**}	-41.983^{*}
	(11.991)	(23.958)
$t_3 = 1$	-4.879^{***}	-7.000^{**}
	(1.768)	(2.871)
$Z = 200$	6.470^{**}	9.975^{*}
	(3.119)	(5.187)
$Z = 400$	18.076^{***}	26.879^{***}
	(3.633)	(6.077)
House Ownership	-1.303	-2.216
	(7.723)	(11.725)
Time	1.377^{**}	2.195^{**}
	(0.651)	(0.986)
Gender	9.492	16.930
	(7.360)	(11.214)
Religion	-2.652	-3.729
	(7.150)	(10.964)
Marital	4.928	5.207
	(9.959)	(15.477)
Age	0.870	1.062
	(0.688)	(1.151)
logSigma		3.804^{***}
		(0.117)
Constant	0.815	-22.045
	(18.047)	(28.999)
Observations		264
Log Likelihood		-951.502
Akaike Inf. Crit.		1,929.005
Bayesian Inf. Crit.		1,975.492

Note: *p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered at the individual level (the level of randomization) and are reported in parentheses. Column 1 reports output from OLS regression. Column 2 reports results using a Tobit specification.

goods. Our paper also pushes the case for applying the contributions of behavioral economics more broadly as compared to an exclusive reliance on the classical nudge type interventions, whose effectiveness has already been demonstrated.

9 Appendix:

9.1 Proofs

Proof of Proposition 1: From (3.13), we know that $\frac{\partial^2 U}{\partial g_i^2} < 0$. Using (3.16) and the implicit function theorem, we have

(a)

$$\frac{\partial g^*}{\partial \lambda} = -\left(-\frac{\partial^2 U}{\partial g_i^2}\right)^{-1} \mu \left[1 - \theta_T r_{t_3} p'(G)\right] \gtreqless 0.$$

(b) For treatment T1, from (2.11), we have $\theta_T = \beta\delta^{t_3} > 0$. Hence,

$$\frac{\partial g^*}{\partial \beta} = \left(-\frac{\partial^2 U}{\partial g_i^2}\right)^{-1} \delta^{t_3} p'(ng^*) \left[(u(Z) + \mu Z - \mu r_{t_3}) + \mu r_{t_3}\lambda\right] > 0. \tag{9.1}$$

In signing (9.1) we have used the fact that the interior solution requires the term in the square brackets on the RHS to be positive. It follows that $\frac{\partial g^*}{\partial(1-\beta)} < 0$. In treatment T2, by contrast, from (2.11), $\theta_T = \delta^{t_3 - t_2}$, hence, the RHS of (3.16) is independent of β, so $\frac{\partial g^*}{\partial(1-\beta)} = 0$.

(c)

$$\frac{\partial g^*}{\partial Z} = \left(-\frac{\partial^2 U}{\partial g_i^2}\right)^{-1} \theta_T p'(ng^*) \left(u'(Z) + \mu\right) > 0.$$

(di)

$$\frac{\partial g^*}{\partial t_3} = \left(-\frac{\partial^2 U}{\partial g_i^2}\right)^{-1} p'(ng^*) \left[(u(Z) + \mu Z - \mu r_{t_3}) + \mu r_{t_3}\lambda\right] \frac{d\theta_T}{dt_3} < 0. \tag{9.2}$$

The sign in (9.2) follows for the following reason. Using the first row of (2.11), if $0 < \delta < 1$, then $\frac{d\theta_T}{dt_3} = \beta\delta^{t_3} ln\delta < 0$; and from the second row of (2.11), $\frac{d\theta_T}{dt_3} = \delta^{t_3 - t_2} ln\delta < 0$ (otherwise, if $\delta = 1$ then $\frac{d\theta_T}{dt_3} = 0$). Thus, for both treatments, an increase in t_3 reduces optimal investment, g^*.

