KCG Working Paper

Challenging Conventional Wisdom:
Experimental Evidence on Heterogeneity and Coordination in Avoiding a Collective Catastrophic Event

Israel Waichman, Till Requate, Markus Karde and Manfred Milinski

No. 19 | September 2019

KCG Working Paper No. 19 | September 2019

KIEL CENTRE FOR GLOBALIZATION

Challenging Conventional Wisdom:

Experimental Evidence on Heterogeneity and Coordination in Avoiding a Collective Catastrophic Event

Israel Waichman, Till Requate, Markus Karde and Manfred Milinski

Abstract: Avoiding a catastrophic climate change event is a global public good characterized by several dimensions, notably heterogeneity between the parties involved. It is often argued that such heterogeneity between countries is a major obstacle to cooperative climate policy. We challenge this belief by experimentally simulating two important heterogeneities, in wealth and loss, when dangerous climate change occurs. We find that under loss heterogeneity the success rate in achieving sufficient mitigation to prevent catastrophic climate change is higher than with homogeneous parties. We also observe that neither endowment heterogeneity nor the combination of endowment and loss heterogeneities lead to significantly different success rates than with homogeneous parties. Our findings suggest that heterogeneities may facilitate rather than hinder successful international climate policy negotiations.

Keywords: global public good, change negotiation, collective-risk social dilemma, endowment heterogeneity, loss heterogeneity, focal point

JEL Classification: C92, D74, H41, Q54

KCG Working Paper No. 19 | September 2019

Israel Waichman
Bard College Berlin
Platanenstr. 24
13156 Berlin
Germany
i.waichman@berlin.bard.edu

Till Requate
Kiel Centre for Globalization
Department of Economics
University of Kiel
Olshausenstraße 40
24118 Kiel
Germany
requate@economics.uni-kiel.de

Markus Karde
Department of Economics
University of Kiel
Olshausenstraße 40
24118 Kiel
Germany
karde@economics.uni-kiel.de

Manfred Milinski
Department of Evolutionary Ecology
Max Planck Institute for Evolutionary Biology
August-Thienemann-Straße 2
24306 Plön
Germany
milinski@evolbio.mpg

About the Kiel Centre for Globalization (KCG): KCG is a Leibniz Science Campus initiated by Christian-Albrechts University of Kiel and Kiel Institute for the World Economy. It works on an interdisciplinary research agenda that evaluates the proliferation of global supply chains as an important aspect of globalization. To this end, the KCG brings together researchers from economics, ethics and management science. KCG is financially supported by the Leibniz Association and the State Government of Schleswig-Holstein. More information about KCG can be found here: www.kcg-kiel.org.

1 Introduction

Within the scientific community it is now widely accepted that the man-made release of greenhouse gases (GHGs) into the atmosphere - notably CO_2 - causes and accelerates global warming (IPCC, 2007, 2013). There is even concern that if GHG gas accumulation exceeds a certain threshold, irreversible effects will ensue, and catastrophic outcomes may occur (Marotzke, 2000; Alley et al., 2003; IPCC, 2013). However, although since the early 1990's (mostly European) countries have engaged in CO_2 emission mitigation, overall CO_2 emissions into the atmosphere continue to increase (IPCC, 2007, 2013). From an economic and social point of view this is not surprising, since climate change mitigation can be considered as a contribution to a public good (Milinski et al., 2006; Raihani and Aitken, 2011), and from economic theory and countless public-good game experiments (see, for example, the review by Ledyard, 1995) we know that the individual benefits fall short of the social (or group) benefit. Since in the short term CO_2 mitigation will inevitably reduce output and slow down economic growth, most countries prefer to do nothing and at best free-ride on others' mitigation efforts instead of reducing their own CO_2 emissions.

While in standard public-good games with continuous payoff functions the (usually unique) equilibrium is inefficient, the introduction of an irreversible threshold that models a 'tipping point' in the earth climate system (Lenton et al., 2008; IPCC, 2013) can change the game substantially. Milinski et al. (2008) suggest modeling a situation like this as a modified threshold public-good game, which they refer to as the "collective-risk social dilemma". This game has two distinct features. First, the outcome is only realized after several contribution rounds. Second, the more wealth a player has accumulated, the more he/she stands to lose in the case of a catastrophe. Typically, such a game has multiple equilibria, one with no party contributing to the public good and several equilibria where the risk of a catastrophe is avoided. The latter differ with respect to burden sharing and hence with respect to the distribution of wealth between the parties. Note that this dilemma still holds even after the "Paris agreement" and was one reason for US president Trump to quit the agreement, as from his point of view the burden sharing plan was unfair. The Paris agreement acknowledges the heterogeneity between developed and developing countries and specifies an amount to be transferred annually to developing countries ($100 billion), but it does not set individual emission targets or sanctions for non-compliance.[1] Therefore, the coordination problem about which country

[1] See Jacquet and Jamieson (2016) for an excellent commentary on the Paris agreement.

will reduce how much to meet the 2°C target still remains.

Our study centers on two prominent heterogeneities between parties in preventing a catastrophic event. It is motivated by the fact that in the climate change social dilemma the parties involved differ in two important respects: (i) wealth and (ii) the extent to which they will be affected by climate change in the future (Frankhauser, 1994; Tol, 2009).[2] In fact, climate-change experts predict that the effect of global warming may be especially devastating for poor countries located in low-latitude areas with high initial temperatures (Mendelsohn et al., 2006; IPCC, 2013). Moreover, poor countries are less well equipped to deal with catastrophic climate-change consequences (adaptation) than their rich counterparts, so their expected loss from a given catastrophe may be higher (Parry, 2007).

As to the related literature on endowment heterogeneity, results from threshold public-good experiments that include at least one (cooperative) Nash equilibrium are inconclusive as to whether or not such heterogeneity hinders success in reaching the threshold. On the one hand, Rapoport and Suleiman (1993) and Bernard et al. (2011) find that endowment heterogeneity is negatively related to contributions to the public good. On the other hand, Croson and Marks (2001) (see also the meta-analysis by Croson and Marks, 2000) find that heterogeneity has no significant impact on contributions. In collective-risk social dilemma experiments related to ours, endowment heterogeneity also yields ambiguous findings. Tavoni et al. (2011) find that in a symmetric treatment 50% of the groups reach the target of €120, while only 20% of the groups achieve it under heterogeneity. Milinski et al. (2011) find that subjects contribute proportionally to the heterogeneous endowments. Burton-Chellew et al. (2013) observe that under endowment homogeneity 87% of the groups reach the target sum, but only 62% under heterogeneity. By contrast, Brown and Kroll (2017) find that endowment heterogeneity does not reduce contributions when the threshold is either certain or uncertain, and even when the probability of loss depends on the contribution level.[3] As to the effect of heterogeneity in loss rates, it has not

[2]In the field, such heterogeneities could be broken into more subtle definitions, for instance, Lange et al. (2010) distinguish between equal per-capita emissions, equal percentage reduction of current emissions, equal ratio between abatement costs and emissions, and equal ratio between abatement costs and GDP. Brick and Visser (2015) distinguish between equal per-capita entitlement to emissions, reduction of current emissions, historical polluter-pays, and future polluter-pays.

[3]Under a known threshold and fixed probability of loss, the success rate is 60% under symmetry and 70% under endowment heterogeneity. Under a known threshold and variable probability of loss, the success rate is 25% under symmetry and 55% under endowment heterogeneity (Brown and Kroll, 2017, Fig. 1, p.162).

yet been tested in this context. Finally, regarding the effect of more than one heterogeneity, Burton-Chellew et al. (2013) conducted two treatments with double heterogeneity in *endowment* and *risk probability*. They find that the success rate is considerably lower in the double heterogeneity treatment in which rich subjects face a lower catastrophic risk than their poor counterparts than under symmetry, endowment heterogeneity, or when the rich members face higher risk than their poor counterparts.

To this end, we conducted a collective risk social dilemma experiment with homogenous (symmetric) conditions, endowment heterogeneity, loss heterogeneity, and also a double heterogeneity in endowment and loss. Our study has four main contributions to make: The first design novelty is studying the pure effect of loss heterogeneity. Even though heterogeneity in expected losses is a well-recognized feature of the climate change dilemma (e.g., Mendelsohn et al., 2006; IPCC, 2013) and despite the large body of literature modeling climate change using a public good game and its variants (e.g., the collective-risk social dilemma), the pure effect of heterogeneity in loss rates or risk (in comparison with homogeneous parties) has, to our knowledge, not been tested before. A second novelty is comparing behavior under theoretically equivalent endowment and loss heterogeneities. In particular, our study is designed so that, for rational participants, endowment and loss heterogeneities result in congruent equilibria. Yet, a heterogeneity in loss rate only matters when a catastrophe occurs, while a heterogeneity in endowment is immediately present. So might it be that these two heterogeneities lead to different behavior despite their theoretical equivalence? As a third contribution, we allude to the question whether heterogeneity facilitates rather than impedes cooperation in the collective-risk social dilemma. This is especially interesting in light of the contradicting results of Tavoni et al. (2011) and Burton-Chellew et al. (2013) versus Brown and Kroll (2017) regarding the effect of endowment heterogeneity on cooperation. Finally, our paper disentangles the effect of joint heterogeneity in endowment and loss rate. We investigate the most realistic type of heterogeneity, i.e., in both wealth and expected loss (e.g., IPCC, 2013), and compare it to each of the single heterogeneities.

We find that the success rate with homogenous parties is no higher than under heterogeneities in endowment or loss rate. In fact, under heterogeneity in loss rate, the success rate in preventing a catastrophic event is significantly higher than with homogenous parties. Even under double heterogeneity in which poor members face a higher loss rate than their rich counterparts, the success rate in preventing a catastrophe is not significantly different from the other treatments. As to overall burden sharing in successful groups, under loss heterogeneity most groups share their burden in accordance with equal contributions (leading to equal earnings). By contrast, under either endowment heterogeneity or double heterogeneity in endowment and loss, most groups settle on proportional contributions to the endowment ratio. While under loss heterogeneity the tendency towards the equal contributions and equal earnings resolution is already revealed in the first round, we observe no strong tendency for any of the focal resolutions in the other two treatments with endowment heterogeneity. Finally, in all treatments, efficiency is considerably higher in groups successfully reaching the threshold than in unsuccessful groups. On the bottom line, in the light of the popular collective-risk social dilemma, our results are the first to indicate that heterogeneity may actually facilitate public-good provision in such a situation.

We proceed as follows: The next section describes the experimental design. Section 3 provides a theoretical analysis of the game and characterizes the equilibria. Section 4 describes the experimental procedure and formulates our *a priori* hypotheses. In Section 5 we present the experimental results and discuss them in Section 6. Finally, Section 7 contains our conclusions.

