* Preliminary version. Please do not circulate.

Are the more vulnerable doomed? Insights from an Experimental Threshold Public Goods game[a]

Sigma Samhita[b]

December 3, 2022

Abstract

Climate change disparately affects different parts of the Earth, with catastrophic consequences in some regions. Coordinated global mitigation efforts can avert this crisis. Nations which are more vulnerable to the catastrophic effects of climate change require substantially higher global mitigation efforts, compared to nations which are less vulnerable. This raises a coordination problem: What should be the target global effort level? We propose a novel method of heterogeneous thresholds and payoffs using a threshold public goods game. In this method, some thresholds provide a return to all players, but other thresholds provide a return to only some players. We experimentally find that heterogeneity in vulnerability significantly reduces contributions and coordination on high effort, compared to a setting with no difference in the vulnerability of players. This problem of low coordination on high effort is mitigated when the less vulnerable player contributes first. This indicates the potential importance of leadership and commitment in increasing global coordination on high effort.

JEL classification: C70; C91; C92; H41; Q54

Keywords— threshold public goods game, heterogeneous threshold public goods games, coordination, climate change

[a]I would like to sincerely thank my supervisor, Prof. Martin Kocher, for his invaluable guidance and time and Prof. Christian Koch for his useful insights and continued support throughout this research. Furthermore, I have also benefited greatly from many fruitful discussions with Prof. Karl Schlag. I also would like to thank Prof. Víctor González Jiménez for his help in understanding the elicitation tests for risk and loss preferences. I am grateful to Boris Knapp for many insightful discussions. I am sincerely thankful to Dr. Geoffrey Castillo, Dr. Philipp Kulpmann and the wonderful lab assistants of the VCEE who were indispensible to the smooth running of the experiment. I also thank the faculty of VGSE for their comments and feedback.

[b]Vienna Graduate School of Economics (VGSE) and Vienna Center for Experimental Economics (VCEE), University of Vienna

1 Introduction

Tackling climate change is the need of the hour. Catastrophic climate change has become a reality with different impacts in different parts of the world. Countries situated in hotter parts, for example in the tropical belt, are the most vulnerable and bear the brunt of climate change. These countries tend to be poorer and less developed than countries situated in cooler parts of the world and are less vulnerable to disastrous climate change. The population in more vulnerable countries has been washed away in floods, their crops burnt by high (and still rising) temperatures or annihilated by locust swarms, storms, and other natural disasters. According to the IMF's World Economic Outlook Report (2017), "Economies of developing countries are facing accelerating climate threats, as evidenced by recent hurricanes in the Caribbean and Latin America and the extreme odds in Asia and Africa." Similar findings are presented by the Global Climate Risk Index by Winges, Hutfils, and Eckstein (2019). Kompas, Ha, and Che (2018) show that if global temperatures increase by 4°C, Malawi would lose 13.6% of its GDP per year, Bangladesh 11.2% and India 14.6% compared to the United Kingdom losing 0.6% and the United States losing 0.8% of their respective GDP per year. The former countries – more vulnerable and poorer – need to spend significant resources fighting increasingly frequent and intense natural calamities. Instead, their resources could have been directed towards much-needed public infrastructure, education and health[1].

Climate change can be solved and, in particular, more vulnerable nations can be helped by sufficient mitigation efforts on a global scale. The question lies in how much effort should be put towards mitigation. More vulnerable countries need significantly more global effort to be safe from catastrophic climate change. However, less vulnerable countries need less effort. Therefore, although coordinating on high mitigation efforts would protect both the less vulnerable and the more vulnerable countries, it is not personally beneficial for the less vulnerable countries to spend their resources on high mitigation efforts. This raises a coordination problem: What should be the target global mitigation effort level?

The aim of this research is to obtain insights on coordination behaviour when agents are differently vulnerable[2]. Several works in the literature have explored the impact of different types of heterogeneity on coordination, such as Chan, Mestelman, Moir, and Muller (1999), Croson and Marks (1999), Burlando and Guala (2005), Buckley and Croson (2006), Tavoni, Dannenberg, Kallis, and Löschel (2011), Feige, Ehrhart, and Krämer (2018), Fellner, Kröger, and Seki

[1]See also Burke and Hsiang (2015), Burke, Davis, and Diffenbaugh (2018) and Diffenbaugh and Burke (2019) for additional findings on how climate change benefits wealthy countries in cooler countries

[2]Mitigation of climate change is a multi-faceted problem. The challenges to mitigation are a result of differences in wealth, vulnerability, abatement costs etc. between nations. We decide to focus on heterogeneity in vulnerability only, because if there exists a coordination problem with one kind of heterogeneity, two kinds of heterogeneity will most likely exacerbate the problem further. Thus, as a starting point, we focus on one kind of heterogeneity, that is, vulnerability

(2019), and Reindl (2022). We use a threshold public goods game to model the aforementioned setting. The non-linear feature of a threshold public goods game can capture the non-linear effects of climate change. In addition, thresholds can naturally model target global effort levels and climate tipping points. To model differences in vulnerability, we use a novel method of heterogeneous thresholds and payoffs. In this method, we employ a multi-threshold public goods game. Lower thresholds model lower mitigation targets and yield a high payoff to the less vulnerable player but a moderate payoff to the more vulnerable player. Higher thresholds model high mitigation targets and high payoffs to both players. The crucial difference between the lower and higher thresholds is that the more vulnerable player only gets the high payoff at the higher threshold, giving her the incentive to target it. However, the less vulnerable player already gets the high payoff at the lower threshold, giving her no monetary incentive to target the higher threshold. Here lies the coordination problem: On which threshold should agents coordinate?

We lend particular focus to coordination behaviour at the higher threshold. Coordination on the higher threshold is crucial to protecting more vulnerable countries. In addition, coordination on the higher threshold is more challenging because less vulnerable parties do not receive an additional return from contribution to this higher threshold. Studying whether a high level of coordination can be achieved and sustained in this asymmetric environment is our central question. By using a threshold public goods game and introducing heterogeneity in thresholds, which is the novelty of this research, our model closely captures the real-world situation we are interested to study.

We design and implement an experiment modelling the aforementioned heterogeneity in vulnerability. This is our main treatment of interest. We implemented two controls: one treatment with less heterogeneity compared to the main treatment, and another treatment with no heterogeneity. Each treatment is repeated for 10 rounds. Theoretically, we find that coordination on the higher threshold is possible in all treatments. However, coordination should decline to the lower threshold in the last round of the heterogeneous treatments, but not in the homogeneous treatment. In addition, the stage game equilibrium predicts coordination on the higher threshold only in the homogeneous treatment. Thus, we expect to observe coordination on the higher threshold less frequently in the heterogeneous treatments, compared to the homogeneous.

We find that heterogeneity significantly impedes high-level coordination. Coordination on the higher threshold is approximately 30 - 40% lower in the heterogeneous treatments, compared to the homogeneous control. Group contributions and payoffs are over 4 points lower in the heterogeneous treatments, compared to the homogeneous. While the difference in contributions is statistically significant, the difference in payoffs is not. However, the negative effect on payoffs is consistent with the finding that the higher threshold, which yields higher group payoffs, is achieved less often in the heterogeneous treatment.

Having concluded that heterogeneity indeed impedes coordination on the higher threshold, we explore possible solutions that can remedy the problem. We consider sequential contributions and role-reversal. Sequential contributions have leadership features, wherein the party that contributes first can set an example for others to follow. A high contribution from the first mover can set a norm of high contributions. Role-reversal is motivated by the social belief that increasing awareness of others' situations and providing the opportunity to experience others' situations can improve cooperation between groups with varying levels of advantage. Furthermore, recent unprecedented heatwaves, floods, and other natural disasters have rocked countries that have traditionally been deemed less vulnerable. We are interested to study whether such natural interventions, beyond the control of a human policy maker, can improve coordination.

Based on the Subgame Perfect Nash Equilibria, we expect to observe coordination only on the lower threshold in the sequential contribution treatments. In the role-reversal treatment, we expect some coordination on the higher threshold, similar to the earlier heterogeneous treatment. We find that imposing a specific sequence of contributions, that is, where the less vulnerable contributes first, significantly improves coordination on the higher threshold. In the treatment where the less vulnerable contributes first, coordination on the higher threshold improved to more than 7 out of 10 rounds. This is a significant improvement over the coordination rates in the heterogeneous treatment with simultaneous contributions, which was approximately 5 out of 10 rounds. This 20% difference is significant. We did not observe improved coordination on the higher threshold in any of the other proposed solutions. In addition, payoffs of the more vulnerable player increase significantly by over 4 points. Thus, we find that sequential contributions where the less vulnerable player contributes first solves the problem of coordination on the higher target effort arising out of differential vulnerability.

While we do not observe improved coordination in the role-reversal treatment, we make an interesting observation. After the reversal, less vulnerable players contribute more, but the more vulnerable players contribute almost equally less, resulting in no difference in coordination or group contributions. There is only a remarkable increase in the difference in the contributions of the less and more vulnerable players in the role-reversal treatment, compared to the heterogeneous, no - reversal benchmark. This finding suggests that the effect of reversal on cooperative behaviour may vary, depending on the individual's initial role. Further research on the reasons for this varying effect would help us understand the strengths and limitations of reversal as a means to promote pro-social behaviour.

The rest of the paper is organised as follows. Section 2 provides an overview of the literature we find most relevant to this work. Section 3 elaborates on the theoretical framework. Sections 4 and 5 detail the experimental design and hypotheses, and the sample and data collection. Section 6 presents the first part of the analysis and the results. Section 7 presents the details of the treatments proposed as solutions. Section 8 presents the results and analysis of the solution treatments. Section 9 details some additional results. Section 10 concludes the paper.

2 Overview of Relevant Literature

Previous works have studied different kinds of heterogeneity in the context of public goods games relevant to climate change. Chan et al. (1999) study a non-linear public goods game with heterogeneous endowment and preferences, and their effect on contributions in the presence of complete and incomplete information. They find that heterogeneity increases coordination, but communication reverses the effects of single or double heterogeneity. Furthermore, heterogeneity increases coordination when information is incomplete. Croson and Marks (1999) study the effect of heterogeneity in threshold valuation, and find higher variance in contributions when valuations are heterogeneous, despite no significant difference in average contributions. Burlando and Guala (2005) use a classification strategy to identify 'types' of subjects, and form homogeneous groups which play the public good game. The authors observe higher average contributions, declining contributions in groups composed of free-riders, but stable and high contributions in groups composed of cooperative and reciprocating players. Buckley and Croson (2006) study a linear public goods game with income and wealth heterogeneity. They find that the same absolute amount is contributed by the less wealthy as by the more wealthy. Thus, the less wealthy contribute a higher percentage of their income to the public good. Tavoni et al. (2011) study a threshold public goods game with heterogeneity in inherited inequality in wealth. They find that inequality makes provision towards the public good harder but pledging (even non-binding pledging) helps. Feige et al. (2018) study a threshold public goods game with an uncertain threshold, heterogeneity in marginal abatement costs and non-binding voting. Think of abatement costs in the context of industrialised countries (like the U.S. and the members of the EU), which have high marginal abatement costs, as these countries have spent significant resources towards abatement already. Developing countries (such as China and India) have comparatively low marginal abatement costs, as these countries have not spent significant resources towards abatement yet. They find that a non-binding unanimous voting procedure results in groups agreeing on an optimal total contribution more frequently along with high rates of compliance, even with heterogeneous marginal contribution costs. However, groups that do not reach such an agreement perform worse than treatments without voting. Groups seem to adhere primarily to a burden-sharing rule that equalises individual contribution costs, even at the cost of the group's total payoff. Mcginty and Milam (2013) experimentally examine provision of linear public goods with decreasing marginal benefits and increasing marginal costs. They find that although the design eliminated the coordination problem at the individual level, over-contribution persisted. Fellner et al. (2019) explore the effect of heterogeneity in external marginal returns, and find impediments to coordination due to heterogeneity. Reindl (2022) explore heterogeneity in vulnerability and wealth in a game introduced by Milinski, Sommerfeld, Krambeck, Reed, and Marotzke (2008). Fischbacher, Teyssier, and Schudy (2012) study how heterogeneous returns and uncertainty about one's own return impact unconditional and conditional contributions in a linear public goods game. They find that heterogeneity in returns decreases unconditional contributions. However, it only affects contributions weakly.

The literature on multi-threshold public goods games is much less compared to the literature on public goods and heterogeneity. Chewning, Coller, and Laury (2001) study a five-player threshold public goods game with one, two, three, or five thresholds. Comparing two and three thresholds, they find that more thresholds initially increase contribution, but contributions drop to below two thresholds after some periods. However, contributions remain higher than in the one-threshold case. Hashim, Maximiano, and Kannan (2011) analyse a game with five players and five thresholds. The authors vary information feedback about other members' contributions to a sub-sample of group members. They find no difference between the random provision of information and non-provision of information. Average contributions improve with targeted treatments. Targeted treatments also reduce coordination waste. Normann and Rau (2014) study a two-person threshold public goods game with one or two thresholds and with sequential or simultaneous moves. They find that the additional threshold increases contributions but does not improve public goods provision. It also lowers payoffs.

Important research in Experimental Economics on climate change has been conducted in recent years. Barrett and Dannenberg (2012) experimentally show that the fear of crossing a dangerous threshold can turn climate negotiations into a coordination game, virtually assuring collective action to avoid the dangerous threshold. They find that these results are robust to uncertainty about the impact of crossing a threshold, but uncertainty about the threshold's location turns the game back to a prisoner's dilemma, causing cooperation to fail. Dannenberg, Löschel, Paolacci, Reif, and Tavoni (2014) experimentally explore how uncertainty in threshold value affects collective action in a series of threshold public goods games. They find that provision is certain when the value of the threshold is certain. However, uncertainty of threshold value makes public good provision hard and contributions variable. Ambiguity makes provision even worse. Early and credible commitment has been shown to help combat threshold uncertainty. Barrett and Dannenberg (2013) show experimentally that reduction in the size of threshold uncertainty may bring about the behavioural change necessary to avoid crossing over the climate tipping points and causing disastrous climate change. Dannenberg, Riechmann, Sturm, and Vogt (2007) study the role of other-regarding preferences (as defined by the Fehr-Schmidt inequality aversion model) in coordination in public goods games. They find that specific composition of pairs significantly influences performance in public goods games. Along similar lines, Dannenberg, Sturm, and Vogt (2010) conduct an experiment with individuals involved in international climate policy to investigate whether climate negotiators have preferences for equity, and whether such preferences would explain different positions in international climate policy. They conclude that the aforementioned differing positions are more motivated by national interests than individual equity preferences.