(dii) In treatment T1, from (2.11), $\theta_T = \beta\delta^{t_3}$, so the RHS of (3.16) is independent of t_2, hence $\frac{\partial g^*}{\partial t_2} = 0$. From (2.11), for treatment T2, $\theta_T = \delta^{t_3 - t_2}$. So, if $0 < \delta < 1$, we have $\frac{d\theta_T}{dt_2} = -\delta^{t_3 - t_2} ln\delta > 0$, hence

$$\frac{\partial g^*}{\partial t_2} = -\left(-\frac{\partial^2 U}{\partial g_i^2}\right)^{-1} p'(ng^*) \left[(u(Z) + \mu Z - \mu r_{t_3}) + \mu r_{t_3}\lambda\right] \delta^{t_3 - t_2} ln\delta > 0,$$

where the sign follows by using $ln\delta < 0$ when $0 < \delta < 1$. When $\delta = 1$, we get the $\frac{\partial g^*}{\partial t_2} = 0$ because $ln1 = 0$.

(e) From Remark 1, we know that $\beta\delta^{t_3} \leq \delta^{t_3 - t_2}$ (and with strict inequality if $\delta < 1$), thus, the weight placed on the time t_3 (positive) payoff in the first order condition (3.16) is lower in treatment T1 relative to treatment T2. Since the first order condition is sufficient, it follows that the contributions in Treatment T1 are relatively lower. ■

Proof of Lemma 1: The objective function (4.1) is continuous and strictly concave in τ_i (see (4.3)); and τ_i belongs to the closed and bounded interval $[0, 1]$. Hence, τ_i^* exists and is unique. Since the first term on the RHS of (4.4) is strictly negative, the second term must be strictly positive for an interior solution. ■

Proof of Lemma 2: We sketch the proof. From Lemma 1, preferences of each voter are single peaked in the tax rate. From (4.4), the first order condition is monotonic in λ_i, hence, the optimal tax rate is monotonic in λ_i (either strictly increasing or strictly decreasing). Indeed, we show that it is strictly decreasing for all reasonable parameter estimates (Example 2). It follows from the median voter theorem that there is a unique solution to the majority voting problem. We have a family of single peaked preferences over the tax rate that are monotonically ordered, across the voters, with respect to the parameter of loss aversion, λ. Ordering the most preferred tax rates of the n voters as $\tau_1^* < .... < \tau_n^*$, the majority voting solution is found to be the median value of these tax rates, denoted by τ_M^*, and this is also the Condorcet winner. ■

Proof of Proposition : Using (4.3), (4.4), and the implicit function theorem, we have

(a)

$$\frac{\partial \tau_i^*}{\partial \lambda} = \left(-\frac{\partial^2 U}{\partial \tau_i^2}\right)^{-1} \mu \left[-Y + \gamma \tau^{\gamma-1} \theta_T r_{t_3}\right] \gtreqless 0.$$

(b) For treatment T3, from (2.11), we have $\theta_T = \beta \delta^{t_3}$, hence,

$$\frac{\partial \tau_i^*}{\partial \beta} = \left(-\frac{\partial^2 U}{\partial \tau_i^2}\right)^{-1} \delta^{t_3} \gamma \tau^{\gamma-1} \left[(u(Z) + \mu Z - \mu r_{t_3}) + \mu r_{t_3} \lambda_i\right] > 0. \tag{9.3}$$

In signing (9.3) we have used the fact that the interior solution requires the term in the square brackets on the RHS to be positive (see Lemma 1). It follows that $\frac{\partial \tau_i^*}{\partial(1-\beta)} < 0$. In treatment T4, by contrast, from (2.11), $\theta_T = \delta^{t_3 - t_2}$, hence, the RHS of (4.3) is independent of β, so in treatment T4, $\frac{\partial \tau_i^*}{\partial(1-\beta)} = 0$.