2 The Experimental Design

Our experimental design builds on Milinski et al. (2008). A group of six subjects can prevent a major loss of private wealth if they jointly succeed in contributing some of their initial endowment to a so-called *prevention account*. More precisely, each group member i is endowed with some amount of initial money a_i. The game lasts exactly 10 rounds. In each round, each subject has to split up 10% of his/her initial endowment between a private account and a prevention account. The total endowment for the six group members is always €240, and to avoid a catastrophe with certainty, €120 have to have accumulated in the prevention account at the end of the 10 rounds.[4] Otherwise a catastrophe will occur with probability p, which can then cause a partial loss $0 < q_i < 1$ of the

[4]Even if the threshold is scientifically uncertain, we chose a *certain* threshold in the experiment for two reasons. First, a particular threshold can be set politically (e.g., it has been suggested that 2°C is no less a political anchor than scientific evidence (Randalls, 2010)). Secondly, under threshold uncertainty it is impossible to disentangle the heterogeneous parties' intent to share the burden (as burden-sharing depends on the expected target sum) from a general tendency to undercompliance.

money in the private account.[5]

We conducted four treatments: symmetric (SY), asymmetric endowment (AE), asymmetric loss (AL), and a double heterogeneity treatment combining asymmetric endowment and loss (AEL).

Endowment homogeneity was implemented so that each subject starts with a_i=€40. Under loss homogeneity, in the case of the group failing to collect the targeted sum, each group member loses a share of $q_i = 0.75$ of his/her private account with a risk probability $p = 0.667$.[6] Under symmetry (SY treatment) both endowment and loss ratios are homogenous. Under endowment heterogeneity (AE treatment), three group members receive a high (low) total endowment of €48 (€32). Under loss heterogeneity (AL treatment), three members face a high (low) loss rate of 0.9 (0.6). Finally, under double heterogeneity (AEL treatment), three group members receive an endowment of €48 and face a low loss rate of 0.6 (€32 and a high loss rate of 0.9). The heterogeneity ratio measuring the difference in expected payoffs between the member types in the case of no contribution to the prevention account is 1.5 in each of the endowment heterogeneity and the loss heterogeneity treatments but is 2.25 in the double heterogeneity treatment. The game structure is illustrated in Figure 1, and the design parameters are displayed in Table 1.

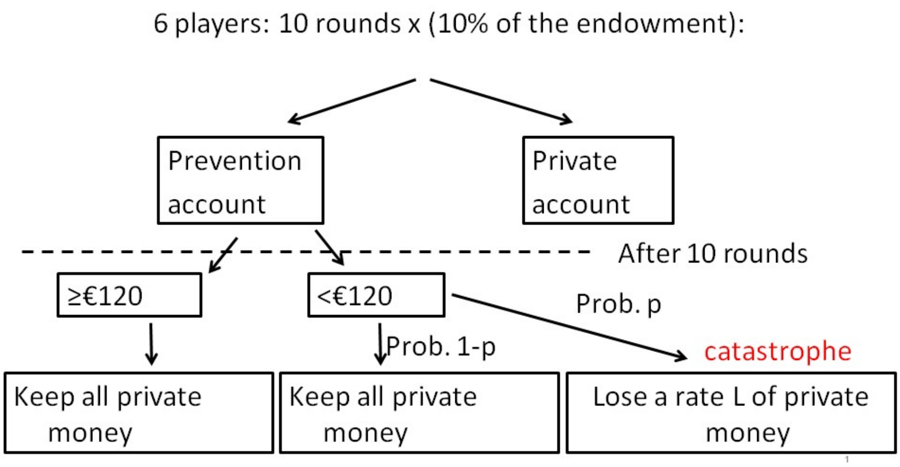


Figure 1: The game structure

[5]Note that to model heterogeneity in loss, it is necessary to choose a loss rate lower than 100%.

[6]Note that this corresponds to the same expected payoff as in the previous studies by Milinski et al. (2008) and Tavoni et al. (2011) for catastrophe probability $p = 0.5$ and $q_i = 1$.

Table 1: The design parameters

Treatment	Endowment (a_i)	Risk prob. (p)	Loss rate (q_i)
Symmetric (SY)	€40	0.66	0.75
Asymmetric Endowment (AE)	€48; €32	0.66	0.75
Asymmetric Loss (AL)	€40	0.66	0.60; 0.90
Double Asymmetry (AEL)	€48; €32	0.66	0.60; 0.90

3 Theoretical Representation and Formal Prediction

3.1 Environment with homogenous players

In the following we describe the theoretical background of the game our experiment is based on and derive some predictions concerning the outcomes. Note that, we do not define strategies in the subgame after each round, but rather characterize the open loop equilibria. We assume an even number n of players or subjects (in our experimental design $n = 6$), all of them receiving an initial money endowment of size a, identical for all subjects. The game consists of T ($= 10$) rounds. To prevent a major loss referred to as a catastrophe, half of the total endowment on average will have to be sacrificed by all players. Let c_{it} be the contribution of player $i = 1, ..., n$ in round t, let $c_i = (c_{i1}, ..., c_{iT})$ be the contribution sequence of player i, and let $c = (c_1, ..., c_n)$ be the *contribution profile* of all players. Moreover we define $\tilde{c}_i = \sum_{t=1}^{T} c_{it}$ as subject i's total contribution over all T rounds. We use p to denote the probability of a loss occurring if total contributions fall short of a particular threshold $A \equiv \frac{n}{2}a$, i.e., if $\sum_{i=1}^{n} \tilde{c}_i < A$, and q to denote the loss rate of the subjects' final wealth if a catastrophe occurs, i.e., a share $(1 - q)$ will be left. Thus at the end of the game, player $i's$ expected earnings will be $p(1 - q)(a- \tilde{c}_i) + (1 - p)(a- \tilde{c}_i) = (1 - pq)(a- \tilde{c}_i)$ if total contributions fall short of threshold A.

Letting y_i be player $i's$ final monetary payoff, his or her utility is denoted by $U(y_i)$. By contrast, we write $\tilde{U}_i(c)$ as subject i's utility resulting from the strategy profile c. Player i's expected utility is then given by

$$\tilde{U}_i(c) = \begin{cases} U\left(a - \tilde{c}_i\right) & \text{if} \quad \sum_{i=1}^{n} \tilde{c}_i \geq A \\ (1-p)U\left(a - \tilde{c}_i\right) + pU\left((1-q)(a - \tilde{c}_i)\right) & \text{if} \quad \sum_{i=1}^{n} \tilde{c}_i < A \end{cases}$$

Equilibria. This game has several equilibria. No contribution by any player is always an equilibrium since $a < \frac{n}{2}a = A$ for $n > 2$, i.e., no player can unilaterally afford to bear the entire burden of preventing the loss. In this case, a player's expected utility is given by

$$U_i(0, ..., 0) = (1 - p)U(a) + pU((1 - q)a)$$

We refer to this equilibrium as a *non-contribution equilibrium*.

However, there can be a continuum of other equilibria. A contribution profile c^* is called a *catastrophe prevention equilibrium* if for all $i = 1, ..., n$ the following holds:

$$0 < c^*_{it} \leq \frac{a}{T} \tag{1}$$

$$\sum_{i=1}^{n} \tilde{c}^*_i = \frac{n}{2}a \equiv A \tag{2}$$

$$U(a - \tilde{c}^*_i) \geq (1 - p)U(a) + pU((1 - q)a) \tag{3}$$

i.e., the certain utility from contributing $(\tilde{c}^*_1, ..., \tilde{c}^*_n)$ within T rounds is as least as large as the expected utility derived from not contributing at all.

In the following subsections we define the games and their corresponding equilibria for heterogeneous environments, i.e., for endowment heterogeneity, loss rate heterogeneity, and both heterogeneities combined.

3.2 Endowment heterogeneity

Let there now be two types of subjects with low and high initial endowments, a_L and a_H, respectively. Let $j(i)$ denote player i's type, and let $U_{i,j(i)}(\tilde{c})$ denote the utility of player i being of type $j(i)$ under contribution profile $\tilde{c}$. We assume now that, in order to prevent a catastrophe, on average half of the endowment will have to be sacrificed by all players. The utility of a type $j = L, H$ player is then given by

$$\tilde{U}_{i,j(i)}(\tilde{c}) = \begin{cases} U\left(a_{j(i)} - \tilde{c}_i\right) & \text{if } \sum_{i=1}^{n} \tilde{c}_i \geq A \equiv \frac{n/2}{2}(a_L + a_H) \\ (1 - p)U(a_j - \tilde{c}_i) + pU((1 - q)(a_j - \tilde{c}_i)) & \text{if } \sum_{i=1}^{n} \tilde{c}_i < A \end{cases}$$

Equilibria. Again, *contributing nothing* by any player is an equilibrium with utility given by

$$\tilde{U}_{i,j(i)}(0,...,0) = (1-p)U\left(a_{j(i)}\right) + pU\left((1-q)a_{j(i)}\right)$$

A profile c^* is a *catastrophe prevention equilibrium* if for all $i = 1, ..., n$:

$$0 < c^*_{it} \leq \frac{a_{j(i)}}{T} \tag{4}$$

$$\sum_{i=1}^{n} \tilde{c}^*_i = \frac{n}{4}(a_L + a_H) \equiv A \tag{5}$$

$$U\left(a_{j(i)} - \tilde{c}^*_i\right) \geq (1-p)U\left(a_{j(i)}\right) + pU\left((1-q)a_{j(i)}\right) \tag{6}$$

3.3 Loss heterogeneity

Under *loss heterogeneity* we assume that initial endowments are equal, but the loss rates denoted by q_L, q_H differ across players. Thus if the player is of type $j = L, H,$, player i's expected utility is given by

$$\tilde{U}_{i,j(i)}(c) = \begin{cases} U\left(a - \tilde{c}_i\right) & \text{if} \quad \sum_{i=1}^{n} \tilde{c}_i \geq A \equiv \frac{n}{2}a \\ (1-p)U\left(a - \tilde{c}_i\right) + pU\left((1 - q_{j(i)})(a - \tilde{c}_i)\right) & \text{if} \quad \sum_{i=1}^{n} \tilde{c}_i < A \end{cases}$$

Equilibria: As before, no contribution by any player is an equilibrium with utility

$$U_{i,j(i)}(0,...,0) = (1-p)U\left(a\right) + pU\left((1-q_{j(i)})a\right)$$

A profile c^* is a *catastrophe prevention equilibrium* if for all $i = 1, ..., n$:

$$0 < c^*_{it} \leq \frac{a}{T}$$

$$\sum_{i=1}^{n} \tilde{c}^*_i = \frac{n}{2}a \equiv A$$

$$U\left(a - \tilde{c}^*_i\right) \geq (1-p)U\left(a\right) + pU\left((1-q_{j(i)})a\right)$$

3.4 Conditions for comparability

To make endowment and loss heterogeneities comparable for experimental subjects, we require that the ratio of loss rates for high and low loss players equals the ratio of high versus low endowments in the endowment heterogeneity environment, i.e.:

$$\frac{q_H}{q_L} = \frac{a_H}{a_L} \tag{7}$$

We also assume that the expected non-contribution payoffs are the same for (i) poor (i.e. low-endowment) players and high-loss players and (ii) for rich (i.e. high-endowment) and low-loss players, respectively. Noting that the expected non-contribution payoff is $(1-p+p\cdot(1-q))a_j = (1-pq)a_j$ for $j = L, H$ (similar for loss heterogeneity), the above assumption can be expressed as

$$(1 - pq)a_L = (1 - pq_H)a \tag{8}$$

$$(1 - pq)a_H = (1 - pq_L)a \tag{9}$$

with $a = (a_L + a_H)/2$.