The existing literature has only focused on coordination on one threshold, and the additional higher threshold is added only to increase provision on the lower threshold. There has been no focus on coordination at the higher threshold itself – how to achieve and sustain it. To the best of our knowledge, we are the first paper to focus on exploring how to achieve and sustain coordination on the higher threshold, which is not straightforward in the presence of heterogeneity, specifically, heterogeneity in vulnerability. We introduce a novel feature of heterogeneous thresholds and payoffs to the otherwise symmetric threshold public goods game, by virtue of which we model heterogeneity in vulnerability.

3 The Framework

We use a standard multi-threshold public goods game to set up the framework of differential returns to differently-vulnerable players at different global effort levels. Thresholds in public goods game can be used to model target global efforts. The consequences of climate change can be non-linear[3] and thus better modelled by a threshold public goods game, rather than a linear one. In particular, a disproportionate effect of achieving or not achieving a target can be modelled in a threshold public goods game. In light of the situation discussed earlier, higher threshold would be required by more vulnerable countries to receive the return that less vulnerable countries would get at the lower threshold. In addition, the less vulnerable players have no incentive to contribute to the higher threshold because they receive no additional return from it. However, the higher threshold, which models a high global mitigation effort, cannot be achieved by the more vulnerable alone. With these features, we describe the general framework below.

Table 1: Utility function of each player depending on total contributions of a group

Less vulnerable			**More vulnerable**		
Case	Return	Utility	Case	Return	Utility
$\Sigma < M_1$	0	0 - x^l	$\Sigma < M_1$	R'	R' - x^m
$M_1 \leq \Sigma$	R_2	R_2 - x^l	$M_1 \leq \Sigma < M_2$	R_1	R_1 - x^m
			$M_2 \leq \Sigma$	R_2	R_2 - x^m

The left panel shows the payoffs of the less vulnerable player, and the right panel shows the payoffs of the more vulnerable player. The payoff for each player depends on the total

[3]See Franzke (2014), Schneider (2004), Huang, Braithwaite, Charlton-Perez, Sarran, and Sun (2022), Rial et al. (2004)

contributions of the group, denoted by Σ. M_1 and M_2 denote the lower and higher thresholds, respectively. In general, a player's utility is the difference between her contribution and the corresponding return from the public good. Consider the less vulnerable player. She receives 0 returns when group contributions are less than M_1 and a return R_2 when group contributions are at least equal to M_1. The resulting utility functions of the less vulnerable player are as follows:

$$\text{when} \quad \Sigma < M_1 : \quad U^l = 0 - x^l \tag{1}$$

$$\text{when} \quad M_1 \leq \Sigma : \quad U^l = R_2 - x^l \tag{2}$$

Similarly, consider the more vulnerable player. She receives a return of R', R_1 or R_2 when group contributions are less than M_1, at least equal to M_1 but less than M_2, or at least equal to M_2, respectively. The resulting utility functions of the more vulnerable player are as follows:

$$\text{when} \quad \Sigma < M_1 : \quad U^m = R' - x^m \tag{3}$$

$$\text{when} \quad M_1 \leq \Sigma < M_2 : \quad U^m = R_1 - x^m \tag{4}$$

$$\text{when} \quad M_2 \leq \Sigma : \quad U^m = R_2 - x^m \tag{5}$$

We make the following reasonable assumptions:

- $x^m, x^l \geq 0$: Each player must contribute 0 or more.
- $M_2 > M_1 > 0$: The higher threshold must be of greater value than the lower threshold. The lower threshold should be greater than 0.
- $R_2 > R_1 > 0 > R'$: Achieving higher thresholds correspond to receiving higher returns, depending on the player's vulnerability.

Keeping the above assumptions in mind, we select the following parameters:

Table 2: Payoffs to each player (in experimental points)

Less vulnerable		**More vulnerable**	
Endowment	20	Endowment	20
$\Sigma < 25$	0	$\Sigma < 25$	-7 (R')
$25 \leq \Sigma$	32 (R_2)	$25 \leq \Sigma < 35$	16 (R_1)
		$35 \leq \Sigma$	32 (R_2)

Each player has an endowment of 20 experimental points. The values of the thresholds M_1 and M_2 are 25 and 35 respectively. Hence, both players together can achieve either threshold, since the sum of their endowments exceeds the value of either threshold. However, no player can achieve a threshold by herself, since her own endowment is less than the value of either threshold. This feature of the game highlights the importance of coordination and allows players to achieve any threshold by coordination. The values of R', R_1 and R_2 are -7, 16 and 32 respectively. A loss of 7 points was chosen for the more vulnerable player in contrast to no loss of points for the less vulnerable player in the event that the lower threshold was missed M_1. The negative payoff illustrates the dire consequences that only more vulnerable countries may face if there are little to no global mitigation efforts. The highest return from the public good is 32. This is received by the less vulnerable when M_1 is achieved, but is received by the more vulnerable only when M_2 is achieved. This framework forms our main treatment of interest, which we denote by $\mathcal{H}_2$.

We introduce two control treatments, $\mathcal{H}_1$ and $\mathcal{H}_0$ with parameters as follows:

Table 3: Treatment-wise payoffs for each player (in experimental points)

	$\mathcal{H}_2$		$\mathcal{H}_1$		$\mathcal{H}_0$	
	Less	**More**	**Less**	**More**	**Less**	**More**
Endowment	20	20	20	20	20	20
$\Sigma < 25$	0	-7	-7	-7	-7	-7
$25 \leq \Sigma < 35$	32	16	32	16	16	16
$35 \leq \Sigma$	32	32	32	32	32	32

$\mathcal{H}_0$ is a completely homogeneous treatment and is a natural control for the heterogeneous $\mathcal{H}_2$. Note that there are two changes in payoffs between $\mathcal{H}_2$ and $\mathcal{H}_0$, namely, the payoffs to the less vulnerable player when contributions are at least equal to M_1 or lower. To better control the effects of payoff changes between $\mathcal{H}_2$ and $\mathcal{H}_0$, we introduce a second control $\mathcal{H}_1$. Thus, there is now only one payoff change between $\mathcal{H}_2$ and $\mathcal{H}_1$ (less vulnerable player's payoff when M_1 is not achieved), and $\mathcal{H}_1$ and $\mathcal{H}_0$ (less vulnerable player's payoff when M_1 is achieved.).

4 Experimental Design

Below, we provide an overview of the timeline of the experiment, that is, the order of tasks which participants engaged in during the experiment. This sequence of tasks remained the same in every session.

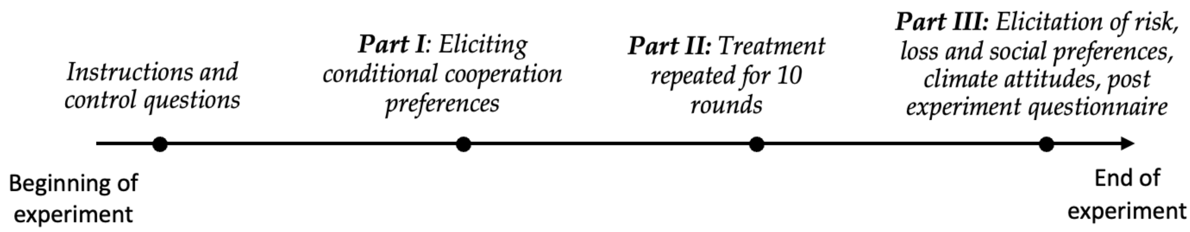

Figure 1: Timeline of tasks in experiment

The experiment began with the reading of the instructions[4], after which participants answered some questions to ensure their understanding of the instructions. All participants answered all questions successfully, confirming that they understood the implications and mechanisms of the payoffs in the game. The decision-making parts of the experiment, that is, ***Parts I, II and III***, were then implemented, details of which are elaborated upon below.

General Instructions about the experiment

Participants were first provided with general instructions about the experiment. Information about the exchange rate from earnings in experimental points to euros, and anonymity of decisions was provided in the general instructions. One of the treatments – $\mathcal{H}_2$, $\mathcal{H}_1$, or $\mathcal{H}_0$ – was randomly chosen to be implemented in the session. Therefore, each subject participated in only one of the treatments. In addition to the role-specific payoffs, participants were told that they will be randomly assigned to the role of a less or more vulnerable player. They will remain in their assigned role throughout the experiment[5]. Furthermore, the basic decision situation – deciding how many points out of the endowment shall be contributed to the public good – and the calculation of one's own total income and the group member's total income was detailed. The total income of a player was the sum of the return she received from the public good, depending on the threshold achieved by her group, and the points left of her endowment after she contributed to the public good. To maintain neutral framing, less and more vulnerable participants were referred to as A and B participants respectively, the public good was referred to as the group account, and thresholds were referred to as targets[6].

[4]Sample instructions for $\mathcal{H}_2$ provided in Appendix. We kept changes in instructions between treatments to a minimum, and thus the instructions for other treatments are largely similar to the instructions for $\mathcal{H}_2$. Instructions for other treatments are available on request

[5]Therefore, from the beginning of the experiment, participants in the heterogeneous treatments $\mathcal{H}_2$ and $\mathcal{H}_1$ knew that in each group, there will be one less and one more vulnerable participant. They could be assigned to either role, and this difference in role and payoffs will hold throughout the experiment. Only in the homogeneous treatment $\mathcal{H}_0$, there was no difference in roles or payoffs among participants in any part of the experiment.

[6]All instructions are provided in the Appendix.

Part I: Eliciting unconditional and conditional cooperation preferences

In this task, we elicited participants' cooperation preferences by adopting the experimental design introduced in Fischbacher, Gächter, and Fehr (2001), which is a variant of the strategy method (Selten (1967)). First, participants were asked to state their *unconditional contribution*. Second, they filled out a *conditional contribution schedule*, which asked them to indicate their contribution decision for every possible contribution by their group member. Each subject made 22 decisions in this task – 1 unconditional contribution decision and 21 conditional contributions decisions for the group member's possible contribution from 0 to 20. The role to which each participant was randomly assigned was revealed at the beginning of *Part I*, that is, before participants made any decision. To calculate a participant's earnings from *Part I*, one participant from each group, irrespective of her role, was chosen at random. The payoff-relevant decision of this randomly chosen participant would be her conditional contribution schedule. The payoff-relevant decision of the other participant in the group would be her unconditional contribution decision. Total group contribution was determined as the sum of the unconditional contribution of the unchosen participant and the conditional contribution of the chosen participant for the value of the other participant's unconditional contribution. Based on this sum, each player received payoffs from the public good as illustrated in Table 3. The total income for each player was the sum of her payoffs from the public good and the points remaining of her endowment, as usual.

We explain this determination of total income in *Part I* with a simple example. Let us consider treatment $\mathcal{H}_2$ and denote two participants in a group by X and Y. Let X be the less vulnerable participant and let Y be the more vulnerable participant. Let us assume that X was chosen randomly. This means that the payoff-relevant decision for X is her conditional contribution schedule, and the payoff-relevant decision of Y is her unconditional contribution. Let us assume that Y made an unconditional contribution of 15. X indicated in her contribution schedule that she would contribute 11 if Y possibly contributed 15. Thus, the total group contribution would be $15 + 11 = 26$. M_1 is achieved since total contributions exceed 25, which is the value of M_1. X receives 32 points from the public good and has 9 points left of her endowment, making her total income $32 + 9 = 41$ points. Y receives 16 points from the public good and has 5 points left of her endowment, making her total income $16 + 5 = 21$ points. Such examples were provided to make the calculation of total income easier to understand[7]. Subjects' earnings from *Part I* were not revealed until the end of the entire experiment. This was done to ensure that there are no overlapping effects of the results of *Part I* on the decisions to be made in *Part II*.

[7] We also made it clear that the numbers in the examples were for illustrative purposes only so as to avoid any anchoring effects.

Part II: Treatment over 10 rounds

At the beginning of *Part II*, each participant was reminded of the role she was assigned to, along with the payoffs corresponding her role depending on the threshold achieved. Participants were randomly matched anew to another subject in the role other their own, which was also common knowledge. For example, after *Part I*, each *less* vulnerable subject was matched with a new *more* vulnerable subject and vice versa. The treatment was repeated for 10 rounds and participants remained matched with the same group member throughout the 10 rounds. The basic decision situation – how much to contribute to the public good out of the endowment – remained the same. The total income of a participant was the sum of her payoff from the public good and the points she retained out of her endowment. Contribution decisions were taken simultaneously by both players in each round. After each round, players were given information about their own and their group member's contribution and payoff in all preceding rounds. After the 10 rounds, one round was randomly chosen to be the payoff-relevant round for both group members. This was to ensure that subjects took the decisions in each round carefully as any round could be their payoff-relevant round.