(c)

$$\frac{\partial \tau_i^*}{\partial Z} = \left(-\frac{\partial^2 U}{\partial \tau_i^2}\right)^{-1} \gamma \tau^{\gamma-1} \theta_T \left(u'(Z) + \mu\right) > 0.$$

(di)

$$\frac{\partial \tau_i^*}{\partial t_3} = \left(-\frac{\partial^2 U}{\partial \tau_i^2}\right)^{-1} \gamma \tau^{\gamma-1} \left[(u(Z) + \mu Z - \mu r_{t_3}) + \mu r_{t_3} \lambda_i\right] \frac{d\theta_T}{dt_3} < 0. \tag{9.4}$$

The sign in (9.4) follows due to the following reason. Using the first row of (2.11), if $0 < \delta < 1$, then $\frac{d\theta_T}{dt_3} = \beta \delta^{t_3} ln\delta < 0$; and from the second row of (2.11), $\frac{d\theta_T}{dt_3} = \delta^{t_3 - t_2} ln\delta < 0$ if $0 < \delta < 1$ (otherwise, when $\delta = 1$, $\frac{d\theta_T}{dt_3} = 0$, because $ln1 = 0$). Furthermore, the interior solution guarantees that the term in the square brackets on the RHS of (9.4) is positive (see Lemma 1). Thus, for both treatments, an increase in t_3 reduces the most preferred tax rate.

(dii) In treatment T3, from (2.11), $\theta_T = \beta \delta^{t_3}$, so the RHS of (4.4) is independent of t_2, hence $\frac{\partial \tau_i^*}{\partial t_2} = 0$. From (2.11), for treatment T4, $\theta_T = \delta^{t_3 - t_2}$, so $\frac{d\theta_T}{dt_2} = -\delta^{t_3 - t_2} ln\delta > 0$ if $\delta < 1$ (and $\frac{d\theta_T}{dt_2} = 0$ if $\delta = 1$). Hence, if $\delta < 1$, then

$$\frac{\partial \tau_i^*}{\partial t_2} = -\left(-\frac{\partial^2 U}{\partial \tau_i^2}\right)^{-1} \gamma \tau^{\gamma-1} \left[(u(Z) + \mu Z - \mu r_{t_3}) + \mu r_{t_3} \lambda_i\right] \delta^{t_3 - t_2} ln\delta > 0.$$

This completes the proof.

(e) The proof is similar to the proof of Proposition 1e, hence, we omit it. ■

9.2 Estimation methods for the behavioral parameters

9.2.1 Convex Time Budgets

We use the method of Convex Time Budgets (CTB) of Andreoni et al. (2015) to estimate time preferences. Consider two time periods t ('sooner') and $t+k$ ('later') with $k>0$. A linear budget set of allocations of monetary rewards to be received at these two times is a line connecting the two points $(x_t, 0)$ and $(0, x_{t+k})$ in a two-dimensional plane. The first point corresponds to receiving a certain amount x_t at time t and nothing at $t+k$. The second point corresponds to receiving a certain amount x_{t+k} at time $t+k$ and nothing at t. Any points on the interior of a budget set represent allocations where the subject receives positive rewards at both dates.

The slope of the budget line represents the intertemporal tradeoff between rewards at two different time periods. In order to identify and estimate the parameters of time preferences, we needs to vary the time periods $(t, t+k)$, the slopes of the budget lines, and the level of the budget lines. Each budget line can be expressed as a set of these numbers. Andreoni et al. (2015) implement the CTB protocol by asking subjects to select a reward schedule (x_t, x_{t+k}) from a set of 6 options that are evenly spaced on the budget line. We follow the same procedure.

Consider quasi-hyperbolic discounting with an instantaneous utility of the constant relative risk aversion (CRRA) form:

$$U(x_t, x_{t+k}) = x_t^{\gamma} + \beta^{1_{t=0}} \delta^k x_{t+k}^{\gamma}, \tag{9.5}$$

where δ is the per-period discount factor; γ is the curvature parameter related to risk aversion; and β is the present bias parameter; the superscript $1_{t=0}$ is an indicator variable to capture the following cases (i) where $t=0$ (the current date), $\beta^{1_{t=0}} = \beta$, and (ii) when $t>0$, $\beta^{1_{t=0}} = 1$. The intertemporal budget constraint is given by

$$x_t + \frac{x_{t+k}}{1+r} = I, \tag{9.6}$$

where $1+r$ is the gross interest rate and I is the time t income available to be allocated to the consumption pair (x_t, x_{t+k}). Maximizing (9.5) subject to (9.6) gives rise to the following intertemporal Euler equation:

$$\frac{x_t}{x_{t+k}} = \left(\beta^{1_{t=0}} \delta^k (1+r)\right)^{\frac{1}{\gamma-1}}. \tag{9.7}$$