Furthermore we need to ensure that even for risk-neutral players a catastrophe-prevention equilibrium exists. Under *endowment heterogeneity* this is the case if total contributions $\hat{c}_L$ and $\hat{c}_H$ for poor and rich players, respectively, exist such that

$$(1 - pq)a_j = a_j - \hat{c}_j, \qquad j = L, H \tag{10}$$

and

$$\hat{c}_L + \hat{c}_H \geq \frac{a_L + a_H}{2} \tag{11}$$

Note that $\hat{c}_j$ can be considered the *maximum willingness to contribute* for players of type $j = L, H$.

Under *loss heterogeneity*, a catastrophe-prevention equilibrium exists if there are total contributions $\check{c}_L$ and $\check{c}_H$ for poor and rich players, respectively, such that

$$(1 - pq_j)a = a - \check{c}_j, \qquad j = L, H \tag{12}$$

and contributions are sufficiently high to prevent the catastrophe, i.e.,

$$\check{c}_L + \check{c}_H \geq a \tag{13}$$

From these conditions we can derive the following result, which limits the choice of parameters:

Proposition 1: *Under the conditions (7) through (13), the loss rates q, q_L, and q_H have to be chosen such that* <u>*risk-neutral*</u> *subjects are indifferent between the non-contribution and the catastrophe-prevention equilibrium, implying that*

p and q have to be chosen such that

$$q = \frac{1}{2p}$$

and the payoff rates in the non-contribution equilibrium are equal for AE and AL, implying

$$\frac{1 - pq_H}{1 - pq_L} = \frac{a_L}{a_H} \tag{14}$$

The unique equilibrium contributions are then given by $\hat{c}_L = pqa_L$ and $\hat{c}_H = pqa_H$ for endowment heterogeneity, and by $\check{c}_L = pq_La$ and $\check{c}_H = pq_Ha$ for loss heterogeneity.

The proof is to be found in Appendix A.1.

3.5 Double heterogeneity

Finally we turn to *double heterogeneity*, where low-endowment (i.e. poor) subjects face high loss rates and the opposite holds for high-endowment (i.e. rich) subjects. We now use q_L and q_H to denote the loss rates for low- and high- endowment subjects. Contrary to the loss heterogeneity model, this implies that $q_L > q_H$. As before, we use $\tilde{q}_j = 1 - pq_j$ to denote the expected survival rate per dollar owned. Obviously, $q_L > q_H$ implies $\tilde{q}_L < \tilde{q}_H$. Moreover, we assume that

$$\tilde{q}_La_L < \tilde{q}_Ha_H/2 \tag{15}$$

This assumption implies that if one subject unilaterally deviates from any catastrophe-prevention equilibrium, the expected deviation payoff for the low-endowment subjects is lower than the expected compliance payoff for the high-endowment subjects. Observe that, conversely, our assumptions imply

$$\tilde{q}_Ha_H < \tilde{q}_La_L/2 \tag{16}$$

The following result shows that under double heterogeneity the only equilibrium is contributing nothing. The proof is depicted in Appendix A.2

Proposition 2: *In the double heterogeneity game under the above assumptions, the non-contribution profile is the only equilibrium.*

3.6 Resolution concepts

In a public good game with heterogeneous players there are numerous possible burden sharing resolutions. The literature offers two relevant allocation principles (see Konow, 2003): "equality of outcomes" and "proportionality between inputs and outcomes." In our game, two burden-sharing resolutions are in line with these fairness principles and hence can be considered prominent resolutions: equality of outcomes is implemented by equal earnings, while proportionality between inputs and outcomes is articulated by contributions proportional to endowments or loss rates. A third prominent resolution since convenient to implement is equality of inputs, carried out by equality of contributions. Hence, and in line with Reuben and Riedl (2013), we consider as benchmarks those three prominent resolutions: *equal contributions*, *proportional contributions*, and *equal earnings*.[7] Under *equal contributions*, each subject contributes the same amount independently of the endowment or loss rate. Under *proportional contributions*, subjects contribute proportionally to the ratio of initial endowments or the loss rates (or both in the double heterogeneity treatment), while under *equal earnings*, subjects contribute in such a way that all group members achieve the same final payoff independently of their initial position. Table 2 shows the payoffs of risk-neutral subjects under no contributions and the different resolution concepts.

Table 2 shows that for risk-neutral players (i) in the symmetric environment (i.e., homogenous players), the unique contribution equilibrium obviously satisfies all three resolution concepts, and (ii) under both endowment and loss heterogeneities the (unique) contribution equilibrium requires proportional contributions, while the other two burden-sharing concepts cannot occur in equilibrium. Finally (iii), under double heterogeneity, none of the resolution concepts constitutes an equilibrium. We extend our analysis briefly to the case of risk aversion.

3.7 Risk aversion

Proposition 1 shows that risk-neutral players are indifferent between the two equilibria and that the contribution equilibria are unique in the endowment and loss heterogeneity treatments. If subjects are risk-averse, which is usually the

[7] Note that such burden-sharing rules are acknowledged in the literature on climate change negotiations. For instance, Kesternich et al. (2014, 2018) and Gallier et al. (2017) test the performance of minimum-contribution mechanisms, *explicitly* specifying these rules.

Table 2: Expected payoffs (and variances) of each member type under the prominent contribution profiles

	No cont.	equal conts.	proportional conts.	equal earnings
		Symmetry (SY)		
All subjects	20	20	20	20
(Var.)	(200)	-	-	-
		Endowment heterogeneity (AE)		
High endowment	24	28	24	20
(Var.)	(288)	-	-	-
Low endowment	16	12	16	20
(Var.)	(128)	-	-	-
		Loss heterogeneity(AL)		
High loss	16	20	16	20
(Var.)	(288)	-	-	-
Low loss	24	20	24	20
(Var.)	(128)	-	-	-
		Double heterogeneity (AEL)		
High endowment & low loss	28.79	28	20.30	20
(Var.)	(202.78)	-	-	-
Low endowment & high loss	12.79	12	19.69	20
(Var.)	(119.67)	-	-	-

Variances of expected payoffs are positive only when the groups do not reach the threshold. Moreover, under double heterogeneity, "proportional conts." refers to contributions proportional to the *two* heterogeneity dimensions. In the case of contributions proportional to *one* heterogeneity dimension, the expected payoffs are 24 and 16 to the high-endowment & low-loss members and the low-endowment & high-loss members, respectively.

case, they will strictly prefer the contribution equilibria characterized above to the non-contribution equilibrium. This in turn implies that other equilibria exist. This is particularly interesting under environments with heterogeneous players because several of the focal resolution outcomes are candidates for equilibrium outcomes.

Corollary 1: *For the risk-averse subject, the sure payoff in the catastrophe-prevention equilibria* $(\hat{c}_L, ..., \hat{c}_L, \hat{c}_H, ..., \hat{c}_H)$ *and* $(\breve{c}_L, ..., \breve{c}_L, \breve{c}_H, ..., \breve{c}_H)$ *for AL and AE, respectively, is strictly preferred to the risky non-contribution equilibrium.*

The (simple) proof is given in Appendix A.3. Note that risk-averse subjects accept a lower certainty equivalent. Therefore, the set of equilibrium outcomes increases when the subjects' risk aversion rises. Accordingly, in the *symmetric* case there is a continuum of equilibrium outcomes satisfying equations (1) - (3).

3.7.1 Endowment heterogeneity

In the same vein as above, for risk-averse subjects there is a continuum of catastrophe-prevention equilibrium outcomes in the endowment heterogeneity environment satisfying equations (4) - (6). Note that a continuous move from the proportional to the equal contributions profile benefits high-endowment players but makes low-endowment players worse off. Since a greater degree of risk aversion lowers the certainty equivalent, it may be the case that for the low-endowment players, the equal contributions yields a utility as least as high as the certainty equivalent. To see how the contribution profile is related to the degree of risk aversion, we discuss two numerical examples.

Example 1: We take the values of our experimental set-up outlined below: $a_L = 32$, $a_H = 48$, $p = 2/3$, $q = 3/4$, and a constant-relative-risk-aversion von Neumann-Morgenstern utility function $U(x) = x^\gamma$ with $0 < \gamma < 1$. Then for extremely high risk aversion only, i.e., risk aversion coefficients close to zero ($\gamma < 10^{-15}$), the equal-contributions profile $(20, ..., 20)$ is an equilibrium outcome. On the other hand, it is easy to compute that, with the same risk coefficient, the high endowment type's maximum willingness to contribute is $\hat{c}_H = 29.28 > 28$. Therefore, for $\gamma < 10^{-15}$ there is a continuum of equilibrium outcomes including all three focal-point resolution concepts. By contrast, for more realistic (relative) risk-aversion parameters such as $\gamma = 0.5$, the set of equilibrium outcomes is (with slight abuse of notation) $\{(c_L, c_H)|\ 13.3 \leq c_L \leq 17.8,$ $22.2 \leq c_H \leq 26.7,\ \sum_{j=1}^{6} c_j = 120\}$, i.e., the set of equilibrium outcomes includes proportional contributions, but not equal contributions or equal earnings.

Example 2: This is similar to example 1, but a lower degree of endowment inequality is chosen by setting $a_L = 36$, $a_H = 44$. Even for $\gamma \leq 0.5$, the equal contributions profile $(20, ..., 20)$ is now an equilibrium outcome. By contrast, the high-endowment type's maximum willingness to contribute is $\hat{c}_H = 36.0$.

Accordingly, there is a continuum of outcomes $\{(c_L, c_H)|\ 15.6 \leq c_L \leq 20,$ $20 \leq c_H \leq 24.4,\ \sum_{j=1}^{6} c_j = 120\}$ where both the proportional and the equal-contributions profile are within the set of equilibrium outcomes, while the equal-earnings profile is not (since $\hat{c}_H \leq 24.4 < 28$).

Hence, the set of potential equilibrium outcomes may shrink or increase in response to game parameters.

3.7.2 Loss and double heterogeneity

Under *loss heterogeneity* with risk-averse players, there is also a continuum of equilibrium outcomes. The interesting question with respect to our experimental analysis is under what circumstances the equal-contributions (= equal-earnings) profile can be an equilibrium. Similar to endowment heterogeneity, for our experimental parameters $q_L = 0.6$ and $q_H = 0.9$ and utility function $U(x) = x^\gamma$, there are only equilibria for extremely low values of γ, i.e., here $\gamma < 6 \cdot 10^{-16}$. By contrast, for reasonable risk-aversion parameters such as $\gamma = 0.5$, loss rates must not be too different between the types (in our case $q_L \geq 0.685$, $q_H \leq 0.815$) to obtain the equal contributions (=equal earnings) equilibrium.

With *double heterogeneity* we have seen that under risk-neutrality no catastrophe-prevention equilibrium exists. Things are not much different under risk aversion. For our experimental parameters, catastrophe-prevention equilibria only exist under extreme risk aversion.