Part III: Elicitation of risk, loss and social preferences, climate attitudes and post experiment - questionnaire

We elicited risk and loss preferences using a decision-making task introduced by Abdellaoui, Bleichrodt, and L'Haridon (2008). Given the potential loss of points to subjects depending on their assigned role, it was important to elicit subjects' loss preferences, in addition to risk preferences. Subjects were asked to choose their preferred lottery between two lotteries, and they made their decision for several pairs of such lotteries. Furthermore, we measured social preferences using the resource allocation task introduced by Murphy, Ackermann, and Handgraaf (2011). Subjects were asked to indicate their preferred allocation of points between themselves and another subject to whom they were newly matched after *Part II*. Each subject made six such allocations, on the bases on which the subject was classified as *Individualistic* (maximises her own payoff), *Competitive* (maximise difference in payoffs), *Prosocial* (maximises joint payoffs or minimises difference in payoffs) or *Altruistic* (sacrifices own payoff for others gain). We also elicited individual attitudes toward climate-related issues with a revised version of the Climate Change Attitude Survey (Christensen and Knezek (2015)). Subjects were asked 15 Likert-type belief-based and intention-based questions about climate change. In the last questionnaire, we asked participants to explain their decision-making process in *Tasks 1* and *2*. Further details are provided in the Appendix.

Sample and Data Collection

The experiment was conducted in the laboratory of the Vienna Center for Experimental Economics, University of Vienna. 118 individuals participated in the experiment, which was completed in 7 sessions. Power calculations suggested a sample size of nine independent observations for each treatment, which implies 18 subjects (9 groups of two participants each) in every treatment. 38 - 40 individuals participated in each treatment, which yielded 19 - 20 independent observations per treatment. The sessions for $\mathcal{H}_0$ lasted around 1 hour. The sessions for $\mathcal{H}_1$ and $\mathcal{H}_2$ lasted around 1 hour and 20 minutes. The average earning ranged between 16 - 19 euros depending on the treatment. The average age of participants was approximately 26 years, with a standard deviation of 5.26 years and ranged between 19 to 54 years. 57% of the sample was female, 40% was male, and 3% identified as non-binary. 25% of the sample reported having a business, finance, or economics related background.

5 Hypothesis

The Pure Strategy Nash equilibrium of the stage game and the Subgame Perfect Nash Equilibrium of the finitely repeated game form the basis of our hypothesis. Let us first consider the stage game. The Pure Strategy Nash equilibrium of the one-shot $\mathcal{H}_2$ and $\mathcal{H}_1$ is $x^l, x^m = (0, 0)$ or $x^l + x^m = 25$ such that $x^m \in \{11, \dots 20\}$, $x^l \in \{5, \dots 14\}$. The Pure Strategy Nash Equilibrium of the one-shot $\mathcal{H}_0$ is $x^l, x^m = (0, 0)$, $x^l + x^m = 25$ such that $x^l, x^m \in \{11, \dots 14\}$, or $x^l + x^m = 35$ such that $x^l, x^m \in \{15, \dots 20\}$. Thus, the stage game prediction for $\mathcal{H}_2$ and $\mathcal{H}_1$ is zero contributions or coordination at M_1, but not M_2. The prediction of the stage game only for $\mathcal{H}_0$ includes coordination on M_2, along with coordination on M_1 and zero contributions.

Let us consider the finitely repeated game. The Subgame Perfect Nash Equilibria of a finitely repeated game are strategies consisting of any combination of its stage game equilibria. Hence, coordination on M_2 is a Subgame Perfect Nash Equilibrium for $\mathcal{H}_0$. Coordinating on M_2 is not a stage game equilibrium in $\mathcal{H}_2$ and $\mathcal{H}_1$. However, using Benoît and Krishna (1985), we construct a strategy to achieve coordination on M_2, which is a Subgame Perfect Nash Equilibrium. We denote this strategy by $\mathcal{S}$. Detailed proof is provided in Appendix section 11.1.

Thus, coordination on M_2 is one of the Subgame Perfect Nash Equilibrium strategies in all treatments $\mathcal{H}_2$, $\mathcal{H}_1$ and $\mathcal{H}_0$. However, coordination on M_2 is not a stage game Nash Equilibrium for $\mathcal{H}_2$ or $\mathcal{H}_1$. In addition, the strategy described earlier to sustain coordination on M_2 in $\mathcal{H}_2$ includes an end game effect – coordination declines to M_1 from M_2 in the last round. This indicates that over the 10 rounds, we should see coordination on M_2 in a lesser number of rounds in $\mathcal{H}_2$ and $\mathcal{H}_1$, compared to $\mathcal{H}_0$. With these expectations, we formally hypothesize the following :[8]:

[8]Behaviourally, we expect some coordination on M_2 in treatments $\mathcal{H}_2$ and $\mathcal{H}_1$ owing to social preferences of some less vulnerable players. This expectation is motivated by extensive findings in the literature about social considerations and other-regarding preferences apart from self-serving biases in people. Classic papers in this literature include Bolton (1991), Bolton and Ockenfels (2000), Fehr and Schmidt (1999) in distributive preferences, Fehr, Kirchsteiger, and Riedl (1993), "Trust, Reciprocity, and Social History" (1995) in reciprocity and Charness and Rabin (2002), Cox and Friedman (2003), Dufwenberg and Kirchsteiger (2004), Falk and

- **Hypothesis 1**: We expect to observation coordination on M_2 in all the treatments. However, we also expect to observe coordination on M_2 more often in $\mathcal{H}_0$, compared to $\mathcal{H}_2$ or $\mathcal{H}_1$.

6 Main Results: Heterogeneity

6.1 Graphical inspection of Mean Group Contributions

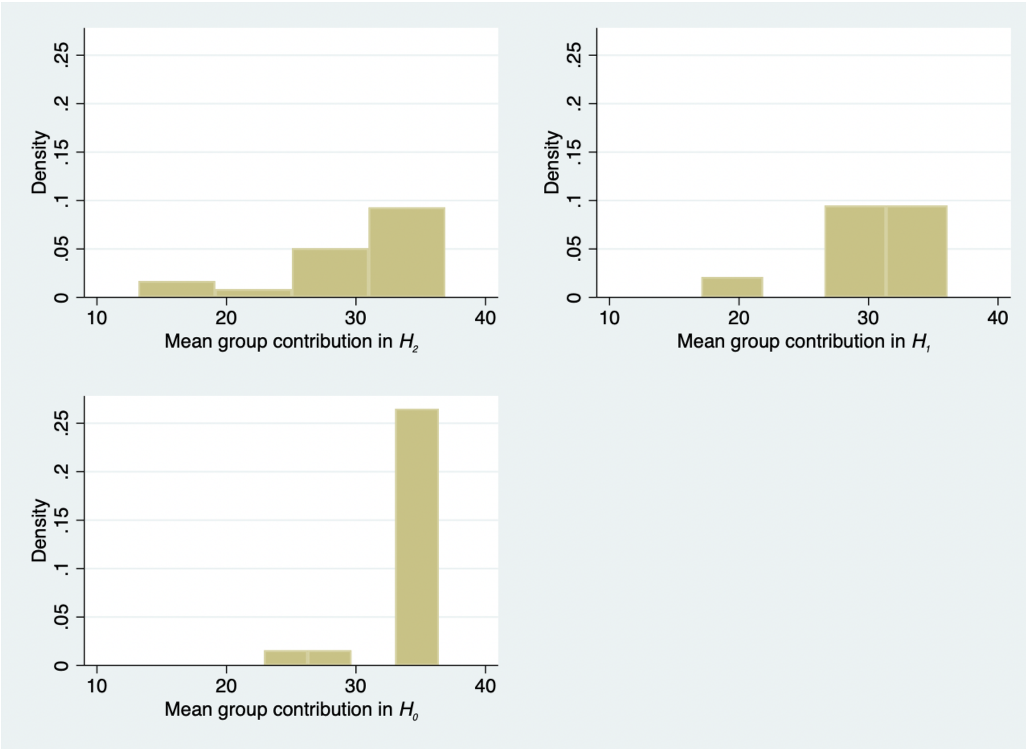

Figure 2: Mean group contribution, by treatment

In Figure 2, we present histograms of the mean group contribution per round in each treatment. The upper left histogram is for $\mathcal{H}_2$, the upper right panel histogram is for $\mathcal{H}_1$, and the lower left histogram is for $\mathcal{H}_0$. In each graph, the x-axis represents the mean group contribution. We observe that the highest contributions (close to 35) are made in $\mathcal{H}_0$. We also observe higher variance in contributions in both the heterogeneous treatments $\mathcal{H}_2$ and $\mathcal{H}_1$, compared to the homogeneous $\mathcal{H}_0$.

Fischbacher (2002) and Falk and Fischbacher (2006) in both. However, the problem of lack of coordination on M_2 is not expected to be completely solved by only the presence of social preferences.

6.2 Coordination rates

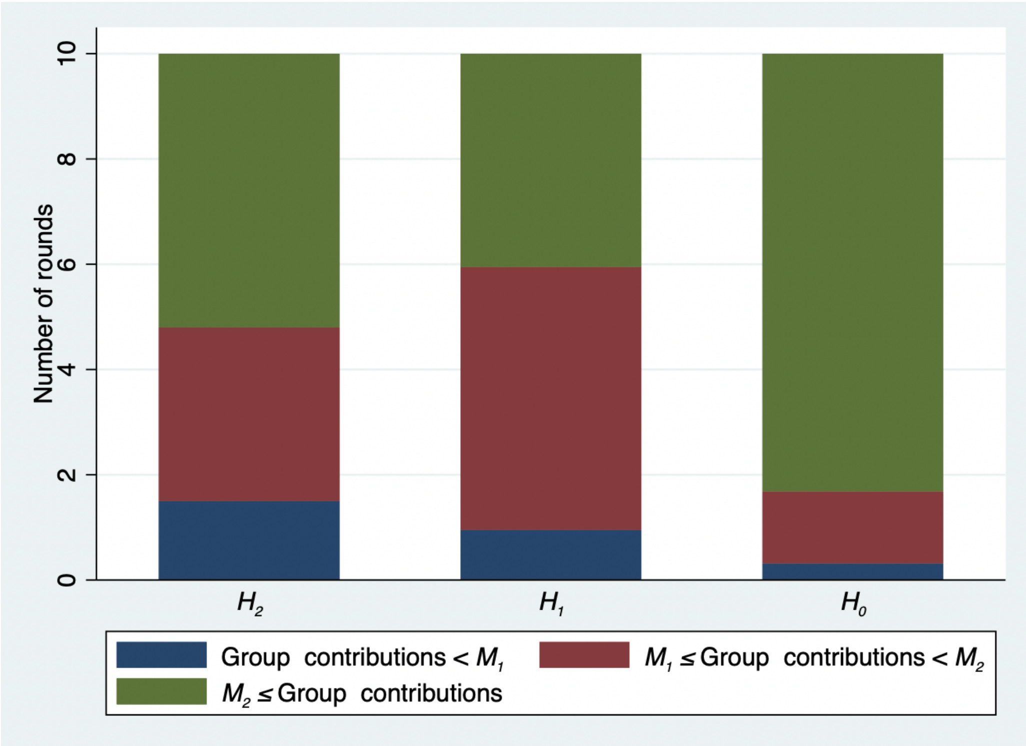

Figure 3: Coordination rates depending on group contributions, by treatment

Figure 3 shows the average number of rounds in which no threshold, M_1, or M_2 is achieved in each treatment. The blue section of each bar graph represents the average number of rounds in which the group contributions did not reach M_1. The red and green sections represent the average number of rounds in which coordination on M_1 and M_2 was achieved, respectively. We observe that coordination on M_2 was the highest in $\mathcal{H}_0$, in more than 8 out of 10 rounds. We use the Mann-Whitney U test to compare coordination rates between treatments. We find that M_2 was achieved significantly less in $\mathcal{H}_2$ and $\mathcal{H}_1$, only in 4 or 5 rounds ($p < 0.01$). We do not find that the difference in failure or success of coordination on M_2 between $\mathcal{H}_2$ and $\mathcal{H}_1$ to be statistically significant ($p > 0.1$)[9].

Coordination on M_1 is the highest in $\mathcal{H}_1$ and was achieved in 5 out of 10 rounds. Comparatively, M_1 was achieved in approximately 3 rounds and 1 round in $\mathcal{H}_2$ and $\mathcal{H}_1$ respectively. There is strong evidence of significant difference in coordination on M_1 between $\mathcal{H}_1$ and $\mathcal{H}_0$ ($p < 0.01$), and between $\mathcal{H}_2$ and $\mathcal{H}_0$ ($p < 0.05$). There is only mild evidence of a significant difference in M_1 coordination between $\mathcal{H}_2$ and $\mathcal{H}_1$ ($p < 0.1$). In general, the coordination rates on M_1 and M_2 are similar between $\mathcal{H}_2$ and $\mathcal{H}_1$, but not between $\mathcal{H}_0$ and either $\mathcal{H}_2$ and $\mathcal{H}_1$.

We also analyse strategies used by each group in the repeated game. We are interested to

[9]Throughout the paper, if we report that a certain result is not statistically significant, it means that the p - value of the test was greater than 0.1

observe whether the strategy $\mathcal{S}$ is adopted for coordination on M_2 in $\mathcal{H}_2$ and $\mathcal{H}_1$. The results are presented in Table 12 in the Appendix. We find that strategy $\mathcal{S}$ is not used by many groups in $\mathcal{H}_2$ or $\mathcal{H}_1$. The more commonly used strategy is sustained coordination on M_2 in all remaining rounds after first successful coordination on M_2. This strategy does not have an end-game effect, that is, no decline in contributions in the last round relative to previous rounds. We also observe some groups which coordinate on M_2 in some rounds, followed by M_1 in the remaining rounds. This strategy allows for coordination on M_2 in some rounds but bears no punishment of zero contributions for the less vulnerable player after the deviation to the lower threshold. This is in contrast to $\mathcal{S}$, which requires the threat of zero contributions in the event of deviation to sustain coordination on M_2.

Thus, we find that coordination on M_2 takes place approximately 4 or 5 times out of 10 in $\mathcal{H}_2$ and $\mathcal{H}_1$, compared to over 8 times in $\mathcal{H}_0$. In other words, heterogeneity reduces coordination on M_2 significantly by 30 - 40% compared to a homogeneous setting.