Andreoni et al. (2015) use different methods for estimating the three parameters β, δ, γ. The first method uses OLS and estimates the parameters in the log-linearized version of Euler equation:

$$\log\left(\frac{x_t}{x_{t+k}}\right) = \frac{\log \beta^{1_{t=0}}}{\gamma-1} + \frac{\log \delta}{\gamma-1} k + \frac{1}{\gamma-1}\log(1+r). \tag{9.8}$$

Under an additive error structure, the underlying preference parameters are recovered via a non-linear combination of the estimated coefficients. Rewriting (9.8) we have

$$\log\left(\frac{x_t}{x_{t+k}}\right) = \gamma_1 + \gamma_2 k + \gamma_3 \log(1+r). \tag{9.9}$$

Thus, the estimates of the three preference parameters (hats over the variables denote estimates) are given by

$$\hat{\gamma} = \frac{1}{\hat{\gamma}_3} + 1, \quad \hat{\delta} = \exp\left(\frac{\hat{\gamma}_2}{\hat{\gamma}_3}\right), \quad \hat{\beta} = \exp\left(\frac{\hat{\gamma}_1}{\hat{\gamma}_3}\right). \tag{9.10}$$

The model in (9.8) can be estimated by OLS. Andreoni et al. (2015) note that the allocation ratio, $\log\left(\frac{x_t}{x_{t+k}}\right)$, is not well defined at corner solutions. To address this issue, one can use the demand function to generate a non-linear regression equation based upon

$$x_t = \frac{I(\beta^{1_{t=0}}\delta^k(1+r))^{\frac{1}{\gamma-1}}}{1 + ((1+r)(\beta^{1_{t=0}}\delta^k(1+r))^{\frac{1}{\gamma-1}}}. \tag{9.11}$$

In our experiment we used non-linear regression to estimate time preferences. For non-convergent cases, we used OLS estimates. Table 5 in the Supplementary Section shows the budget sets that we used to estimate the preference parameters.

9.2.2 Loss aversion parameter

Task 1 gives us an estimate of the parameter of the CRRA utility function; see Section 9.2.1. In task 2 we elicit certainty equivalents of two lotteries in order to estimate the loss aversion parameter. We use the bisection procedure with 6 steps to find the value of an outcome $l > 0$ such that the subject expresses the following indifference, given a predetermined value of $x > 0$:[42]

$$L \equiv (-l, 0.5; x, 0.5) \sim (0, 1). \tag{9.12}$$

The lottery on the LHS, L, gives a 50–50 chance of gaining x or losing l. In our experiment, we use two different values of $x \in [100, 500]$ and we compute the loss aversion for each value. We take the average loss aversion across these two values as the final measure that we use for our empirical analysis.[43] The lottery on the RHS of (9.12) is a value of 0 with certainty. Starting with the lottery $(0, 1)$, and given a value for x, we elicit the value of l that will make subjects indifferent to the lottery, L. We take the status-quo value (0 received with probability 1) as the reference point; indeed the status-quo typically provides a satisfactory approximation to the reference point (Kahneman and Tversky, 2000; Dhami 2019, Vol. 1).

Consider the standard utility function under prospect theory, with a reference point of 0[44]

$$v(x) = \begin{cases} x^\gamma & \text{if } x \geq 0 \\ -\lambda(-x)^\gamma & \text{if } x < 0 \end{cases}. \tag{9.13}$$

In (9.13), the parameter $\gamma \in (0, 1)$ captures the curvature of the utility function and empirical estimates indicate that γ is close to 1 (Dhami, 2019, Vol. 1).[45] The parameter λ is the parameter

[42] This method draws on Abdellaoui (2000), Dhami et al. (2023a) and Dhami et al. (2023b).