3.7.3 Summary

In sum, our analysis indicates that for reasonably risk averse individuals (i) under symmetry, the cooperative equilibrium is strictly preferred to the non-cooperative equilibrium. (ii) under either endowment or loss heterogeneity, the only resolution-contribution profile that belongs to the equilibrium set is proportional contributions to the endowment or loss heterogeneity, respectively. This equilibrium is also strictly preferred to the non-contribution equilibrium. Finally (iii), under double heterogeneity in endowment and loss rate, none of the focal resolution contribution profiles belongs to the set of equilibrium outcomes. Thus, the unique equilibrium is the non-cooperative equilibrium.

4 Procedure and Hypotheses

In the following we describe the experimental procedure and formulate our research hypotheses following the theoretical analysis.

4.1 Experimental procedure

A total of 510 student subjects from different departments of the universities of Kiel and Rostock, Germany, participated in the experiment. This produced 21 independent groups each in the SY, AL, and AE treatments, and 22 groups in the AEL treatment. The experiment was programmed and conducted using the z-Tree experimental program (Fischbacher, 2007). The program has three stages: In the first stage, instructions are displayed on the computer screen, including control questions to verify that subjects have understood the instructions. In the second stage, subjects make decisions in the 10-round game (where, after each contribution round, subjects are informed about both the individual and the accumulated contributions of their group members). If, in the final stage, the amount that has accumulated in the prevention account is less than €120, the program determines randomly whether or not a partial loss in the private account (mimicking a "catastrophe") will occur. In either case, final payoffs are displayed to the subjects.[8]

4.2 Research hypotheses

After introducing the design parameters, theoretical model, and resolution concepts, we are now ready to formulate our *a priori* hypotheses. The overarching hypothesis is that the more focal a burden-sharing resolution is, the smaller is the coordination problem and hence the higher is the success rate in preventing a catastrophe.

An apparent feature of the symmetric (SY) treatment is that it has one focal burden-sharing contribution profile, satisfying all three resolution concepts (i.e., *equal contributions* coincides with *proportional contributions* and yields *equal earnings*). This focal burden-sharing rule is supported by equilibrium play under risk-neutral and risk-averse preferences.[9]

By contrast, in the treatments with heterogeneity, group members looking to reach an equilibrium that eliminates the risk of a catastrophe face a coordination problem as to how to share the burden between the two types of group members. Accordingly, in the AE and AL treatments we chose design parameters ensuring that for risk-neutral subjects the unique cooperative equilibrium requires *contributions proportional to the endowments and loss rates*. Moreover, from the three resolution contribution profiles only *proportional contributions* is an equilibrium for (reasonably) risk-averse subjects.

In the double heterogeneity (AEL) treatment, by contrast, it is not obvious how to share the burden between the two types of group members. For a selfish individual with risk-neutral (or not too risk-averse) preferences, contributing nothing is the only equilibrium. Thus we can now summarize our *a priori* hypotheses regarding success rates and burden sharing in successful groups.

[8] All computer screens are provided as supplementary information.

[9] We refer here to catastrophic risk. Obviously games with imperfect information (due to simultaneous moves) also entail strategic uncertainty as to the other members' actions.

Hypothesis 1 [success rate]: *We expect a higher success rate in reaching the threshold under symmetry than under loss or endowment heterogeneity. Moreover, we expect similar success rates under loss and endowment heterogeneity. Finally, we expect the lowest success rate under double heterogeneity.*

Hypothesis 2 [Burden sharing in successful groups with heterogeneous member types]: *Under either loss or endowment heterogeneity we expect successful groups to share their burden around the proportional contributions resolution.*

5 Results

We start by reporting the results with respect to success rates and burden sharing achieved by successful groups in the different treatments. Then we investigate the first-round contributions to learn about decisions that are only based on dispositions, beliefs, and priors (without yet being influenced by others' choices and experience). Next, we study the contribution dynamics over time and in the last round. Finally, to learn about overall efficiency, we inspect the payoffs across treatments and between groups successful and unsuccessful in preventing a catastrophic event.

5.1 Success in reaching the threshold

We start with our main result regarding success rates in reaching the threshold and overcoming the catastrophe. Figure 2 shows success rates across treatments. Surprisingly, we find that the SY treatment yields the lowest success rate in reaching the prevention target (38%). The AE and AEL treatments display 57% and 59% success rates, respectively. The highest success rate is observed in the AL treatment, where 67% of the groups reach the target. Formally, using a Chi-square test we find that the success rate is higher in the AL than in the SY treatment ($p = 0.06$), but we observe no significant differences among the other treatments.

Result 1 [success rate]: *Success rate is highest under loss heterogeneity and lowest under symmetry. There is no difference in success rates between the heterogeneous treatments. Interestingly, the success rate under double heterogeneity is no lower than in any other treatment.*

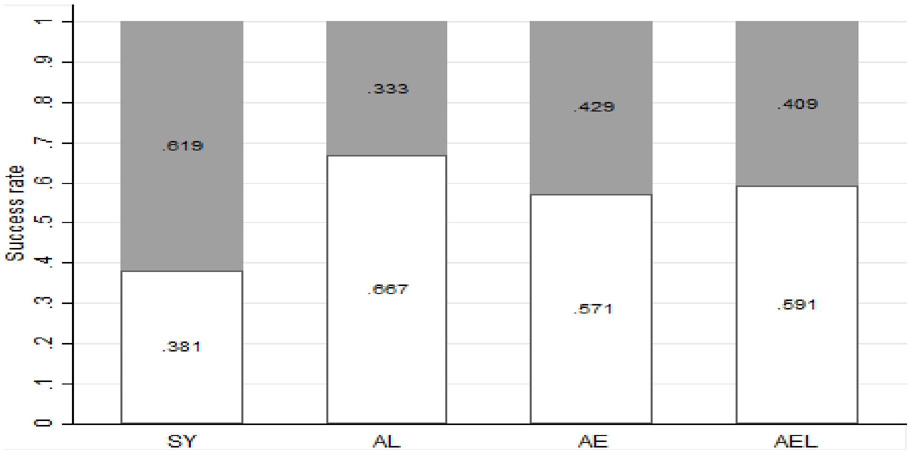
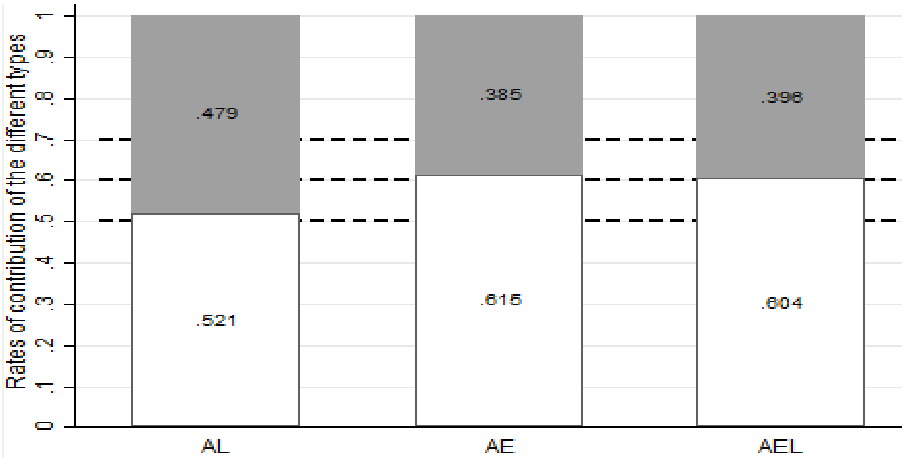


Figure 2: Success (light bars) and fail (dark bars) rates in the different treatments

5.2 Burden-sharing

Next we inspect burden sharing by each successful group to learn whether there is convergence in the different treatments to one of the burden-sharing contribution profiles. In the AL treatment, the *equal contributions=equal earnings* contribution profile requires that the contribution share of high-loss members be equal to that of the low-loss members (50% for each), while *proportional contributions* requires that the high-loss members contribute 60% of the total amount. In the AE and AEL treatments, *equal contributions* requires that the contribution share of high-endowment members be equal to that of their low-endowment counterparts (i.e., 50% each), while *equal earnings* requires that the contribution share of high-endowment members be equal to 70% of the total amount. *Proportional contributions to the endowment ratio* requires the high-endowment members to contribute 60% of the burden. In the AEL treatment, *proportional contributions to both endowment and loss rates* requires the high-endowment & low-loss members to contribute about 71% of the total burden share.

Figure 3 shows the relative share of total contributions to the prevention account in successful groups. The average contribution share of high- and low-loss subjects in AL groups is [52%:48%], corresponding quite closely to the *equal-contributions* (and *equal-earnings*) burden-sharing profile. By contrast, the average contribution share of high-endowment and low-endowment subjects in the AE treatment is [61%:39%], corresponding closely to the *proportional-contributions* burden-sharing rule. Finally, the contribution ratio between high-endowment & low-loss subjects and low-endowment & high-loss subjects in the



The light bars indicate the contribution rates of "high types" (i.e., high-loss, high-endowment, and high-endowment & low-loss members in the AL, AE, and AEL treatments, respectively). The horizontal dashed lines denote the prominent contribution-sharing rules. The top line indicates *equal earnings* in the AE and AEL treatments. The middle line represents *proportional contributions* to the loss or endowment ratio in the AL, AE, and AEL, respectively. The bottom line denotes *equal contributions* in all treatments (as well as *equal earnings* in the AL treatment).

Figure 3: Relative share of total contributions to the prevention account by the different types of successful groups in the heterogeneous treatments

AEL treatment is [60%:40%], corresponding to contributions *proportional to the endowment ratio*.[10]

However, group members do not share the burden in exact accordance with the three above-mentioned contribution profile resolutions. Moreover, the average burden sharing across groups may not provide an accurate description of actual burden sharing. Consider, for instance, the extreme case where half of the groups share the burden around *equal earnings*, while the other half around *equal contributions*. In this case we receive that, on average, the groups share the contribution burden around the *proportional contributions* resolution (although none of the groups share the burden close to this contribution rule). Therefore, we classify strategies in each group as *equal contributions*, *proportional contributions*, or *equal earnings* depending on their distance from the respective benchmark strategy. Table B.1 in the Appendix shows the burden share of each successful group and its corresponding classification according to the focal resolutions.

Starting with the AL treatment, we classify a contribution profile as *equal*

[10]Formally, using a Median test we find that the contribution ratio in the AL treatment is no different from the *equal-contributions* (= equal-earnings) burden-sharing rule ($p = 0.43$), but is different from the *proportional-contributions* burden-sharing rule ($p < 0.01$). In addition, we find that the contribution ratios in the AE and AEL treatments are no different from the *proportional (to endowment)-contributions* burden-sharing rule ($p = 0.18$ and $p = 0.50$, respectively). The comparisons with other burden-sharing rules yield $p < 0.01$.