6.3 Treatments effects on individual and group contributions and payoffs

We compare group contributions and group payoffs across treatments with a random-effects panel data regression model. The results are tabulated below in Table 4. Summary statistics on contributions and payoffs are provided in Table 9 in Appendix. We do not compare individual contributions between $\mathcal{H}_2$, $\mathcal{H}_1$ and $\mathcal{H}_0$ since there are more and less vulnerable players in the heterogeneous treatments, but there are no such players in the homogeneous treatment.

Table 4: Random-Effects Panel Data regression results

	Dependent variable			
	Group contributions		Group payoffs	
	(1)	(2)	(3)	(4)
$\mathcal{H}_2$	-4.68^{***}	-4.87^{***}	-4.40	-4.64
	(1.52)	(1.63)	(3.76)	(3.91)
$\mathcal{H}_1$	-4.53^{***}	-4.45^{**}	-4.03	-3.94
	(1.34)	(1.36)	(3.22)	(3.27)
$\mathcal{H}_2$	-0.15	-0.41	-0.37	-0.69
(Base: $\mathcal{H}_1$)	(1.74)	(1.93)	(3.62)	(3.98)
Controls	No	Yes	No	Yes

Note: Standard errors in parentheses. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Base category is $\mathcal{H}_0$ for the first 2 rows. Controls include session size and round number. Estimates are robust to the exclusion of the first and the last round.

In Table 4, we estimate four random-effects panel data regression models, the results of which are tabulated in each column of the table. Columns (1), (2) and columns (3), (4) present the results of the aforementioned regression where the dependent variable is group contributions and group payoffs respectively. The coefficients reported in columns (2) and (4) are obtained after controlling for the session size and the round number. The reported estimates measure the treatment effect on group contributions or payoffs, relative to the benchmark. The benchmark treatment for the first two rows is $\mathcal{H}_0$. The benchmark treatment for the third row is $\mathcal{H}_1$ to estimate the treatment effect of $\mathcal{H}_2$ treatment compared to $\mathcal{H}_1$. The number of observations in the regression was 590. There were 20 groups and thus 20 independent observations in $\mathcal{H}_2$ and $\mathcal{H}_1$. There were 19 independent observations in $\mathcal{H}_0$. Each group was observed for 10 periods.

We find significantly lower group contributions in both heterogeneous treatments $\mathcal{H}_2$ and $\mathcal{H}_1$, compared to the homogeneous $\mathcal{H}_0$ ($p < 0.01$). Group payoffs are lower in $\mathcal{H}_2$ and $\mathcal{H}_1$ compared to $\mathcal{H}_0$. This effect is in the right direction given the significantly lower contributions in $\mathcal{H}_2$ and $\mathcal{H}_1$ compared to $\mathcal{H}_0$. However, we do not find this difference in group payoffs between $\mathcal{H}_2$ or $\mathcal{H}_1$ and $\mathcal{H}_0$ to be statistically significant[10]. We also find little, statistically insignificant difference in the sum of contributions and payoffs between $\mathcal{H}_2$ and $\mathcal{H}_1$.

Thus, we summarise the **main findings** from this analysis below:

- Heterogeneity **hurts coordination at the higher threshold significantly**.
- In addition, heterogeneity **significantly reduces group contributions** compared to the homogeneous setting.
- Heterogeneity **reduces group payoffs** compared to a homogeneous setting, although not significantly.

7 Proposed solutions

Having experimentally confirmed that heterogeneity reduces coordination at the higher threshold, we explore two possible solutions to this problem: sequential order in contributions, and reversal of roles.

7.1 Sequential contributions

The existing literature on public goods games finds that sequential contributions improve provision compared to simultaneous contributions[11]. In climate change actions, contribution decisions can take place sequentially, for example, in instances of leadership. Some countries can take on a leadership role and contribute first, setting an example for contributors who follow.

[10] Mann Whitney U test results comparing group payoffs shows a significant difference ($p < 0.05$) between $\mathcal{H}_2$ and $\mathcal{H}_0$, and $\mathcal{H}_1$ and $\mathcal{H}_0$.

[11] see Normann and Rau (2014), Coats, Gronberg, and Grosskopf (2009), Bolle (2016), Bolle (2014)

From the point of view of a policy maker, in principle, sequential contributions have a strong 'commitment' aspect. The first mover makes the contribution decision first, and thus essentially commits himself to his contribution decision. If this contribution decision is high, it would have the potential to encourage high contribution from other players as well. Therefore, we were interested in observing whether sequential contributions would increase coordination on M_2 in $\mathcal{H}_2$, compared to simultaneous contributions.

We implement two sequential move versions of $\mathcal{H}_2$. In one version, the less vulnerable player makes the contribution decision first in every round, followed by the more vulnerable player. We denote this treatment by *LFirst*. The second version is vice versa of the first and is denoted by *MFirst*. Who moves first is predetermined exogenously and is known to all players at the beginning of the 10 rounds. We find that in *LFirst* and *MFirst*, there is a unique Subgame Perfect Nash Equilibrium, in which the first mover contributes 5 points and the second mover contributes 20 points. Thus, we expect that groups would only coordinate on M_1 in all rounds.

- **Hypothesis 3**: There should be no coordination on M_2 in any round in *LFirst* and *MFirst*.

7.2 Role-reversal

Increasing awareness and giving people the opportunity to experience others' situations are tools frequently aimed at improving cooperation between groups with varying levels of advantage. Experimental research in dictator games, bargaining games and trust games find increased trust and pro-social behaviour when roles are reversed[12]. In light of the recent unprecedented heat waves and floods across 'less vulnerable' countries, the world is now increasingly experiencing the non-linear effects of climate change. We find this intervention to be a realistic reflection of the increasingly blurring categorisation of vulnerability. We find it interesting to explore how this intervention – one that is beyond the control of a human policy maker – can impact coordination. We implement role-reversal in a modified version of $\mathcal{H}_2$. Instead of each player remaining in her role for all 10 rounds, players switch roles and receive payoffs according to their new role after the first 5 rounds. The less vulnerable player reaps the payoffs of a more vulnerable player, and vice-versa. We denote this treatment by *Reversal*. In this paper, to the best of our knowledge, we are the first to explore reversal in vulnerability and implement it in a threshold public goods setting. The theoretical predictions for the stage game and the finitely repeated *Reversal* game as the same as the predictions for $\mathcal{H}_2$ (proof presented in Appendix section 11.2). Thus, similar to our expectations for $\mathcal{H}_2$, we expect to observe coordination on M_2 in some rounds during the repeated game. However, since coordination on M_2 is not a stage game Nash Equilibrium, we expect to observe coordination on M_2 in a number of rounds similar to $\mathcal{H}_2$.

[12]See Lange, Schmitz, and Schwirplies (2022) and references therein for a detailed overview

- **Hypothesis 4**: We expect to observe coordination on M_2 in some rounds in *Reversal*. We do not expect coordination rate on M_2 to be different between $\mathcal{H}_2$ and *Reversal*.

7.3 Sample and Data collection

The experiment was conducted at the Vienna Center for Experimental Economics laboratory at the University of Vienna. 120 subjects participated in the solution treatments, which were completed over 6 sessions. Session sizes varied between 16 to 20 participants. The average age was around 25 years and ranged between 18 and 62 years. 65% of the subjects were female, and 35% were male. 15% of the sample reported having a business, finance, or economics related background. The timeline of tasks in each solution treatment was the same as in $\mathcal{H}_2$. Each session lasted approximately 90 minutes. The average earning ranged between 16 - 19 euros depending on the treatment.

8 Main Results: Solution Treatments

8.1 Graphical inspection of Mean Group Contributions

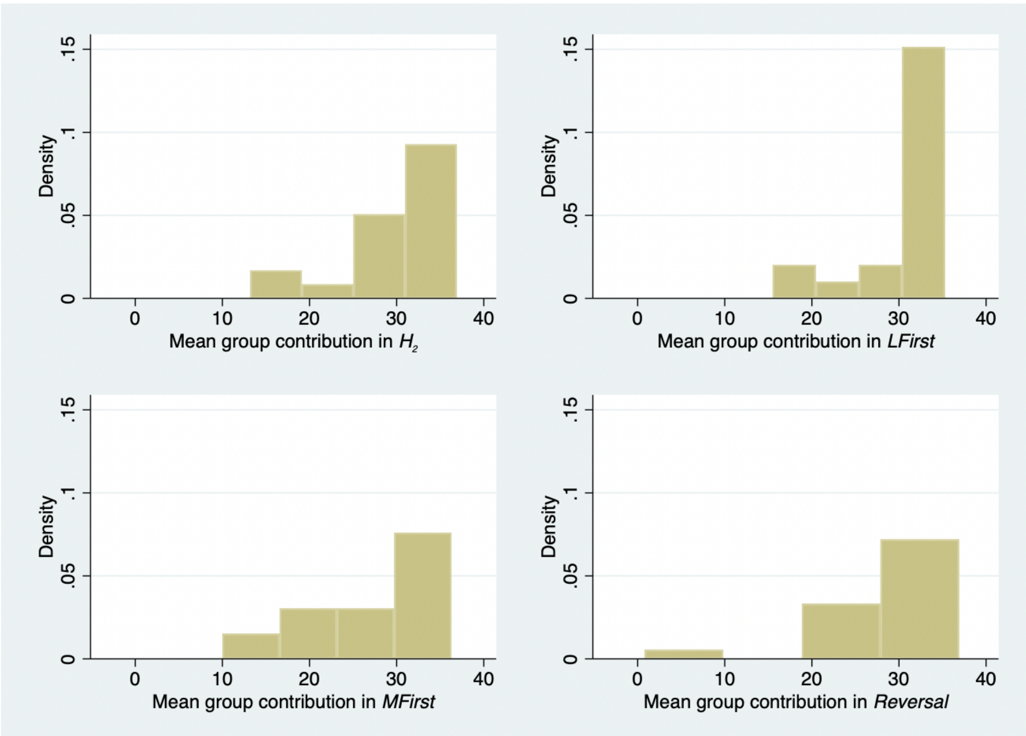

Figure 4: Mean group contribution, by treatment

In Figure 4, we present histograms of the mean group contribution per round in each treatment. The upper left histogram is for $\mathcal{H}_2$, the upper right panel histogram is for *LFirst*, the lower left histogram is for *MFirst* and the lower right histogram is for *Reversal*. In each graph, the x-axis represents the mean group contribution. We observe the highest contributions in *LFirst*. We also observe that the distribution of contributions of *MFirst* and $\mathcal{H}_2$ is largely similar. We also observe that the variance of contributions is the lowest in *LFirst* and the highest in *Reversal*.

8.2 Coordination rates

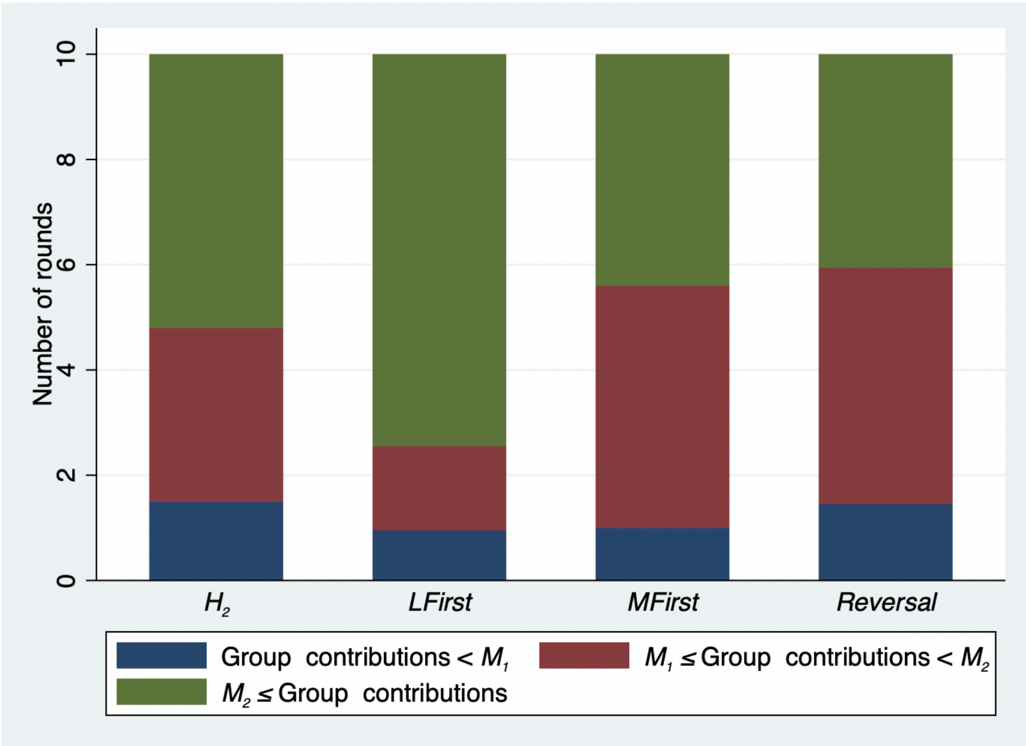

Figure 5: Coordination rates depending on group contributions, by treatment

With the same colour legend as in Figure 3, we compare the coordination rates when M_1 is missed, on M_1 and on M_2 in each of the solution treatments with the benchmark $\mathcal{H}_2$. Recall that $\mathcal{H}_2$ is the treatment where we detected the problem of lack of coordination on M_2. We use the Mann-Whitney U test to compare coordination rates between treatments, as in Section 6.2. Coordination on M_2 is observed in more than 7 out of 10 rounds in *LFirst*, which is the highest among all solution treatments. In fact, *LFirst* is the only solution treatments to have significantly higher levels of coordination on M_2 ($p < 0.05$). The levels of coordination on M_2 observed in *LFirst* are not significantly different from the levels observed in $\mathcal{H}_0$. In other words, the levels of coordination on M_2 in *LFirst* is almost as good as the levels observed in the homogeneous setting. We find no significant differences in coordination on any threshold when comparing $\mathcal{H}_2$ to *MFirst* and *Reversal*. We also observe no difference in the number of groups who follow strategy $\mathcal{S}$ in *Reversal* and $\mathcal{H}_2$. Therefore, imposing sequential contributions where the less vulnerable contribute first may be an effective solution to the lack of coordination on M_2 posed by heterogeneity.