[43] In theory, the parameter of loss aversion is independent of the level of income, but our method takes account of the possibility of such dependence and, hence, creates a more robust measure.

[44] For the rationale for such a utility function, its empirical basis, and its axiomatic foundations, see Dhami (2019, Vol. 1)

[45] In principle, one could introduce different power parameters for gains and losses and estimate them separately. However, the empirical evidence shows that these power parameters are approximately identical (Dhami, 2019, Vol. 1). A similar comment applies to the probability weighting function, which we also take to be identical in gains and losses.

of loss aversion; $\lambda > 1$ indicates loss aversion and $0 < \lambda < 1$ indicates loss tolerance. The classical studies on loss aversion suggest that the median value of loss aversion is $\lambda \approx 2.25$ (Kahneman and Tversky, 1979, 2000; Tversky and Kahneman, 1992) and we review some of the more recent estimates in the introduction. The prospect theory evaluation of the two lotteries in (9.12) is[46]

$$PT(L) = w(0.5)v(-l) + w(0.5)v(x);\ PT(0,1) = v(0) = 0. \tag{9.14}$$

Using (9.13), (9.14), the indifference $L = (-l, 0.5; x, 0.5) \sim 0$ implies that for a subject who uses prospect theory, $PT(L) = PT(0,1)$, or $-w(0.5)\lambda l^{\gamma} + w(0.5)x^{\gamma} = 0$. Rearranging this expression, we have $\lambda = \frac{w(0.5)}{w(0.5)}\left(\frac{x}{l}\right)^{\gamma}$, or

$$\lambda = \left(\frac{x}{l}\right)^{\gamma},$$

which gives the required estimate of loss aversion.

References

[1] Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. Management Science 46(11): 1497-1512.

[2] Allcott, H. (2011a). Rethinking real-time electricity pricing. Resource and Energy Economics 33 (4): 820–842.

[3] Allcott, H. (2011b). Social norms and energy conservation," Journal of Public Economics 95 (9): 1082– 1095.

[4] Andreoni, J., Kuhn, M. A., and Sprenger, C. (2015). Measuring time preferences: A comparison of experimental methods. Journal of Economic Behavior & Organization 116: 451–464

[5] Andreoni, J., Sprenger, C. (2012). Estimating time preferences from convex budgets. American Economic Review 102 (7): 3333–3356.

[6] Barrett, S., Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. Proceedings of the National Academy of Sciences 109, 17372e17376.

[7] Brown, A. L., Imai, T., Vieider, F., and Camerer, C. F. (2023). Meta-Analysis of Empirical Estimates of Loss-Aversion. Forthcoming in Journal of Economic Literature.

[8] Calzolari, G., Casari, M. and Ghidoni, R. (2018) Carbon is forever: A climate change experiment on cooperation. Journal of Environmental Economics and Management 92: 169–184

[9] Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., Tavoni, A., (2015). On the provision of public goods with probabilistic and ambiguous thresholds. Environmental Resource Economics 61(3), 365–383.

[10] Dengler, S., Gerlagh, R., Trautmann, S. T., and van de Kuilen, G. (2018). Climate Policy Commitment Devices. Journal of Environmental Economics and Management 92: 331-343

[46] For the standard methods in applying prospect theory, see Kahneman and Tvesky (2000) and Dhami (2019, Vol. 1).

[11] Dhami, S. (2019). The foundations of behavioral economic analysis: Behavioral Decision Theory. Volume 1. Oxford University Press: Oxford.

[12] Dhami, S. (2019). The foundations of behavioral economic analysis: Behavioral Time Discounting. Volume 3. Oxford University Press: Oxford.

[13] Dhami, S. (2023). Green Technology Adoption, Complexity, and the Role of Public Policy: A Simple Theoretical Model. CESifo Working Paper No. 10364, Available at SSRN: https://ssrn.com/abstract=4416793 or http://dx.doi.org/10.2139/ssrn.4416793.