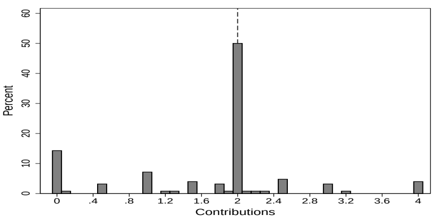
contributions and *equal earnings* if the high-loss members contribute 45%-55% of the burden share. Similarity, we classify a contribution profile as *proportional contributions* if high-loss members contribute 55%-65% of the burden share. From the total of 14 groups successful in reaching the target, 57% correspond to the *equal-contributions=equal-earnings* contribution profile, whereas 35% correspond to the *proportional-contributions* resolution profile.

Turning to the AE and AEL treatments, we classify a contribution profile as *equal contributions* (*equal earnings*) if the high-endowment members contribute 45%-55% (65%-75%) of the burden share. *Contributions proportional to one of the heterogeneity dimensions* requires the high-endowment members to contribute a total share of 55%-65%. In the AE treatment, 75% of the groups are classified as *proportional contributions* (while only 8% and 17% are classified as *equal contributions* or *equal earnings*). In the AEL treatment, 84% of the groups are classified as *contributions proportional to one heterogeneity dimension*. Furthermore, one group (8%) is classified as *equal contributions* and another group (8%) is classified as either *equal earnings* or *contributions proportional to both heterogeneity dimensions* (as these two outcomes coincide). We can now formulate our next result.

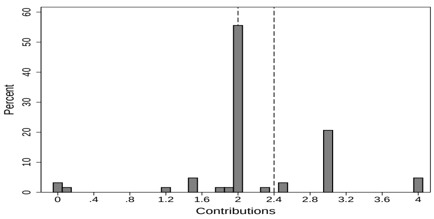
Result 2 [Burden sharing in successful groups with heterogeneous member types]: *Under loss heterogeneity, most groups share the burden according to the equal-contributions and equal-earnings resolution. But a sizable portion of the groups share according to contributions proportional to the loss ratios. By contrast, most successful groups under endowment heterogeneity coordinate on contributions proportional to the endowment ratio. The burden-sharing classification under double heterogeneity largely resembles that of endowment heterogeneity, with almost all groups sharing the burden according to the one-dimension (endowment) heterogeneity.*

5.3 First-round contributions

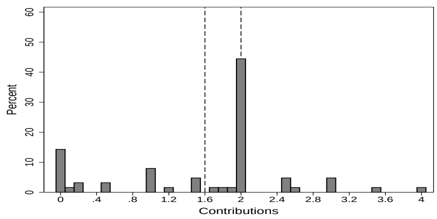
Recall that contributions in the collective risk social dilemma are made sequentially, so overall burden sharing only reveals average behavior over time. In this regard, subjects' first round contribution decisions are based only on their predispositions, initial beliefs, and priors without as yet being influenced by the decisions of their group members. Figure 4 shows first-round contributions in the different treatments. We inspect these contributions to learn (i) whether subjects actually consider the three focal burden-sharing profiles to be genuine candidates for coordination when making their contribution decisions, and (ii) whether the subjects' initial perception of the success chances in reaching the threshold differ across treatments (heterogeneities).



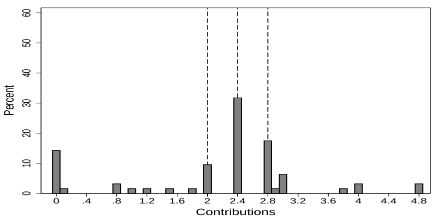
(a) SY treatment



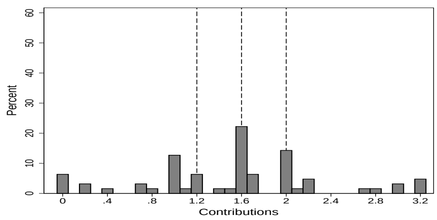
(b) AL treatment: high-loss members



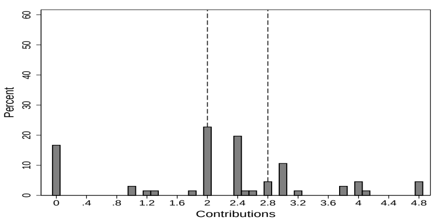
(c) AL treatment: low-loss members



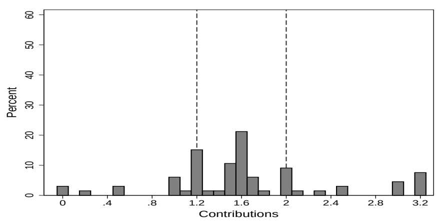
(d) AE treatment: high-endowment members



(e) AE treatment: low-endowment members



(f) AEL treatment: high-endowment & low-loss members



(g) AEL treatment: low-endowment & high-loss members

Figure 4: Contribution decisions in the first round of the different treatments (the vertical dashed lines indicate the focal contributions)

In the SY treatment, exactly 50% of the subjects contributed €2 to the prevention account (corresponding to all three burden-sharing rules). In the AL treatment too, exactly 50% of the subjects contributed €2 corresponding to the *equal-contributions* and *equal-earnings* resolutions. Moreover, no subject contributed an amount corresponding to the *proportional-contributions* share (€1.6 and €2.4 for low-loss and high-loss subjects, respectively). By contrast, in the AE treatment the *proportional-contributions* profile is the most frequently observed choice, selected by 26.98% of all subjects, while 11.90% of the subjects contributed amounts corresponding to the *equal-contributions* profile (€2) and another 11.90% contributed according to the *equal-earnings* resolution (€1.2 and €2.8 for low- and high-endowment subjects, respectively). Thus, the outcome is considerably less focal in this treatment compared to the resolution that emerges in the SY and AL treatments.

Finally, in the AEL treatment we observe that 15.90% of the subjects contributed €2 corresponding to the *equal-contributions* resolution, while 9.84% contributed amounts corresponding to the *equal-earnings* resolution (which is also approximately the same as the *proportional contributions profile taking into account the two dimensions of heterogeneity*). Moreover, we find that 17.42% of the subjects contributed *proportionally to the endowment ratio* (by choosing amounts of €1.6 and €2.4, respectively). Hence, in the first round under double heterogeneity no focal contribution profile emerges, which is in accordance with our prediction.

On the whole, we find that in the comparable SY, AL, and AE treatments 50%, 50%, and 51% of the subjects, respectively, contributed amounts corresponding to the prominent sharing rules.[11] Moreover, Figure 4 shows that the highest spikes in first-round contributions are in these resolutions. So it appears that the classification of burden-sharing rules is largely supported by the subjects' decisions.

The second question we could answer using first-round data is whether subjects' initial perception of success differ across treatments. Note that once a subject suspects that his/her group will not reach the target, he/she should stop contributing to the prevention account (this is observable in later rounds whenever it became clear that a group would not manage to reach the target). So the rate of zero contributions in the initial round provides a measure for the initial perception of the heterogeneities. In this respect, we find that 14%, 8%, 10%, and 7% of the subjects contribute zero in the first round of the SY, AL, AE, and AEL treatments, respectively (using a Chi-square test we find no difference across treatments). Accordingly, the fact that in the initial round there is no significant difference in the rate of non-contributors across treatments suggests that subjects do not have different priors concerning success rates in the

[11] In fact, 31%, 33%, 35%, and 30% of the contributions over all rounds correspond exactly to one of the focal resolution outcomes in the SY, AL, AE, and AEL treatments, respectively.

different treatments. We summarize or findings as follows:

Result 3 [First-round contribution behavior]: (i) *First-round contribution behavior under symmetry is strikingly similar to that under loss heterogeneity with 50% of the members in both treatments choosing to contribute €2 to the prevention account. (ii) First-round contribution behavior under endowment heterogeneity is not different than under double heterogeneity, in both treatments one obvious focal contribution strategy does not emerge. (iii) Moreover, the rates of no contributions to the prevention account are not different across treatments, implying that subjects in the different treatments do not a priori predict different chances in reaching the threshold across treatments.*

5.4 First-round contributions as predictors of success

Figure 5 presents the average first-round group contributions, differentiating between groups eventually successful and unsuccessful in reaching the threshold. From this figure, it is obvious that contributions in the first round are significant predictors of success in reaching the target. This observation is formally established using logit model estimations.[12]

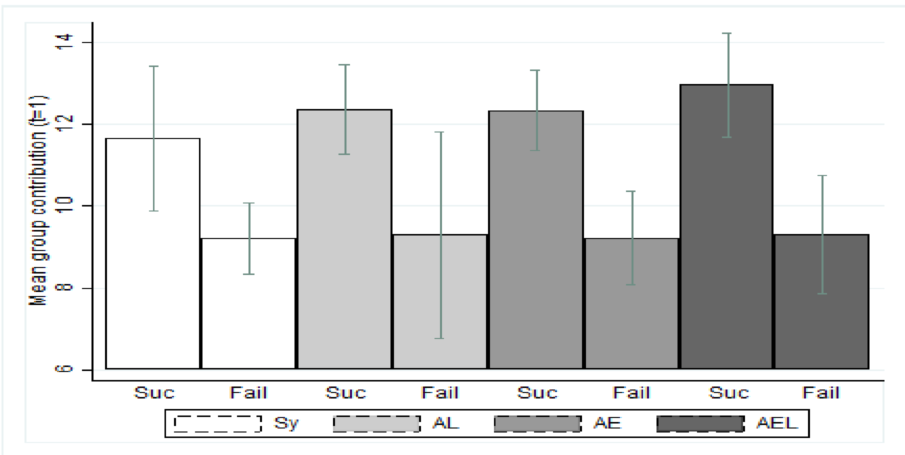
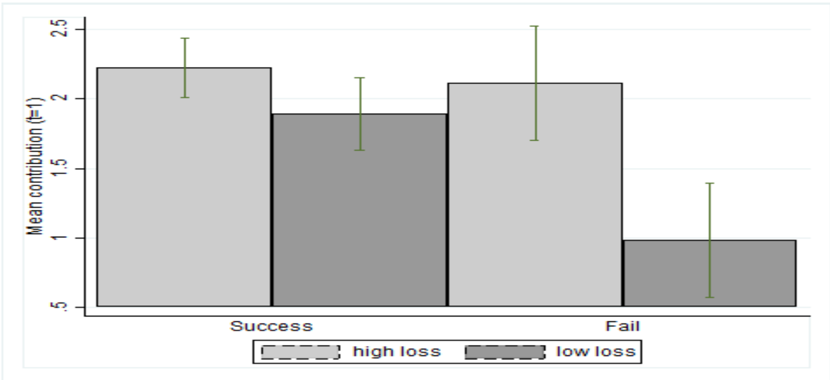


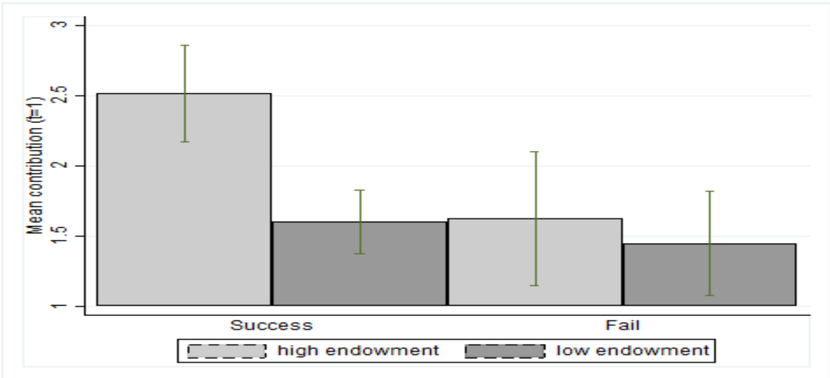
Figure 5: Average group contributions (and std. err.) in the initial round, per successful (Suc) and unsuccessful (Fail) groups