8.3 Treatment effects on individual and group contributions and payoffs

Having concluded that coordination on M_2 is the highest in *LFirst*, we test whether this translates to significant differences in individual and group contributions, and payoffs. We estimate this treatment effect using a random-effects panel data model. We observe each unique group, which remains composed of the same subjects, for 10 rounds. This yields a total of 800 observations. The results of the regression are reported in Table 5 below.

Table 5: Random-Effects Panel Data regression results: Solution treatments

	Dependent Variable					
	Group Contributions			Group Payoffs		
	All Rounds	1 - 5	6 - 10	All Rounds	1 - 5	6 - 10
LFirst	1.28			5.68		
	(1.95)			(4.17)		
MFirst	−2.30			4.17		
	(2.30)			(3.67)		
Reversal	−0.45	−1.27	0.36	−0.78	−2.05	0.47
	(2.26)	(2.33)	(2.56)	(3.71)	(4.35)	(4.04)
N	800	400	400	800	400	400

	Dependent Variable					
	Individual Contributions			Individual Payoffs		
	All Rounds	1 - 5	6 - 10	All Rounds	1 - 5	6 - 10
LFirst	0.68			4.24*		
	(1.14)			(2.20)		
MFirst	−2.57*			2.51		
	(1.46)			(2.22)		
Reversal	−1.55	−0.15	−2.94**	3.67*	−2.55	9.91***
	(1.27)	(1.39)	(1.47)	(2.04)	(2.48)	(2.38)
Less	−1.29	−0.71	−1.87**	7.62***	7.66***	7.58***
	(0.86)	(0.86)	(0.93)	(1.46)	(1.33)	(1.82)
LFirst x *Less*	−0.13			−2.97*		
	(0.99)			(1.74)		
MFirst x *Less*	2.78*			−1.05		
	(1.60)			(1.92)		
Reversal x *Less*	2.60***	−1	6.21***	−8.3***	2.9	−19.5***
	(0.99)	(1.41)	(1.28)	(1.80)	(2.13)	(2.67)
N	1600	800	800	1600	800	800

Note: Standard errors in parentheses. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$. Under each dependent variable, first column reports the estimates for All Rounds, the second column for rounds 1 - 5 (before reversal) and the third column for rounds 6 - 10 (after reversal). In the upper panel, base category is $\mathcal{H}_2$. In the lower panel, the base category is *More vulnerable* in $\mathcal{H}_2$. Controls for round number and session size included. Estimates robust to exclusion of first and last round.

In Table 5, the upper panel reports the regression results with group contributions and group payoffs as the dependent variable. The lower panel reports the regression results with individual contributions and payoffs as the dependent variable. The first column under each of the dependent variables reports the coefficients in the regressions for all rounds.The second and the third columns under each dependent variable report the coefficients in the regressions for the first five and the last five rounds rounds separately. This is important for the *Reversal* treatment, where we compare the treatment effect of *Reversal* before and after the reversal with the comparative rounds in $\mathcal{H}_2$. In the table, we estimate the treatment effects separately for the first five and the last five rounds only for *Reversal* relative to $\mathcal{H}_2$. For *LFirst* and *MFirst*, we estimate the treatment effects over all rounds. Summary statistics of contributions and payoffs in *LFirst*, *MFirst* and *Reversal* are provided in the Appendix in Tables 10 and 11.

We do not find significant effects for group contributions and group payoffs for any of the solution treatments. Regarding individual payoffs, we find that the more vulnerable earns over 4 points more in *LFirst* compared to $\mathcal{H}_2$ ($p < 0.10$), but does not contribute significantly more. In *MFirst*, the more vulnerable player contributes around 3 points less ($p < 0.10$), but this does not significantly improve her earnings. We also observe that the difference in payoffs between the more vulnerable and less vulnerable players is lesser in *LFirst* compared to $\mathcal{H}_2$. Thus, we conclude that *LFirst* significantly improves coordination on M_2 and the payoffs of the more vulnerable, which makes it a good solution to the problem of coordination on M_2[13].

Similarly, we analyse the treatment effect of *Reversal* on individual and group contributions and payoffs. Since there is a reversal after five rounds, we studied the treatment effects in all rounds, as well as separately before and after the reversal. We find no significant differences in group contributions and payoffs between *Reversal* and $\mathcal{H}_2$. This is consistent with our previous finding that there are no differences in the coordination rates on M_1 or M_2 between *Reversal* and $\mathcal{H}_2$. In all rounds, we find that the more vulnerable players do not contribute a significantly different amount in *Reversal*, but earn approximately 4 points more ($p < 0.1$). Comparing $\mathcal{H}_2$ and *Reversal*, we find little difference between them in the first five rounds but a significant difference in the last five rounds, i.e., after the reversal. After the reversal, the difference in contributions between the less and more vulnerable players rises to over 6 points, compared to a difference of just 1 point in the last five rounds of $\mathcal{H}_2$. Particularly, less vulnerable players contribute over 3 points more but more vulnerable players contribute 3 points less. Given that the coordination levels between *Reversal* and $\mathcal{H}_2$ remain similar, one would expect the difference in contributions between players to also remain somewhat similar. We find it puzzling that the difference in contributions is significantly larger after the reversal. The difference in payoffs between the players is reduced in *Reversal* compared to $\mathcal{H}_2$, which is to be expected

[13]We acknowledge that there is only mild evidence ($p < 0.1$) that the more vulnerable player's earnings are significantly higher in *LFirst* compared to $\mathcal{H}_2$. This may be due to low power. We would need more observations to draw a definitive conclusion on whether *LFirst* significantly increases payoffs of the more vulnerable player compared to $\mathcal{H}_2$.

because of the reversed payoffs. However, the increased difference in contributions after the reversal, especially when there is no change in coordination, is a strange but interesting finding.Although the literature emphasises the positive effects of role-reversal, the aforementioned finding indicates that the effect of reversal may depend on one's initial role. While reversal may bring about a positive change in the advantaged group, this positive effect may be absent in the disadvantaged group. The reason for this observation may be a combination of guilt of the less vulnerable and the desire to make up for lost payoffs of the more vulnerable. It could also simply be the more vulnerable punishing the less vulnerable for coordination on M_1 before the reversal. Further research needs to be conducted to ascertain the potentially differing effects of role-reversal depending on one's initial role and to better understand the reasons for such an effect.

To summarise, our **main findings** from this section are as follows:

- Sequential contributions where the less vulnerable player moves first is the only solution treatment with **increased coordination on** M_2. This level of coordination on M_2 is comparable to that in a non-heterogeneous environment. It also **increases the payoffs of the more vulnerable player** significantly.
- Thus, sequential contributions where the less vulnerable player contributes first is a **good solution** to the lack of coordination on M_2 due to heterogeneity.
- Reversal leads to more contributions from the less vulnerable, but equally less contributions from the more vulnerable. This cancels out any opportunity for more coordination on the higher threshold. Thus, the **effect of reversal on coordination may vary depending on one's initial role**, in contrast to a general positive effect shown in the literature.

9 Other Results

9.1 Unconditional and conditional contributions

We implement a test introduced by Fischbacher et al. (2001) to identify cooperative preferences of players. Details of the elicitation task are described in Section 4. We adopt a new classification strategy to identify types of players, given their responses in the task. We divide the contribution schedule into three ranges: $\{0, 1, \ldots, 4\}$, $\{5, 6, \ldots, 14\}$, $\{15, 16, \ldots, 20\}$. $\{0, 1, \ldots, 4\}$ includes the possible contributions of the other player such that if one contributes all of her endowment, the group contributions would still fall short of M_1. $\{5, 6, \ldots, 14\}$ includes the possible contributions of the other player so that coordination at M_1 can be achieved. Similarly, $\{15, 16, \ldots, 20\}$ includes the possible contributions of the other player so that coordination on M_2 can be achieved.

We classified subjects into one of the following types:

1. **Perfectly self-interested**: If a player exhibits selfish preferences, we should make the following contributions for each range of the other player's possible contributions.

   - $\{0, 1, \ldots, 4\}$: Both the less vulnerable and the more vulnerable players contribute 0.
   - $\{5, 6, \ldots, 14\}$: Both the less vulnerable and the more vulnerable players contribute exactly to M_1.
   - $\{15, 16, \ldots, 20\}$: Less (More) vulnerable players contribute exactly to M_1 (M_2).

   Thus, players who follow the pattern mentioned above in contributions are classified as **Perfectly self-interested**. Any contributions in excess of the threshold value would reduce payoffs and thus, Perfectly self-interested player exactly contribute so that sum of contributions is exactly equal to the threshold value and her own payoff is maximised, given the possible contributions of the other group member.
   Less vulnerable participants who were classified as Perfectly self-interested contributed 0 for all possible contributions in the range {0,4}. For possible contributions greater than 4, they made contributions so that the group contribution would exactly equal the value of M_1. They did not target M_2 because they personally do not receive an extra return by contributing to this higher threshold, but would have to contribute a higher amount since the threshold value of M_2 is higher than M_1. Hence, contributing to M_2 would reduce the payoff of a less vulnerable participant compared to if he contributes to M_1, and does not align with the self-centred interests of the Perfectly self-interested player.

   More vulnerable participants who were classified as Perfectly self-interested contributed 0 for all possible contributions in the range {0,4}. For possible contributions in the range {5,14}, they made contributions so that the group contribution would exactly equal the value of M_1. Given the possible contribution of the less vulnerable participant in this range and the maximum contribution of the more vulnerable player, M_2 cannot be achieved. For possible contributions greater than 14, they made contributions so that the group contribution would exactly equal the value of M_2. More vulnerable participants receive the highest payoff when M_2 is achieved, and thus it is in their best interest to coordinate on M_2 if the other member of the group makes sufficient contributions. This contribution pattern is also the same for players in $\mathcal{H}_0$, since they have the same incentive structure as the more vulnerable players in $\mathcal{H}_1$ and $\mathcal{H}_2$.

2. **Imperfectly self-interested with inequality aversion**: A player is classified as Imperfectly self-interested with inequality aversion if the following conditions are fulfilled:

   - The contribution decisions of the player do not exactly match the pattern followed by Perfectly self-interested players.

- The player exhibits aversion to large inequalities between her own and her group member's contributions. She increases her contributions to meet the next higher threshold only if the difference in contributions between her and her group member is lower than the largest possible difference in contributions, which allows the next higher threshold to be achieved. For example, when possible contributions are in the range {5,14}, this type of player will contribute 0 for a possible contribution of 5. Only when possible contributions are greater than 5, typically 9 or 10, and thus the difference in contributions between herself and her group member is lower, this type of player will contribute to meet M_1.

3. **Imperfectly self-interested without inequality aversion**: If the contribution decisions of a player do not exactly match the pattern followed by Perfectly self-interested players, and the player *does not* exhibits aversion to large inequalities between her own and her group member's contributions, we classify the player as *Imperfectly self-interested without inequality aversion.*

4. **Other**: Subjects whose contributions decisions were not readily interpretable were classified as Other. There were only 3 participants in this category out of 118 (see Table 6).

The number of subjects per category in each treatment is tabulated in the Appendix in Table 6.

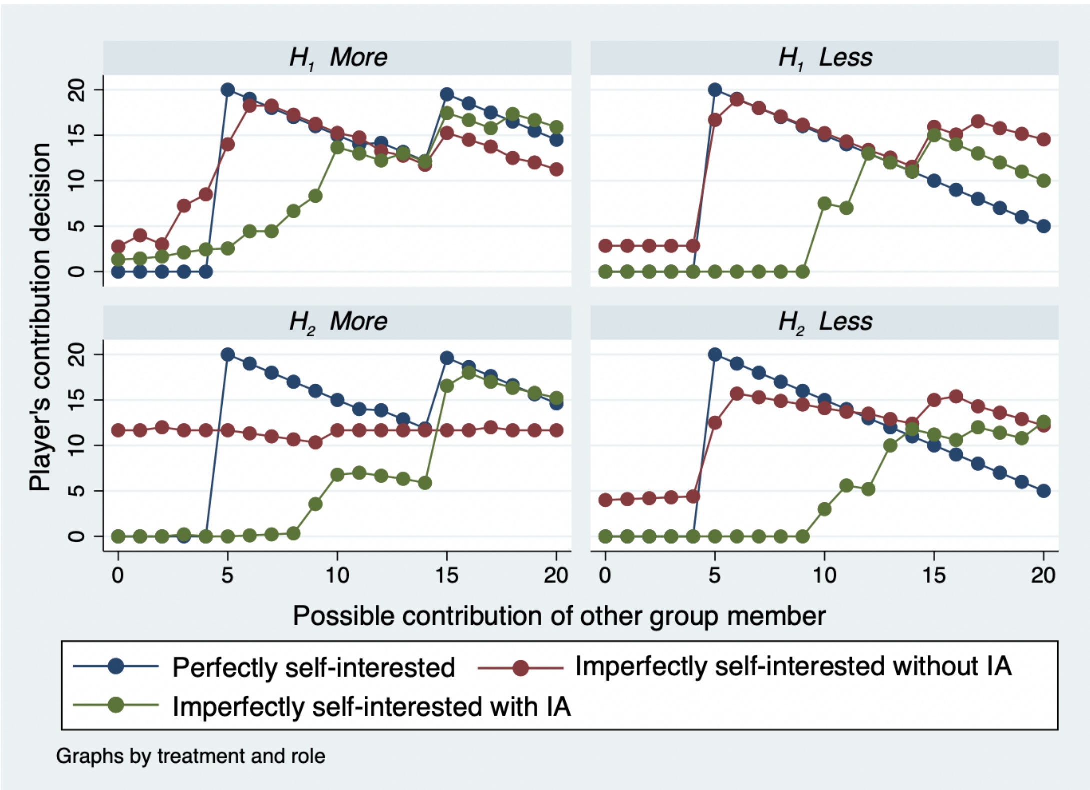

Figure 6: Average contribution pattern of each type by role in $\mathcal{H}_1$ and $\mathcal{H}_2$

In Figure 6, the upper panel shows the average behaviour of different types of players, according to our classification, separately for less vulnerable and more vulnerable players in $\mathcal{H}_1$. Similarly, the lower panel shows the average behaviour of different types of players, separately for less vulnerable and more vulnerable players, in $\mathcal{H}_2$. The blue, red, and green line graphs depict the contribution patterns of Perfectly self-interested, Imperfectly self-interested without inequality aversion, and Imperfectly self-interested with inequality aversion players, respectively. As described earlier, we observe that for possible contributions in the range {0,4}, Perfectly self-interested players contribute 0, regardless of the role. For possible contributions greater than 4, less vulnerable players contribute to achieve exactly M_1. More vulnerable players contribute to achieve exactly M_1 or M_2, depending on the contribution of the less vulnerable player. Comparing the average behaviour of different types of the less vulnerable player, we do not observe any striking differences between $\mathcal{H}_1$ and $\mathcal{H}_2$. Comparing the average behaviour of different types of the more vulnerable player between $\mathcal{H}_1$ and $\mathcal{H}_2$, we observe that Imperfectly self-interested players without inequality aversion tend to be fairly unconditional in their contributions. In addition, among the more vulnerable players, Imperfectly self-interested players with inequality aversion contribute less in $\mathcal{H}_2$ compared to $\mathcal{H}_1$, for possible contributions of the less vulnerable player in the range {5,14}.