[14] Dhami, S., Hajimoladarvish, N., and Konstantinos, G. (2023a) Precautionary Savings, Loss Aversion, and Risk: Theory and Evidence. CESifo Working Paper No. 10570.

[15] Dhami, S., Hajimoladarvish, N., Mamidi, P. (2023b) Loss Aversion and Tax Evasion: Theory and Evidence CESifo Working Paper 10224.

[16] Dhami, S., Wei, M., and al-Nowaihi, A. (2023c) Classical and Belief-Based Gift Exchange Models: Theory and Evidence. Games and Economic Behavior 138: 171-196.

[17] Dietz, T., Fitzgerald, A., Shwom, R. (2005). Environmental values. Annual Review of Environmental Resources 30(1), 335–372.

[18] Gachter, S., Johnson, E. J., and Herrmann, A. (2022). Individual-level loss aversion in riskless and risky choices. Theory and Decision 92(3-4): 599-624.

[19] Gerlagh, R. and Liski, M. (2018). Consistent climate policies. Journal of the European Economic Association, European Economic Association 16(1): 1-44

[20] Ghidoni, R, Clazolari, G., Casari, M. (2017) Climate change: Behavioral responses from extreme events and delayed damages. Energy Economics 68: 103–115

[21] Gustafsson M, Biel A, Gärling T (2000) Egoism bias in social dilemmas with resource uncertainty. Group Processes & Intergroup Relations 3(4): 351–365.

[22] Heinz, N. and Koessler, A-K. (2021) Other-regarding preferences and pro-environmental behaviour: An interdisciplinary review of experimental studies. Ecological Economics 184: 106987.

[23] Imai, T. Rutter, T. A., Camerer, C. F. (2021) Meta-Analysis of Present-Bias Estimation using Convex Time Budgets. The Economic Journal 131(636): 1788–181

[24] IPCC (2014). Climate change 2014: mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

[25] Jessoe, K., and Rapson, D. (2014). Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use. American Economic Review 104(4): 1417–38.

[26] Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. Econometrica 47(2): 263–91.

[27] Kahneman, D., Tversky, A. (Eds.). (2000). Choices, values and frames. Cambridge: Cambridge University Press.

[28] Milinski, M., Sommerfeld, R.D., Krambeck, H.-J., Reed, F.A., Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. Proceedings of the National Academy of Sciences U. S. A. 105(7): 2291–2294.

[29] Nordhaus, W. D., and Boyer, J. G. (2000). Warming the World: Economic Models of Global Warming. Cambridge, MA: MIT Press.

[30] Ostrom, E. (1990) Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press: Cambridge.

[31] Rapoport A, Budescu DV, Suleiman R, Weg E (1992) Social dilemmas with uniformly distributed resources. In: Liebrand WBG, Messick DM, Wilke HAM (eds) Social dilemmas: theoretical issues and research findings. Pergamon Press, Oxford, pp. 43–57.

[32] Schultz, W., Zelezny, L. (1999). Values as predictors of environmental attitudes: evidence for consistendy across 14 countries. Journal of Environmental Psychology 19(3), 255–265.

[33] Schwartz, D., Fischhoff, B., Krishnamurti, T., Sowell, F. (2013). The Hawthorne effect and energy awareness. Proceedings of the National Academy of Sciences of the United States of America. 110. 10.1073/pnas.1301687110.

[34] Stern, N. (2007). The Economics of Climate Change: The Stern Review. Cambridge, UK: Cambridge University Press.

[35] Stern, N. (2008). The Economics of Climate Change. American Economic Review 98(2): 1-37.

[36] Stern, P.C., (2000). Toward a coherent theory of environmentally significant behavior. Journal of Social Issues 56 (3), 407–424. https://doi.org/10.1111/0022-4537.00175.

[37] Stern, N. (2022). A time for action on climate change and a time for a change in economics. The Economic Journal 132(644): 1259–1289.

[38] Thaler, R. H. and Benartzi, S. (2004). Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving Journal of Political Economy 112(S1): S164–S187.

[39] Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. Journal of Risk and Uncertainty 5(4): 297–323.