We are now interested in whether we can identify the member types in the heterogeneous treatments that are responsible for success or failure in reaching

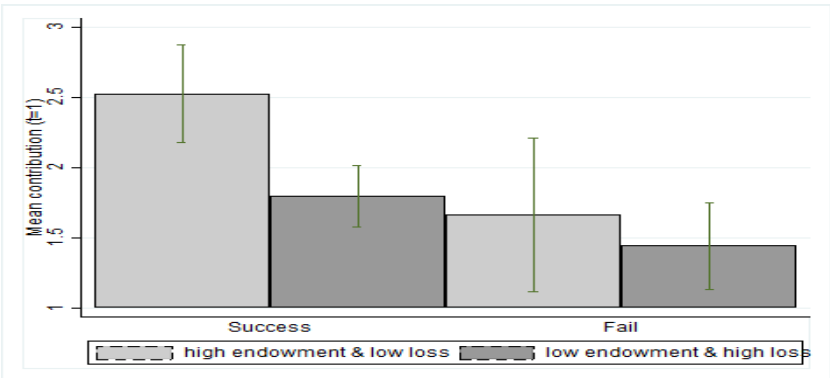
[12]The results are presented in Table B.2 in the Appendix. We were using groups as independent observations. Moreover, due to Ai and Norton (2003), who suspect significant interaction effects in non-linear probability models, we estimate the model for each treatment separately.



(a) AL treatment



(b) AE treatment



(c) AEL treatment

Figure 6: Average first-round individual contributions (and std. err.) of group-member types in successful and unsuccessful groups

the threshold. To this end, Figure 6 presents the average contributions of member types in the first round of eventually successful and unsuccessful groups in reaching the threshold. Figure 6 (a) shows that under loss heterogeneity (AL), high-loss members contribute similar amounts in successful and unsuccessful groups. But contributions from low-loss members are considerably higher in successful than in unsuccessful groups. This implies that the key for successfully reaching the threshold is for low-loss members not to contribute considerably lower amounts than their high-loss counterparts. Figure 6 (b) illustrates that under endowment heterogeneity (AE), low-endowment members contribute similar amounts in successful and unsuccessful groups. Hence, reaching the target is crucially predicted by the contributions of high-endowment members. Finally, Figure 6 (c) suggests that under double heterogeneity (AEL), reaching the target depends on the contributions of high-endowment & low-loss members (because in the initial round low-endowment & high-loss members do not contribute significantly different amounts in eventually successful or unsuccessful groups).[13] These observations indicate that success is largely predicted by the contribution decisions of "fortunate members" (those who are richer or stand to lose less from a catastrophe). We can summarize our next findings as follows:

Result 4 [First-round contributions as predictor of success]: *In all treatments, higher first-round contributions significantly predict success in reaching the threshold and preventing a catastrophe. Moreover, success depends on the willingness of those fortunate members either to compensate the poor (in the case of endowment or double heterogeneity) or to ignore the heterogeneity (in the case of loss heterogeneity).*

5.5 Contributions over time and in the last round

Next we turn our attention to individual contributions throughout the rounds, and in the final round (only for those groups that still have a chance of reaching the threshold). The variable $\text{GINI}_G(t = T)$ measures the inequality in commutative contributions between members of some group G up to round T.

$$\text{GINI}_G(t = T) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \| \sum_{t=1}^{T} c_{i,t} - \sum_{t=1}^{T} c_{j,t} \|}{2n\,(n-1) \sum_{t=1}^{T} \overline{c}_t} \quad ,$$

where $\sum_{t=1}^{T} c_{i,t}$ and $\sum_{t=1}^{T} c_{j,t}$ are the accumulated contributions of any two group members i and j in round T, and $\sum_{t=1}^{T} \overline{c}_t$ is the arithmetical sample mean of the accumulated group members' contributions in that round, while $n(= 6)$ is

[13] These results have been formally established using a two-sided Wilcoxon Mann Whitney U (MWU)-test. High-endowment members in the AE treatment and high-endowment & low-loss members in the AEL treatment contribute more in successful groups than in unsuccessful groups (both at $p = 0.003$). Also, low-loss members in the AL treatment contribute more in successful groups than in unsuccessful groups ($p = 0.0008$). We observe no other differences in the contributions of the different member types between unsuccessful and unsuccessful groups.

the group size.

In addition, in the heterogeneous treatments the variable L/H-ratio$_G(t = T)$ measures the difference in commutative contributions between the two member types in a given *heterogeneous* group G up to round T.

$$\text{L/H-ratio}_G(t = T) = \frac{\sum_{l=1}^{L} \sum_{t=1}^{T} c_{l,t}}{\sum_{h=1}^{H} \sum_{t=1}^{T} c_{h,t}} \quad ,$$

where $\sum_{t=1}^{T} c_{h,t}$ and $\sum_{t=1}^{T} c_{l,t}$ denote the accumulated contributions of "high" and "low" subjects, respectively, in round T.[14]

Table 3 shows random-effect estimations for the evolution of contributions in rounds 2-10. We observe that in all treatments the lower the inequality is in accumulated contributions (i.e., the lower the GINI), the higher the subsequent contributions (controlling for the amount currently stocked in the prevention account) must be. However, the most interesting result is the effect of the contribution ratio between the different member types (L/H-ratio). It appears that the lower this rate is in the AE and AEL (i.e., the more the high-endowment members contribute compared to their low-endowment counterparts), the higher the additional contribution of a group member will be. However, the opposite effect occurs in the AL treatment. The more low-loss members contribute in comparison with high-loss members (i.e., the contributions of low-loss and high-loss members is nearer to equal), the higher the subsequent individual contribution will be. This accords with Result 4. Hence, we can summarize as follows.

Result 5a [Contribution dynamics over time]: *Individual contributions depend on the amount accumulated in the prevention account. Moreover, the lower the inequality between group members in previous contributions, the higher the subsequent individual contributions will be. In addition, we find that under endowment heterogeneity and double heterogeneity, the more the high-endowment and high-endowment & loss-loss members contribute compared to their counterparts, the higher subsequent individual contributions will be. Under loss heterogeneity, the higher the contributions from the low-loss members, the higher subsequent individual contributions will be.*

Finally, although success rates in reaching the target are lowest under symmetry, we observe that 61%, 76%, 61%, and 63% of the groups in the SY, AL,

[14]High type is defined as a high-loss or high-endowment member in the AL and AE treatments and a high-endowment & low-loss member in the AEL treatment.

Table 3: Random-effect estimations of contribution decisions over time

	All Treatments	Heterogeneity	
Group accumulated $(t-1)$	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)
GINI $(t-1)$	-1.48***	-1.88***	-
	(0.43)	(0.68)	
L/H-ratio $(t-1)$	-	-	-0.12**
			(0.05)
AE	0.12	-	-
	(0.20)		
AL	-0.00	-0.13	-0.73***
	(0.11)	(0.20)	(0.21)
AEL	-0.28	-0.40*	-0.10
	(0.19)	(0.24)	(0.12)
GINI $(t-1)$ × AE	-0.33	-	-
	(0.69)		
GINI $(t-1)$ × AL	-0.04	0.34	-
	(0.45)	(0.65)	
GINI $(t-1)$ × AEL	0.68	1.03	-
	(0.65)	(0.78)	
L/H-ratio $(t-1)$ × AL	-	-	1.00***
			(0.22)
L/H-ratio $(t-1)$ × AEL	-	-	0.05
			(0.05)
Round	-0.05	-0.04	-0.04
	(0.06)	(0.07)	(0.07)
Round squared	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)
Constant	2.27***	2.43***	1.91***
	(0.14)	(0.25)	(0.14)
Observations	4590	3456	3456
R^2 (overall)	0.1856	0.2009	0.1928
Wald-χ^2	383.93***	305.05***	554.18***

The dependent variable is contribution to the prevention account in round t. "Group accumulated $(t-1)$" denotes the amount in the prevention account after round t-1. "GINI $(t-1)$" denotes the inequality in accumulated contributions after round $t-1$. Similarly, L/H-ratio $(t-1)$ denotes the contribution ratio in accumulated contribution between low- and high-type subjects after round $t-1$. Robust standard errors are clustered across groups. Finally, *, **, and *** denote significance levels between 5% and 10%, between 1% and 5%, and at 1% or less, respectively.

AE, and AEL treatments, respectively, were "in the game" before the last round (no difference across treatments, Chi-square test), i.e., they could still reach the threshold. Among these groups, 87%, 92%, and 92% eventually reached the target in the AL, AE, and AEL treatments, but only 61% in the SY treatment. So the question is what happens in the last round that distinguishes symmetry from the heterogeneous treatments. By including only individuals from groups who reached the last round with at least €96 but less than €120 in the prevention account, Figure 7 indicates that the main determinant for success in the

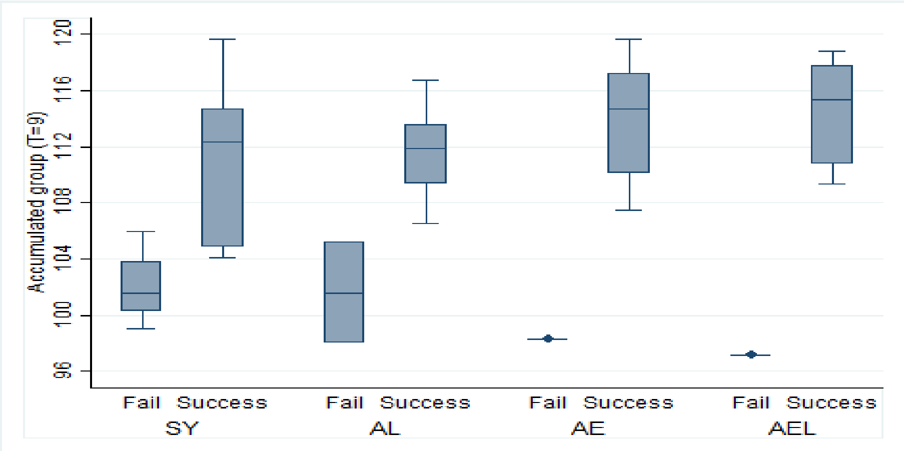


Figure 7: Box plots of accumulated amounts in the prevention account before the last round of the eventually successful and unsuccessful groups that are still "in the game"

last round is the amount accumulated in the prevention account at the beginning of that round.[15] But while in the heterogeneous treatments the distinction between successful and unsuccessful groups is fully predicted by the accumulated prevention account, this is not the case in the SY treatment, implying a coordination problem in this symmetric treatment. We can summarize our findings as follows.