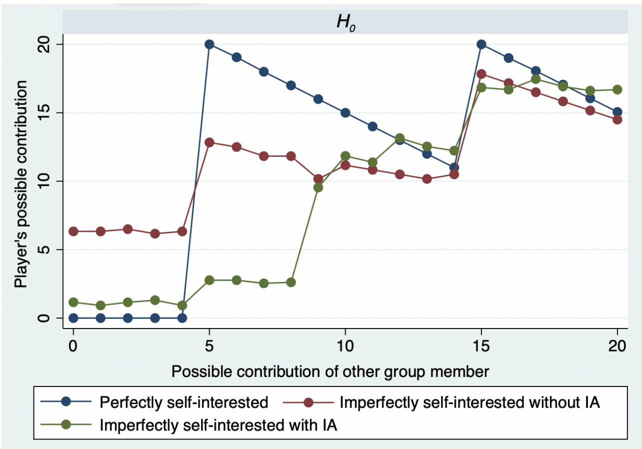

Figure 7: Average contribution pattern of each type in $\mathcal{H}_0$

Figure 7 shows the average behaviour of different types of players, according to our classification, in $\mathcal{H}_0$. There is no heterogeneity between group members in $\mathcal{H}_0$. We observe that Perfectly self-interested players contribute to M_1 and M_2, depending on the possible contribution of the other group member. Imperfectly self-interested players without inequality aversion over-contribute for possible contributions in the range {0,4}, compared to Imperfectly self-interested players with inequality aversion and Perfectly self-interested players. For possible contributions in the range {5,9 }, we observe that players classified as Imperfectly self-interested with IA contribute the lowest, whereas perfectly self-interested players contribute the highest. There is a notable smaller difference in average contributions of different types of players for possible contributions greater than 8.

Table 6: Number of subjects per category in each treatment

	$\mathcal{H}_2$		$\mathcal{H}_1$		$\mathcal{H}_0$
	Less	*More*	*Less*	*More*	All
Perfectly self-interested	5	8	5	6	17
Imperfectly self-interested without inequality aversion	10	3	13	4	6
Imperfectly self-interested with inequality aversion	5	9	2	9	13
Other				1	2
Total	40		40		38

Table 6 tabulates the number of players of each type in each treatment, by role. In all treatments, we observe that there are more *Imperfectly self-interested* (irrespective of inequality aversion) than *Perfectly self-interested* players. We also observed a greater proportion of more vulnerable players who exhibited a clear aversion to inequality compared to less vulnerable players. This is not surprising, given the difference in incentives of both players. In addition, we did not find any relationship between a player's unconditional contribution decision and her type. Most subjects chose a high unconditional contribution between 13 and 20, irrespective of treatment. We also did not find any relationship between our cooperative preferences classification and social preferences classification with the Social Values Orientation measure (Murphy et al. (2011)), which was elicited in *Part III* of the experiment. Most of our subjects, regardless of treatment, were either *Altruistic* (players who sacrifice their own payoff for the gain of others) or *Prosocial* (players who maximise joint payoffs or minimise difference in payoffs). Details are presented in Table 15) in the appendix.

We tried to study how coordination varied across groups composed of different types of players. However, we did not have enough observations to perform a meaningful analysis in this regard. To estimate the effect of a player's type on her contributions in the repeated game, we regressed individual contributions on player type using a random-effects panel data model, controlling for the number of participants in the session and round number. Each unique participant is observed for 10 rounds in the repeated game, making panel data estimation suitable for this analysis. We estimated the effect of a player's type on her contributions the repeated game separately for less vulnerable and more vulnerable players, and separately for each treatment. We largely find no effect of player type on contributions in the repeated game in any treatment or role. Average contribution in $\mathcal{H}_2$, $\mathcal{H}_1$ and $\mathcal{H}_0$ is between 15 and 17 points. For this range, in Figures 6 and 7, we observe little difference in the average contributions between types. Thus, given this average contribution, players' contributions in the repeated game are in accordance with their contributions in the cooperation preferences elicitation task.

9.2 Round wise analysis of contributions

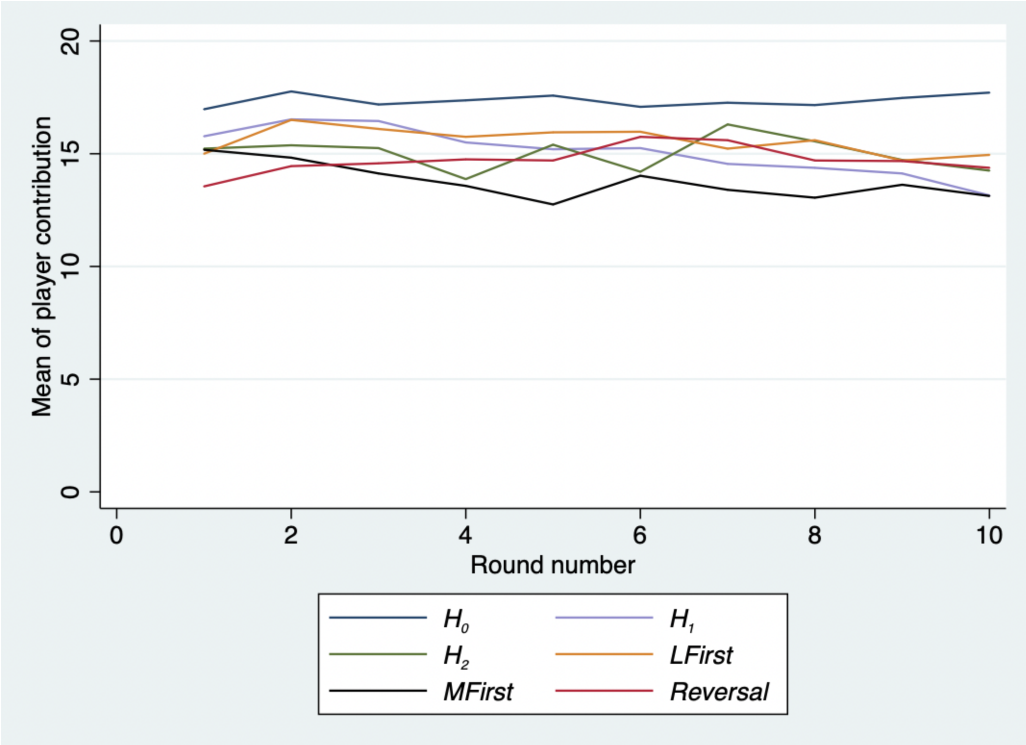

Figure 8: Mean contribution per round by treatment

In Figure 8, we present the mean contribution per round in all treatments. The x-axis denotes round numbers from 1 to 10 and the y-axis denotes the mean contribution of an average player regardless of whether she is a low or more vulnerable player. We observe that contributions are the highest in $\mathcal{H}_0$ consistently in each round. The contributions in other treatments lie in a range similar to one another. Interestingly, we do not find evidence of a significant decline in contributions with repetition, as has been observed in linear public goods games.

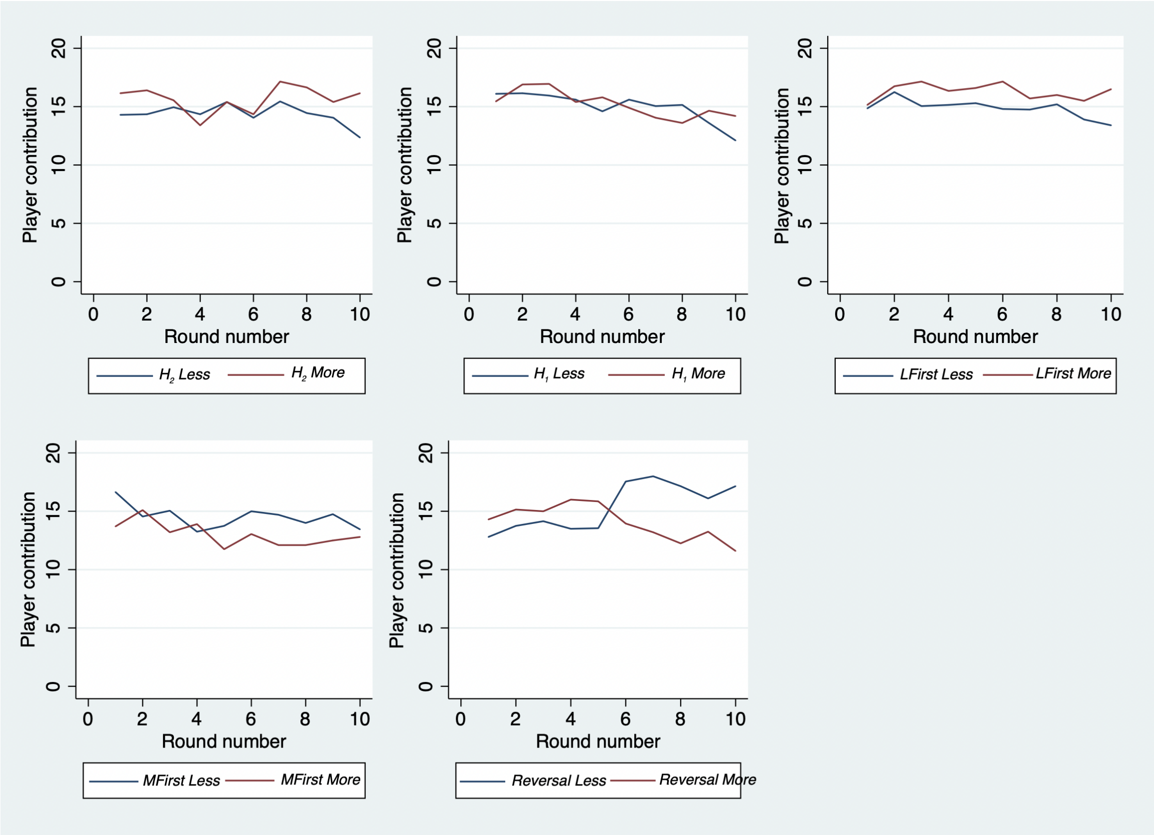

Figure 9: Mean contribution per round by role

In Figure 9, we present the mean contribution per round in every treatment by role. The blue and red lines in each graph show the contributions of the less and more vulnerable players, respectively. From the top left to the right, we present the graphs for $\mathcal{H}_2$, $\mathcal{H}_1$ and *LFirst*. From the bottom left to the right, we present the graphs for *MFirst* and *Reversal*.

Consistent with our above findings, we do not observe a significant decline in contributions of the less or more vulnerable participants. The most notable observation is in the graph for *Reversal*, where we observe a significant increase in the difference in contributions after the reversal. We do not observe this pattern in other treatments. This observation is consistent with our findings of the panel data regressions.

We summarise our **additional findings** below:

1. We find that quite a number of less vulnerable participants showed preferences for coordination in M_2, despite the lack of direct monetary incentive. This is in line with the coordination rates on M_2 being non-zero in $\mathcal{H}_2$ and $\mathcal{H}_1$ in the repeated game.

2. In contrast to previous findings in the literature, we do not observe a significant decline in contributions over repetition.

10 Conclusion and Discussion

To study how differential climate vulnerability affects coordination on climate change mitigation targets, we present a novel threshold public goods game with heterogeneous thresholds and payoffs. We use a threshold public goods game to model target global effort levels. In addition, we make a novel contribution by introducing heterogeneity in thresholds. The use of threshold public goods games as well as threshold heterogeneity closely capture the real-world situation we are interested to study. We conduct an experiment to gain insights on how coordination behaviour is affected by this heterogeneity. We find that environments with no heterogeneity are significantly better at sustaining high level coordination. However, heterogeneity impedes high-level coordination. Heterogeneity also leads to significantly lower contributions and reduces payoffs.

As a natural next step, we explored possible solutions that can improve coordination, specifically at the higher threshold. We consider two sequential move treatments, where an exogenously pre-determined player makes the contribution decision first in every round. We also consider a treatment in which the roles are reversed. In particular, halfway through the ten rounds, less vulnerable players switch roles with more vulnerable players and will now receive the payoffs of the more vulnerable (and vice-versa). We find that coordination on the higher threshold is improved when the less vulnerable contributes first. There is also a significant improvement in the payoffs of the more vulnerable player. Thus, allowing the less vulnerable player to contribute first is a potentially good solution to the lack of high-level coordination resulting from heterogeneity. We do not find any significant effect of the other solution treatments on coordination rates, contributions or payoffs.