Result 5b [Contributions in the last round]: *There are no differences between the treatments in the percentage of groups that can still prevent a catastrophe before the last round. Moreover, in the heterogeneous treatments the amount in the prevention account before the last round perfectly predicts success in reaching the threshold, which is not the case under symmetry.*

5.6 Efficiency

Finally, we turn to the question of overall efficiency in the different treatments. Table 4 shows the average expected payoffs after the 10th round but before nature (the random mechanism) draws "loss" or "no loss." It indicates that in each treatment the average expected payoff in successful groups is significantly higher than in unsuccessful groups. Also, we find no differences across treatments in pairwise comparison of expected payoffs in successful groups. However, for unsuccessful groups, we find a significant difference between the SY and AEL

[15]Using a logit model estimation, we find that in the heterogeneous treatments the accumulated amount in the prevention account perfectly predicts success in reaching the threshold.

treatments (two-sided MWU test, $p = 0.03$). The higher payoff for unsuccessful groups in the AEL treatment compared to the SY treatment stems from the fact that in the former the distinction between successful and unsuccessful groups evolves quite early in the game, so the group members (in unsuccessful groups) stop contributing earlier. In the SY treatment, the distinction between successful and unsuccessful groups evolves later on average than in the other treatments.

Table 4: Mean (and standard deviation) of total contributions to the prevention account and expected payoffs per treatment

Treatment:	SY	AL	AE	AEL
Total contributions to the prevention account				
Successful groups	123.30 (0.80)	124.56 (0.80)	123.75 (0.37)	124.05 (0.62)
unsuccessful groups	83.78 (8.89)	66.54 (14.51)	68.91 (10.46)	62.14 (6.55)
significance	***	***	***	***
Expected payoffs				
Successful groups	19.45 (0.38)	19.24 (0.50)	19.37 (0.22)	19.32 (0.38)
unsuccessful groups	13.01 (2.67)	14.69 (3.37)	14.25 (2.61)	15.75 (1.66)
significance	***	***	**	***

"Expected payoffs" are expected payoffs after the last round but before the lottery. Notably, ** and *** denote significance differences in a two-sided MWU test (between 1% and 5%, or less than 1%, respectively) in total contributions (upper panel) or expected payoffs (lower panel) between successful and unsuccessful groups.

Result 6 [Overall efficiency]: *In all treatments efficiency is significantly and considerably higher for successful than for unsuccessful groups in reaching the threshold and preventing a catastrophe. Moreover, there are no differences in overall efficiency between successful groups across treatments. However, the overall efficiency of unsuccessful groups is lower under symmetry than under double heterogeneity. The reason is that, under symmetry, unsuccessful groups are determined considerably later than under double heterogeneity, leading to a waste of resources in that treatment.*

6 Discussion

Our main finding is that heterogeneity between group members is not necessarily an obstacle to climate-change mitigation. Our result indicating that endowment heterogeneity does not reduce the success rate in reaching the threshold appears to contrast with Tavoni et al. (2011) and Burton-Chellew et al. (2013), but is consistent with Brown and Kroll (2017). In fact, success rates are quite high (statistically indistinguishable from the symmetric case) even under double heterogeneity where poor members face a higher loss rate than the rich. This result is somewhat at odds with Burton-Chellew et al. (2013), who find that in comparison with symmetry or endowment heterogeneity, double heterogeneity in endowments and risk probabilities reduces the success rate in reaching the thresholds. Arguably, our most remarkable result is that heterogeneity in loss rates may actually facilitate cooperation.

How can we explain the fact that heterogeneity favors group success? With homogeneous parties there is one obvious focal burden-sharing rule. But it seems plausible that if a subject does not comply with the "fair" equal-sharing norm, his/her group members have little inclination to compensate for that subject's attempt to free-ride. This is supported by the observation that, compared to the treatments with heterogeneities, considerably more groups with a chance of reaching the target before the last round fail to do so in the symmetric treatment. By contrast, since the equilibrium prediction under endowment or loss heterogeneity requires that the fortunate members (in endowment and loss) compensate their less fortunate counterparts, small deviations from any burden-sharing rule may be more generously forgiven than deviations from the obviously "fair" burden-share profile under symmetry.

Under endowment heterogeneity we observe that overall burden sharing is typically in line with a *proportional-contributions* profile. This finding is in accordance with our theoretical predictions. Under loss heterogeneity we find that overall burden sharing is more in line with the *equal-contributions* (=*equal-earnings*) resolution than with *proportional contributions*. This evidence is at odds with our theoretical predictions and also suggests that heterogeneity in loss rates is not perceived in the same way as heterogeneity in endowments. In fact, the distribution of first-round contributions under loss heterogeneity is strikingly similar to that under symmetry. In both treatments, 50% of the team members choose to contribute €2 to the prevention account (under loss heterogeneity not even one subject chose the *proportional-contributions* amount).

Hence, the considerably higher success rate under loss heterogeneity compared with symmetry may be due to two opposing effects. On the one hand, first-round contributions suggest that the *equal-contributions* and *equal-earnings* resolution is as focal under loss heterogeneity as under symmetry. On the other hand, due to built-in heterogeneity under loss heterogeneity, it appears that even when low-loss members are not sharing the burden equally, high-loss members will sometimes compensate for that (in fact, the burden sharing in 35% of the successful groups is classified as *proportional contributions*). In other words, the two treatments share a similar focal resolution norm, but loss heterogeneity has the additional advantage that it makes subjects more willing to forgive deviations from that norm. In support of the first point, i.e., that subjects do not really perceive loss heterogeneity differently from symmetry, the distributions of contributions indicate that subjects do not perceive the endowment heterogeneity differently than double heterogeneity in endowment & loss. Hence they seem to ignore the dimension of loss heterogeneity (this feature is consistent throughout the analysis).

How can we explain the behavioral differences between heterogeneity in endowments and in expected loss? One possible explanation may be that once subjects believe they can avoid the catastrophe, the heterogeneity in loss rates is no longer relevant and they are back in a completely symmetric situation. In that case, *equal contributions*, which also result in *equal earnings*, may be considered the fair outcome. In other words, many subjects seem to focus only on the final payoff in this subgame. In support of this explanation, we observe that under loss heterogeneity only 8% of the participants contribute nothing in the first round, whereas the remaining 92% revealed their preferences for catastrophe avoidance. Another possible and related explanation builds on the endowment effect (Kahneman et al., 1990): subjects focus on their endowments (which are equal) and ignore possible (heterogeneous) losses. Note that with this state of mind, the three allocation principles (equality of contributions/earnings and proportionality between inputs and outcomes) coincide. A third explanation for the empirical results is through social preferences. For example, using the inequity-aversion model (Fehr and Schmidt, 1999), it can be shown that the *proportional-contributions* profile can only be an equilibrium under relatively unlikely parameters, whereas an *equal-contributions=equal-earnings* profile is much more likely.[16]

Finally, the (relatively) high success rates in the treatments with heterogeneities are remarkable since in our design reaching the target does not yield higher group earnings than contributing nothing. Note that only in using these

[16]Fehr and Schmidt (1999) consider the following utility function:

$$U_i = \Pi_i - \alpha_i \max\{\Pi_j - \Pi_i, 0\} - \beta_i \max\{\Pi_i - \Pi_j, 0\},$$

where Π_i and Π_j are the respective payoffs of agents i and j, and α_i and β_i are parameters indicating dissatisfaction from inequality disadvantage and advantage, respectively. The *proportional-contributions* profile is not an equilibrium as long as $\beta_i/\alpha_i \leq 3/2$. On the other hand, for an *equal-contributions=equal-earning* profile to be an equilibrium, $\beta_i \geq 0.33$ is required.

design parameters, we can achieve the same unique catastrophe-prevention equilibrium under both endowment and expected loss heterogeneities. Any other design with a Pareto dominant contribution equilibrium will either not be unique or not equivalent for both heterogeneities. In our design, the equilibrium is the same for both heterogeneities satisfying only one of the focal resolutions, namely proportional contributions. Another reason for making the contribution equilibria as unattractive as possible is that, given the already high success rates with our conditions, the more attractive the contribution equilibrium is, the less likely it would be to identify possible effects due to heterogeneities. In fact, our design choice was enforced by an early pilot study where we set the contribution equilibrium to Pareto-dominate the non-contribution equilibrium. We found that in all treatments success rates were close to 100% and hence such a study is not suitable for revealing differences due to heterogeneities. Note, however, that since subjects are typically risk-averse (e.g., see Holt and Laury, 2002 and numerous subsequent studies), they strictly prefer the contribution equilibrium to the non-contribution equilibrium. Indeed, we observe that almost all subjects (between 86% and 93%) contribute to the public good in the first round of the game, empirically revealing preferences for the cooperative equilibrium (in accordance with risk aversion). Moreover, since in all groups there are some members who contribute to the prevention account in the first round, after that round the contribution equilibrium Pareto dominates the non-contribution equilibrium.[17] Finally, our design is supported by the observation that in all our treatments successful groups earn *significantly and considerably* higher payoffs than unsuccessful groups. So, *de facto* a contribution equilibrium dominates the non-contribution equilibrium.

7 Concluding Remarks

In this paper we use a variant of a threshold public-good game to study how heterogeneities across parties affect the provision of a common good: the prevention of a catastrophe. This game is used to model the international climate change dilemma (e.g., Milinski et al., 2008; Tavoni et al., 2011; Barrett and Dannenberg, 2012; Jacquet et al., 2013). Our findings give rise to greater optimism regarding successful outcomes in the provision of threshold public goods. First, we find

[17]Note also that in one respect our setting is in line with that of Tavoni et al. (2011). In their design, contribution and non-contribution profiles are payoff equivalent over ten rounds. It is only due to the sunk cost from contributions to the first three (passive) rounds that the contribution equilibrium Pareto-dominates the non-contribution equilibrium.

that, compared to a situation with homogenous parties, heterogeneity (with respect to either wealth, loss rates, or both) does not reduce contributions and success rates in preventing a catastrophe. Under heterogeneity, rich or severely affected members may feel more responsible and less concerned about the exact contributions of others. This result holds true even in the double heterogeneity case, where rich group members face lower loss-rates than poor members. In fact, under double heterogeneity (arguably the closest situation to the true international climate-change dilemma), it appears that in order to reach the target, (i) rich members must compensate the poor, and (ii) they should only compensate proportionally for one dimension of heterogeneity, namely wealth heterogeneity. This finding may be considered good news for supporters of international climate-change negotiations because it seems to reduce complexity. Parties that differ in many respects may reach an agreement by coordinating their mitigation efforts proportionally to their respective wealth or their ability to pay.

References

Ai, C. and Norton, E. C. (2003). Interaction terms in Logit and Probit models. *Economics Letters*, 80(1):123–129.

Alley, R. B., Marotzke, J., Nordhaus, W. D., Overpeck, J. T., Peteet, D. M., Pielke, R. A., Pierrehumbert, R. T., Rhines, P. B., Stocker, T. F., Talley, L. D., and Wallace, J. M. (2003). Abrupt climate change. *Science*, 299:2005–2010.