Despite no differences in coordination, we find that there is a significantly higher difference in contributions after the reversal, compared to the non-reversal benchmark. This is an interesting finding. Given that there is no difference in coordination, one would not expect a significant increase in the difference in contributions between players. The less vulnerable player contributes more after the reversal, but the more vulnerable contributes equally less, essentially removing any opportunity for improved coordination. Thus, in contrast to the generally positive effect of reversal found in the literature, we find that reversal may have differing impacts on coordination, depending on the individual's initial role. Further research may shed light on the reasons for this increased difference.

This study offers important insights on human behaviour and coordination decisions in the presence of heterogeneity[14]. From the point of view of a policy maker, sequential contributions

[14]We do not claim that our results can be exactly extrapolated to the actual problem, but rather they offer understanding on identifying the problematic reason for non-coordination in the presence of heterogeneity, and what can potentially be done by a policy maker to improve coordination. The problem still has a human component, which is why insights on human behaviour will nevertheless be helpful.

where the least vulnerable contribute first can potentially be a viable solution to the lack of consensus on high global effort. Surprisingly, despite strong suggestions in the literature about the efficacy of role-reversal in enhancing pro-social behaviour like coordination, we find no such evidence in our data. Thus, despite the fact that some less vulnerable countries may be facing the brunt of climate change in recent times, it may not necessarily translate to a sustained consensus on high global mitigation effort. Moreover, further research needs to be conducted to understand how the effect of reversal may vary depending on one's initial role. This research would have important implications in the evaluation of role-reversal as a coordination enhancing mechanism.

Although we find a potential solution to the problem of coordination with high effort, there is a great area of potential research where this solution can be applied to real-world climate negotiations. The problem of climate change and its heterogeneous effects are more pressing than ever. Continued research on additional solutions, as well as extending our model to capture additional aspects of climate change, will help us effectively tackle this extraordinary problem.

References

Abdellaoui, M., Bleichrodt, H., & L'Haridon, O. (2008, 06). A tractable method to measure utility and loss aversion under prospect theory. *Journal of Risk and Uncertainty*, *36*, 245-266.

Ajzen, I. (1991, 12). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179-211. doi: 10.1016/0749-5978(91)90020-T

Ajzen, I. (2002, 04). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, *32*, 665-683. doi: 10.1111/j.1559-1816.2002.tb00236.x

Ajzen, I., & Fishbein, M. (1980, 01). Understanding attitudes and predicting social behavior. *Journal of Experimental Social Psychology*, *278*.

Barrett, S., & Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences*, *109*(43), 17372-17376.

Barrett, S., & Dannenberg, A. (2013, 12). Sensitivity of collective action to uncertainty about climate tipping points. *Nature Climate Change*, *4*.

Benoît, J.-P., & Krishna, V. (1985, 02). Finitely repeated games. *Econometrica*, *53*, 905-22.

Bolle, F. (2014). On a class of threshold public goods games : With applications to voting and the kyoto protocol.

Bolle, F. (2016, 07). Passing the buck on the acceptance of responsibility. *Research in Economics*, *71*.

Bolton, G. (1991, 01). A comparative model of bargaining: Theory and evidence. *American Economic Review*, *81*, 1096-136.

Bolton, G., & Ockenfels, A. (2000, 02). Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*, 166-193.

Buckley, E., & Croson, R. (2006, 05). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics*, *90*, 935-955.

Burke, M., Davis, W., & Diffenbaugh, N. (2018, 05). Large potential reduction in economic damages under un mitigation targets. *Nature*, *557*.

Burke, M., & Hsiang, S. (2015, 10). Global non-linear effect of temperature on economic production. *Nature*, *527*.

Burlando, R., & Guala, F. (2005, 02). Heterogeneous agents in public goods experiments. *Experimental Economics*, *8*, 35-54.

Chan, K., Mestelman, S., Moir, R., & Muller, R. (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics*, *2*(1), 5-30.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*(3), 817–869.

Chewning, E., Coller, M., & Laury, S. (2001, 03). Voluntary contributions to a multiple threshold public good. *Research in Experimental Economics*, *8*, 47-83.

Christensen, R., & Knezek, G. (2015, 10). The climate change attitude survey: Measuring middle school student beliefs and intentions to enact positive environmental change. *International Journal of Environmental and Science Education*, *10*, 773-788.

Coats, J., Gronberg, T., & Grosskopf, B. (2009, 02). Simultaneous versus sequential public good provision and the role of refunds — an experimental study. *Journal of Public Economics*, *93*, 326-335.

Cox, J., & Friedman, D. (2003, 01). A tractable model of reciprocity and fairness. *SSRN Electronic Journal*.

Croson, R., & Marks, M. (1999, 08). The effect of heterogeneous valuations for threshold public goods: An experimental study. *Risk Decision and Policy*, *4*, 99-115.

Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2014, 05). On the provision of public goods with probabilistic and ambiguous thresholds. *Environmental and Resource Economics*, *61*.

Dannenberg, A., Riechmann, T., Sturm, B., & Vogt, C. (2007, 07). Inequity aversion and individual behavior in public good games: An experimental investigation. *ZEW discussion paper*, *47*.

Dannenberg, A., Sturm, B., & Vogt, C. (2010, 01). Do equity preferences matter for climate negotiators? an experimental investigation. *Environmental and Resource Economics*, *47*, 91-109.

Diffenbaugh, N. S., & Burke, M. (2019). Global warming has increased global economic inequality. *Proceedings of the National Academy of Sciences*, *116*(20), 9808-9813.

Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, *47*(2), 268-298.

Falk, A., & Fischbacher, U. (2002, 01). The economics of reciprocity: Evidence and theory. In

(p. 207-233).
Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, *54*(2), 293-315.
Fehr, E., Kirchsteiger, G., & Riedl, A. (1993, 02). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics*, *108*, 437-59.
Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.
Feige, C., Ehrhart, K.-M., & Krämer, J. (2018, 06). Climate negotiations in the lab: A threshold public goods game with heterogeneous contributions costs and non-binding voting. *Environmental and Resource Economics*, *70*, 1-20.
Fellner, G., Kröger, S., & Seki, E. (2019, 10). Public good production in heterogeneous groups: An experimental analysis on the relation between external return and information. *Journal of Behavioral and Experimental Economics*, *84*, 101481.
Fischbacher, U., Gächter, S., & Fehr, E. (2001, 06). Are people conditionally cooperative? evidence from a public good experiment. *Economics Letters*, *71*, 397-404.
Fischbacher, U., Teyssier, S., & Schudy, S. (2012, 01). Heterogeneous reactions to heterogeneity in returns from public goods. *Social Choice and Welfare*, *43*.
Franzke, C. L. E. (2014). Nonlinear climate change. *Nature Climate Change*, *4*(6), 423-424.
Hashim, M., Maximiano, S., & Kannan, K. (2011, 07). Information targeting and coordination: An experimental study. *SSRN Electronic Journal*.
Huang, W., Braithwaite, I., Charlton-Perez, A., Sarran, C., & Sun, T. (2022, 03). Non-linear response of temperature-related mortality risk to global warming in england and wales. *Environmental Research Letters*.
Kompas, T., Ha, P., & Che, N. (2018, 07). The effects of climate change on gdp by country and the global economic gains from complying with the paris climate accord. *Earth's Future*, *6*. doi: 10.1029/2018EF000922
Lange, A., Schmitz, J., & Schwirplies, C. (2022, 02). Inequality, role reversal and cooperation in multiple group membership settings. *Experimental Economics*, *25*, 68–110.
Mcginty, M., & Milam, G. (2013, 04). Public goods provision by asymmetric agents: Experimental evidence. *Social Choice and Welfare*, *40*.
Milinski, M., Sommerfeld, R., Krambeck, H.-J., Reed, F., & Marotzke, J. (2008, 03). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 2291-4.
Murphy, R., Ackermann, K., & Handgraaf, M. (2011, 12). Measuring social value orientation. *Judgment and Decision Making*, *6*, 771-781.
Normann, H.-T., & Rau, H. (2014, 01). Simultaneous and sequential contributions to step-level public goods: One vs. two provision levels. *Journal of Conflict Resolution*, *59*.
Reindl, I. (2022, 08). Wealth and vulnerability to climate change: An experimental study on burden sharing among heterogeneous agents. *Environmental and Resource Economics*,

82.

Report, I. M. F. (2017). *World economic outlook october 2017*.

Rial, J., Pielke Sr, R., Beniston, M., Claussen, M., Canadell, J., Cox, P., ... Salas, J. (2004, 07). Nonlinearities, feedbacks and critical thresholds within the earth's climate system. *Climatic Change*, *65*.

Schneider, S. (2004, 10). Abrupt non-linear climate change, irreversibility and surprise. *Global Environmental Change*, *14*, 245-258.

Selten, R. (1967, 01). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. *Beiträge Zur Experimentellen Wirtschaftsforschung*, 136-168.

Sheppard, B., Hartwick, J., & Warshaw, P. (1988, 02). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research*, *15*. doi: 10.1086/209170

Sinatra, G., Kardash, C., Taasoobshirazi, G., & Lombardi, D. (2012, 01). Promoting attitude change and expressed willingness to take action toward climate change in college students. *Instructional Science*, *40*, 1-17. doi: 10.1007/s11251-011-9166-5

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011, 07). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 11825-9.

Trust, reciprocity, and social history. (1995). *Games and Economic Behavior*, *10*(1), 122-142.

Winges, M., Hutfils, M.-L., & Eckstein, D. (2019, 12). Global climate risk index 2019. who suffers most from extreme weather events? weather-related loss events in 2017 and 1998 to 2017.

11 Appendix

11.1 Proof that strategy $\mathcal{S}$ is a Subgame Perfect Nash Equilibrium in $\mathcal{H}_2$ and $\mathcal{H}_1$

Consider the following notation. x^i and x^j denotes the contribution of player i and j respectively, τ and t denote time periods, L and M denote the less and more vulnerable players respectively. Consider the strategy $\mathcal{S}$, $\forall\ \tau < t < 10$, $\forall\ i, j \in \{L, M\}$, $i \neq j$:

$$x^i(h_t) = \begin{cases} M_2 - x^j & \text{if } \ x^j_\tau \geq M_2 - x^i_\tau, \\ 0 & \text{otherwise} \end{cases}$$

$$x^i(h_{10}) = \begin{cases} M_1 - x^j & \text{if } \ x^j_9 \geq M_2 - x^i_9, \\ 0 & \text{otherwise} \end{cases}$$

The strategy is a grim-trigger strategy. According to this strategy, both players coordinate on M_2 in all but the last round. Coordinate on M_1 using the stage game Nash Equilibrium if there has been no deviation in the previous rounds. If there has been a deviation from M_2 in any of the rounds before the last round, contribute 0 for all remaining rounds, including the last round. Let us consider the case where the less vulnerable player deviates to M_1 in the ninth round. The more vulnerable player is always strictly better off contributing to M_2, and has no incentive to deviate to M_1. So, to check for deviations, it is sufficient to consider the less vulnerable player's payoffs. If the less vulnerable player deviates to M_1 in the ninth round, he triggers the punishment, and both players contribute 0 in the tenth. However, if he had continued to coordinate on M_2 in the ninth round, both players would have also coordinated on M_1 in the last round. The one-off extra payoff from deviation to the less vulnerable is 10 points. However, in each of the following rounds, he loses out on receiving 32 points which he could have gotten had he coordinated on M_2. In other words, for any contribution of the more vulnerable player, the sum of the payoffs of the less vulnerable player from coordination on M_2 followed by M_1 is strictly greater than his payoff from coordination on M_2, followed by 0. Thus, no player has any incentive to deviate from the strategy to coordinate on M_2 till the ninth round and M_1 in the tenth round. In case of deviation, no player has any incentive to deviate from the punishment strategy of zero contributions till the last round because $x^m = x^l = 0$ is a Pure Strategy Nash Equilibrium. Therefore, coordination on M_2 can be sustained in $\mathcal{H}_2$ and $\mathcal{H}_1$ using the aforementioned strategy, and it is a Subgame Perfect Nash Equilibrium.

11.2 Proof that strategy $\mathcal{S}$ is a Subgame Perfect Nash Equilibrium in *Reversal*

The above proof also holds true for the *Reversal* treatment. No player will want to deviate in the ninth round. The more vulnerable player is strictly better off contributing to M_2 and the less vulnerable player is threatened by zero contributions in the event of a deviation, which also gives him no incentive to deviate from M_2. By the same argument, no player will deviate in the 8th, 7th or 6th round, that is, after the reversal. In the last round before the reversal or in any of the preceding rounds, the less vulnerable will have no incentive to deviate for the aforementioned reason. Furthermore, the reversal will make her a more vulnerable player, which means that the punishment of zero contributions will hurt her more. Therefore, strategy $\mathcal{S}$ can be sued to sustain coordination on M_2 in *Reversal*.

11.3 Elicitation of Risk and Loss preferences

We elicited risk and loss preferences using a decision-making task set down by Abdellaoui et al. (2008). To elicit risk preferences, we elicited the certainty equivalence of a lottery of the form (X,0.5; 0,0.5) and compared it to the expected value of the lottery. We elicited the certainty equivalence of three such lotteries. The value of X in the three lotteries was 12, 24 and 36. We chose values of X guided by the range of possible payoffs which a player can earn in *Parts I* and *II*. If the certainty equivalence of at least two of the three lotteries was greater (lesser) than the expected value of the lottery, the participant was classified as risk-averse (risk-seeking). If the certainty equivalence of at least two out of the three lotteries was equal to the expected value of the lottery, the participant was classified as risk-neutral. Otherwise, participants were classified as mixed. To elicit the certainty equivalence of a lottery, we used the bisection method. We illustrate its use with an example below:

Table 7: An illustration of the bisection method

Iteration	Certainty Equivalence elicitation	Loss Aversion elicitation
1	$(\mathbf{12}, \mathbf{0.5}; \mathbf{0}, \mathbf{0.5})$ vs 6	$(\mathbf{6}, \mathbf{0.5}; \mathbf{-6}, \mathbf{0.5})$ vs 0
2	$(12, 0.5; 0, 0.5)$ vs **9**	$(6, 0.5; -9, 0.5)$ vs **0**
3	$(12, 0.5; 0, 0.5)$ vs **7**	$(6, 0.5; -10, 0.5)$ vs **0**
4	$(12, 0.5; 0, 0.5)$ vs **6**	$(6, 0.5; -11, 0.5)$ vs **0**
Indifference	6	-11

An iteration is a choice question. With every successive iteration, we get closer to the certainty equivalence and the loss aversion preference. In iteration 1, participants choose between a lottery and a sure payoff, which is the expected value of the lottery. The participant's choice is highlighted in bold. Depending on the participant's choice, the value of the sure payoff is increased or decreased until the certainty equivalence is narrowed down. The size of the change is half the size of the change in the previous iteration. Furthermore, we rounded down the sure payoff to obtain an integer value. Continuing this method for 4 iterations was enough to yield the indifference value. For the elicitation of loss aversion, the first iteration presented a choice between a sure payoff 0 and a lottery (Y,0.5; Z,0.5). Y is the indifference value of the lottery (X,0.5; 0,0.5) and Z=−Y. The bisection method illustrated earlier is followed to obtain the indifference value of the mixed lottery (Y,0.5; Z,0.5). In iteration 1, if lottery (Y,0.5; Z,0.5) is chosen over the sure payoff, the value of Z in the second iteration is the midpoint of 2Z and Z. The remaining iterations follow the bisection method illustrated earlier as usual. We obtain the loss aversion coefficient using the procedure described in **??**. We find that there is no difference in the percentage of participants exhibiting preferences for loss-aversion or gain-seeking between treatments.

11.4 Measuring Social Preferences

Social preferences were elicited using the six primary slider tasks as outlined in Murphy et al. (2011). Each participant was matched anew with another participant in the room, who was referred to as 'the other player'. Each slider task consisted of allocating points between one's own self and the other player. A person's SVO, which is a measure of her social preference, was computed as follows. The mean allocation for one's own self (denoted by A_s) and the mean allocation for the other player (denoted by A_o) is computed. Then 50 is subtracted from each mean in order to "shift" the base of the resulting angle to the center of the circle (50, 50) rather than having its base start at the Cartesian origin (0,0). Finally, we compute the inverse tangent of the ratio of the means and obtain a single index of a person's SVO. Thus,

$$\text{SVO}^{\circ} = \arctan\left(\frac{A_o - 50}{A_s - 50}\right) \tag{6}$$

Depending on the SVO degrees thus computed, we can classify a participant into one of the following social preferences categories, following the bounds of the SVO angles provided by Murphy et al. (2011).

- **Altruist**: An altruistic person would choose to maximise the allocation of the other player. Her SVO° would be an angle greater than 57.15°.
- **Prosocial**: A prosocial person would choose to maximise joint gain and would be intolerant to inequality. Her SVO° would be an angle between 22.45° and 57.15°.
- **Individualist**: An individualistic person would choose to maximise her own allocation. Her SVO° would be an angle between −12.04° and 22.45°.
- **Competitive**: A competitive person would choose to maximise the difference in her own allocation and the other member's allocation. Her SVO° would be an angle less than −12.04°.

Table 8: SVO types in each treatment

Treatments	**Altruist**	**Prosocial**	**Competitive**
$\mathcal{H}_2$	27	11	0
$\mathcal{H}_1$	28	10	2
$\mathcal{H}_0$	25	12	1
LFirst	26	13	1
MFirst	29	9	2
Reversal	31	7	2

We observe that about 70% of our participants were Altruists, regardless of treatment. We find some Prosocial participants, very few Competitive participants and no Individualistic participants, also irrespective of treatment.

11.5 Climate Change Attitudes Survey

We also elicited attitudes towards climate change using a modified version of the Climate Change Attitudes survey detailed in Christensen and Knezek (2015). It consists of 15 statements, and participants indicate their most preferred option on a five-point Likert scale. The authors classify the questions in the survey as belief-based on intention-based. The questions we used in this survey are outlined below. Belief-based questions are marked as *belief*, and intention-based questions are marked as *intention*.

1. I believe our climate is changing (*belief*).
2. I am concerned about global climate change (*belief*).
3. I believe there is evidence of global climate change. (*belief*)
4. Global climate change will impact our environment in the next 10 years. (*belief*)
5. I would support my country helping other countries fight global climate change. (*intention*)
6. The actions of individuals can make a positive difference in global climate change. (*belief*)
7. Human activities cause global climate change. (*belief*)
8. Climate change has a negative effect on our lives. (*belief*)
9. We cannot do anything to stop global climate change. (*intention*)
10. I would want the government in my country to take global climate change seriously. (*intention*)
11. Combined global effort is key to tackling global climate change. (*belief*)
12. I think most of the concerns about environmental problems have been exaggerated. (*intention*)
13. Things I do have no effect on the quality of the environment. (*intention*)
14. It is a waste of time to work to solve environmental problems. (*intention*)
15. There is not much I can do that will help solve environmental problems. (*intention*)

The questions were classified as *belief* or *intention* based, in accordance with published theory and research-based recommendations (Ajzen (1991), Ajzen (2002), Ajzen and Fishbein (1980), Sheppard, Hartwick, and Warshaw (1988), Sinatra, Kardash, Taasoobshirazi, and Lombardi (2012) and other related studies).

Table 9: Summary statistics: $\mathcal{H}_2$, $\mathcal{H}_1$ and $\mathcal{H}_0$

	$\mathcal{H}_2$	$\mathcal{H}_1$	$\mathcal{H}_0$
# rounds M_1 achieved	3.3	5	1.37
	(3.08)	(3.52)	(2.11)
# rounds M_2 achieved	5.2	4.05	8.32
	(3.44)	(3.89)	(2.96)
Mean contribution: *Less*	14.37	14.99	34.71
	(3.44)	(2.75)	(3.26)
Mean contribution: *More*	15.66	15.19	
	(3.77)	(2.86)	
Mean payoff: *Less*	32.83	33.30	62.44
	(7.46)	(5.11)	(10.60)
Mean payoff: *More*	25.20	25.10	
	(7.30)	(6.37)	

This table shows the summary statistics of mean individual contributions and payoffs in a round, number of rounds M_1 is achieved (# rounds M_1 achieved) and the number of rounds M_2 is achieved (# rounds M_2 achieved) out of a total of 10 rounds, in treatments $\mathcal{H}_2$ and $\mathcal{H}_1$. Since there are no less or more vulnerable players in $\mathcal{H}_0$, we report mean group contributions and payoff in $\mathcal{H}_0$, instead of individual contributions and payoffs. Standard error in parentheses.

Table 10: Summary statistics: *LFirst* and *MFirst*

	LFirst	*MFirst*
# rounds M_1 achieved	1.6	4.6
	(1.81)	(3.51)
# rounds M_2 achieved	7.4	4.4
	(3.25)	(4.09)
Mean contribution: *Less*	14.86	14.51
	(3.11)	(4.21)
Mean contribution: *More*	16.28	13.02
	(3.39)	(5.07)
Mean payoff: *Less*	34.09	34.28
	(3.88)	(3.05)
Mean payoff: *More*	29.45	27.72
	(6.43)	(5.64)

This table shows the summary statistics of mean individual contributions and payoffs in a round, the number of rounds M_1 is achieved (# rounds M_1 achieved) and number of rounds M_2 is achieved (# rounds M_2 achieved) out of a total of 10 rounds, in treatments *LFirst* and *MFirst*. Standard error in parentheses.

Table 11: Summary statistics: *Reversal* and $\mathcal{H}_2$ by first 5 and last 5 rounds

	Reversal		$\mathcal{H}_2$	
	First 5 rounds	Last 5 rounds	First 5 rounds	Last 5 rounds
# rounds M_1 achieved	2.35	1.8	1.7	2.45
	(1.49)	(1.70)	(1.68)	(1.73)
# rounds M_2 achieved	2.15	1.5	2.35	2.85
	(2.05)	(1.76)	(2.35)	(2.85)
Mean contribution: *Less*	13.55	17.19	14.67	14.07
	(4.85)	(4.42)	(3.40)	(4.20)
Mean contribution: *More*	15.26	12.85	15.38	15.94
	(4.58)	(4.85)	(4.33)	(3.96)
Mean payoff: *Less*	32.37	24.03	31.89	33.77
	(6.62)	(7.98)	(8.27)	(7.29)
Mean payoff: *More*	21.81	35.95	24.23	26.19
	(7.44)	(6.47)	(8.16)	(8.04)

This table shows the summary statistics of mean individual contributions and payoffs in a round, the number of rounds M_1 is achieved (# rounds M_1 achieved) and the number of rounds M_2 is achieved (# rounds M_2 achieved) out of a total of 10 rounds in treatment *Reversal*, separately for the first five rounds (before reversal) and the last five rounds (after reversal). We have also provided the same statistics for the benchmark $\mathcal{H}_2$ treatment for comparison. Standard errors in parentheses.

Table 12: Strategies over 10 rounds

Strategy	Description	$\mathcal{H}_2$	$\mathcal{H}_1$	$\mathcal{H}_0$
1	M_1 or 0 contributions in all rounds	2	5	1
2	M_2 in rounds 1-10	7	4	15
3	$\mathcal{S}$	3	1	0
4	M_2 in some rounds, then M_1 forever	3	5	0
5	Other	5	5	3
	N	20	20	19

This table shows the number of groups in each treatment $\mathcal{H}_2$, $\mathcal{H}_1$ and $\mathcal{H}_0$ which followed the corresponding strategy. Strategies 2, 3 and 4 include strategies in which coordination on M_2 first takes place in the first round or after the first few rounds. This is done to obtain a clear classification as well as to allow subjects to take the first initial rounds to familiarise themselves with the game and their group member. Strategy $\mathcal{S}$ as defined earlier consists of coordination on M_2 in rounds 1-9, M_1 in round 10 or 0 contributions if there has been a deviation in rounds 1-9.

Table 13: Unconditional contribution by category and treatment

Unconditional contribution	$\mathcal{H}_0$		$\mathcal{H}_1$				$\mathcal{H}_2$				Total
	All players		Less		More		Less		More		
	P	I	P	I	P	I	P	I	P	I	
0	0	0	0	0	0	0	0	1	0	1	2
5	0	0	0	0	0	0	1	1	1	0	3
6	0	0	0	0	0	1	1	0	0	0	2
7	0	0	0	0	0	0	0	0	0	1	1
9	0	0	0	0	0	1	0	0	0	0	1
10	0	0	1	1	0	2	0	1	0	0	5
12	0	1	0	0	0	0	0	0	1	0	2
13	1	2	1	1	1	1	1	3	2	0	13
14	0	0	0	2	0	2	0	0	0	0	4
15	2	0	1	0	0	0	0	1	1	1	6
16	0	2	0	1	0	0	0	0	0	1	4
17	1	2	0	3	0	1	1	2	0	1	11
18	11	10	2	6	3	4	0	5	3	3	47
19	0	1	0	0	0	0	0	0	0	0	1
20	2	1	0	1	2	1	1	1	0	4	13
Total	17	19	5	15	6	13	5	15	8	12	115

This table shows the unconditional contribution of subjects observed in each category by treatment and type. Under columns $\mathcal{H}_1$ and $\mathcal{H}_2$, there are sub-columns for Less and More vulnerable players, each of which is then divided into sub-columns P and I which stand for *Perfectly self-interested and Imperfectly self-interested* respectively. Since there is no difference in vulnerability in the $\mathcal{H}_0$ treatment, its column is divided into P and I only.

Table 14: Types and social preferences ala Social Values Orientation measure

SVO type	$\mathcal{H}_2$		$\mathcal{H}_1$				$\mathcal{H}_0$			
	All players		Less		More		Less		More	
	P	I	P	I	P	I	P	I	P	I
Altruistic	13	10	4	11	3	9	4	9	4	10
Competitive	1	0	0	0	1	1	0	0	0	0
Prosocial	3	9	1	4	2	3	1	4	4	2
Total	17	19	5	15	6	13	5	13	8	12

This table shows a subject's type according to the Social Values Orientation (SVO) measure by treatment and type. Under columns $\mathcal{H}_1$ and $\mathcal{H}_2$, there are sub-columns for Less and More vulnerable players, each of which is then divided into sub-columns P and I which stand for *Perfectly self-interested and Imperfectly self-interested* respectively. Since there are no differently vulnerable players in the $\mathcal{H}_0$ treatment, its column is divided into P and I only.

Table 15: Effect of player type on contributions in the repeated game

Type	Subsample				
	$\mathcal{H}_2$, *Less*	$\mathcal{H}_2$, *More*	$\mathcal{H}_1$, *Less*	$\mathcal{H}_1$, *More*	$\mathcal{H}_0$
Type I	2.58	3.47	2.78^{**}	−0.77	−0.77
	(1.80)	(3.27)	(1.35)	(1.98)	(0.83)
Type II	0.92	3.30	0.63	−1.12	3.19
	(2.22)	(2.07)	(2.15)	(1.62)	(0.63)
Type I	1.66	0.16	2.15	0.34	−1.09
(Base: Type II)	(2.11)	(3.19)	(1.95)	(1.85)	(0.88)
N	200	200	200	200	380

For the purpose of this table, we denote Imperfectly self-interested players without inequality aversion as Type I and Imperfectly self-interested players with inequality aversion as Type II. Base category, unless otherwise stated, is Perfectly self-interested. Note: Standard errors in parentheses. $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.