Barrett, S. and Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences USA*, 109:17372–17376.

Bernard, M., Reuben, E., and Riedl, A. (2011). On the focality of multiple fairness principles. Mimeo.

Brick, K. and Visser, M. (2015). What is fair? an experimental guide to climate change. *European Economic Review*, 74:79–95.

Brown, T. C. and Kroll, S. (2017). Avoiding an uncertain catastrophe: Climate change mitigation under risk and wealth heterogeneity. *Climatic Change*, 141:155–166.

Burton-Chellew, M., May, R., and West, S. (2013). Combined inequality in wealth and risk leads to disaster in the climate change game. *Climatic Change*, 120:815–830.

Croson, R. and Marks, M. (2000). Step returns in threshold public goods: A meta- and experimental analysis. *Experimental Economics*, 2:239–259.

Croson, R. and Marks, M. (2001). The effect of recommended contributions in the voluntary provision of public goods. *Economic Inquiry*, 39:239–249.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114:817–868.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiment. *Experimental Economics*, 10:171–178.

Frankhauser, S. (1994). The social costs of greenhouse gas emissions: An expected value approach. *Energy Journal*, 15:157–184.

Gallier, C., Kesternich, M., and Sturm, B. (2017). Voting for burden sharing rules in public goods games. *Environmental and Resource Economics*, 67(3):535–557.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92:1644–1655.

IPCC (2007). Summary for policymakers. In Metz, B., Davidson, O. R., Bosch, P. R., Dave, R., and Meyer, L. A., editors, *Climate Chhange 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2013). Summary for policymakers. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jacquet, J., Hagel, K., Hauert, C., Marotzke, J., Röhl, T., and Milinski, M. (2013). Intra-and intergenerational discounting in the climate game. *Nature Climate Change*, 3(12):1025.

Jacquet, J. and Jamieson, D. (2016). Soft but significant power in the Paris agreement. *Nature Climate Change*, 6(7):643–646.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98:1325–1348.

Kesternich, M., Lange, A., and Sturm, B. (2014). The impact of burden sharing rules on the voluntary provision of public goods. *Journal of Economic Behavior and Organization*, 105:107–123.

Kesternich, M., Lange, A., and Sturm, B. (2018). On the performance of rule-based contribution schemes under endowment heterogeneity. *Experimental Economics*, 21:180–204.

Konow, J. (2003). Which is the fairest one of all? a positive analysis of justice theories. *Journal of Economic Literature*, 41:1188–1239.

Lange, A., Löschel, A., Vogt, C., and Ziegler, A. (2010). On the self-interested use of equity in international climate negotiations. *European Economic Review*, 54(3):359–375.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Kagel, J. H. and Roth, A. E., editors, *The Handbook of Experimental Economics*, pages 111–194. Princeton University Press, Princeton.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J. (2008). Tipping elements in the earth's climate system. *Proceedings of the National Academy of Sciences USA*, 105:1786–1793.

Marotzke, J. (2000). Abrupt climate change and thermohaline circulation: Mechanisms and predictability. *Proceedings of the National Academy of Sciences USA*, 97:1347–1350.

Mendelsohn, R., Dinar, A., and Williams, L. (2006). The distributional impact of climate change on rich and poor countries. *Environment and Development Economics*, 11:159–178.

Milinski, M., Röhl, T., and Marotzke, J. (2011). Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Climatic Change*, 109:807–814.

Milinski, M., Semmann, D., Krambeck, H. J., and Jochem, M. (2006). Stabilizing the earth's climate is not a losing game: Supporting evidence from public goods experiments. *Proceedings of the National Academy of Sciences USA*, 103:3994–3998.

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

Parry, M. L. (2007). *Climate Change 2007: Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Fourth Assessment Report of the IPCC Intergovernmental Panel on Climate Change*, volume 4. Cambridge University Press.

Raihani, N. and Aitken, D. (2011). Uncertainty, rationality and cooperation in the context of climate change. *Climatic Change*, 108:47–55.

Randalls, S. (2010). History of the 2°c climate target. *Wiley Interdisciplinary Reviews: Climate Change*, 1:598–605.

Rapoport, A. and Suleiman, R. (1993). Incremental contribution in step-level public goods games with asymmetric players. *Organizational Behavior and Human Decision Process*, 55:171–194.

Reuben, E. and Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, 77:122–137.

Tavoni, A., Dannenberg, A., Kallis, G., and Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences USA*, 108:11825–11829.

Tol, R. S. J. (2009). The economic effects of climate change. *Journal of Economic Perspectives*, 23:29–51.

A Proofs of Propositions

A.1 Proof of Proposition 1

First we ignore (7) and solve the equation system (8), (9), (10) for $j = L, H$, and (9), for $j = L, H$ to obtain

$$q_L = \frac{(2pq-1)a_H + a_L}{(a_L + a_H)p} \tag{A.1}$$

$$q_H = \frac{(2pq-1)a_L + a_H}{(a_L + a_H)p} \tag{A.2}$$

For maximum willingness to contribute, we obtain

$$\hat{c}_L = pqa_L \tag{A.3}$$

$$\hat{c}_H = pqa_H \tag{A.4}$$

$$\breve{c}_L = pq_L a \tag{A.5}$$

$$\breve{c}_H = pq_H a \tag{A.6}$$

If we now substitute (A.1) and (A.2) into (7), we obtain $q = 1/2p$. Substituting this into (A.3), (A.4), and (A.5), (A.6), respectively, we can verify that the equations (11) and (13) are satisfied. Thus, for $q = 1/2p$, maximum willingness to contribute is just enough to prevent the climate catastrophe. Therefore by definition of maximum willingness to contribute, risk-neutral players are indifferent between the two equilibria.

A.2 Proof of Proposition 2

Note that under selfish preferences and risk neutrality, maximum willingness to pay for preventing the loss event is implicitly given by $a_j - \tilde{c}_j = \tilde{q}_j a_j$, yielding

$$\tilde{c}_j = (1 - \tilde{q}_j)a_j = pq_j a_j \tag{A.7}$$

$j = L, H$. Summing up the contributions of both players gives us

$$\tilde{c}_L + \tilde{c}_H = p(q_L a_L + q_H a_H) \tag{A.8}$$

Using $q_H = 2q - q_L$, we can rewrite this as

$$\tilde{c}_L + \tilde{c}_H = p(q_L a_L + (2q - q_L)a_H) \equiv f(q_L) \tag{A.9}$$

Using $pq = 1/2$ for $q_L = q$ we would obtain

$$\tilde{c}_L + \tilde{c}_H = pq(a_L + a_H) = \frac{a}{2} \tag{A.10}$$

But $f'(q_L) = p[a_L - a_H] < 0$. Therefore

$$\tilde{c}_L + \tilde{c}_H = p(q_L a_L + (2q - q_L)a_H) < \frac{a}{2}$$

for $q_L > q > q_H$.

A.3 Proof of Corollary 1

We give the proof for the AE game. For AL the proof is similar.

Since $q = 1/2p$, we obtain

$$a_j - \hat{c}_j = [1 - p + p(1 - q)]a_j = (1 - pq)a_j = \frac{a_j}{2}, \qquad j = L, H$$

i.e., for the risk-neutral subject, the monetary payoffs are equal in both types of equilibria. This implies that for a risk-averse subject we have

$$U(a_j - \hat{c}_j) = U((1-pq)a_j) = U((1-p+p(1-q))a_j > (1-p)U(a_j) + pU((1-q)a_j)$$

B Further figures

Table B.1: Classification of successful group according to the focal resolution

Treatment	Share of high-type	Classification
AL	0.51	*equal contributions=equal earnings*
AL	0.49	*equal contributions=equal earnings*
AL	0.59	*proportional contributions*
AL	0.61	*proportional contributions*
AL	0.58	*proportional contributions*
AL	0.49	*equal contributions=equal earnings*
AL	0.46	*equal contributions=equal earnings*
AL	0.52	*equal contributions=equal earnings*
AL	0.46	*equal contributions=equal earnings*
AL	0.44	no classification
AL	0.60	*proportional contributions*
AL	0.59	*proportional contributions*
AL	0.48	*equal contributions=equal earnings*
AL	0.48	*equal contributions=equal earnings*
AE	0.62	*proportional contributions*
AE	0.60	*proportional contributions*
AE	0.65	*proportional contributions*
AE	0.60	*proportional contributions*
AE	0.48	*equal contributions*
AE	0.64	*proportional contributions*
AE	0.69	*equal earnings*
AE	0.61	*proportional contributions*
AE	0.59	*proportional contributions*
AE	0.59	*proportional contributions*
AE	0.66	*equal earnings*
AE	0.64	*proportional contributions*
AEL	0.64	*proportional contributions*
AEL	0.63	*proportional contributions*
AEL	0.63	*proportional contributions*
AEL	0.59	*proportional contributions*
AEL	0.64	*proportional contributions*
AEL	0.60	*proportional contributions*
AEL	0.66	*equal earnings*
AEL	0.62	*proportional contributions*
AEL	0.61	*proportional contributions*
AEL	0.59	*proportional contributions*
AEL	0.56	*proportional contributions*
AEL	0.50	*equal contributions*
AEL	0.58	*proportional contributions*

The table indicates the contribution shares of high-type members (defined as high-loss or high-endowment members in the AL and AE treatments and high-endowment & low-loss members in the AEL treatment). The classification criteria are described in section 5.2.

Table B.2: Logit model estimations of success probability

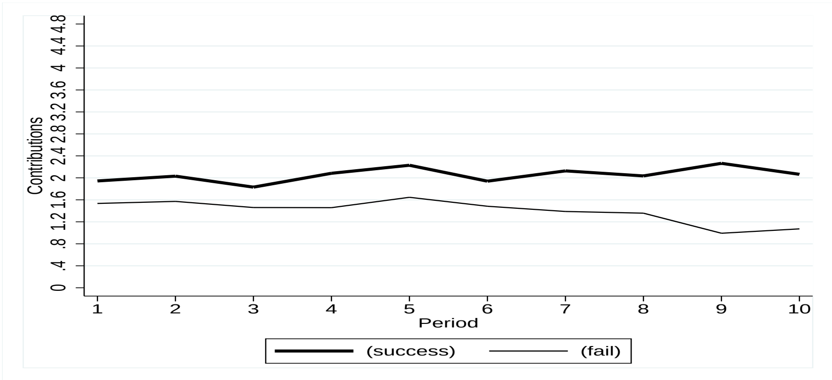
	SY	AL	AE	AEL
Contributions ($t = 1$)	0.90**	0.56**	1.37**	0.91**
	(0.40)	(0.25)	(0.57)	(0.37)
Cons	-9.67**	-5.48**	-14.47**	-9.61**
	(4.14)	(2.73)	(6.13)	(3.95)
Obs	21	21	21	22
Wald chi2(1)	8.65***	7.21***	14.93***	13.08***
Pseudo R^2	0.30	0.26	0.52	0.43

The dependent variable: success rate in reaching the €120 threshold. Robust standard errors are clustered across groups. Finally, *, **, and *** denote significance levels between 5% and 10%, between 1% and 5%, and at 1% or less, respectively.

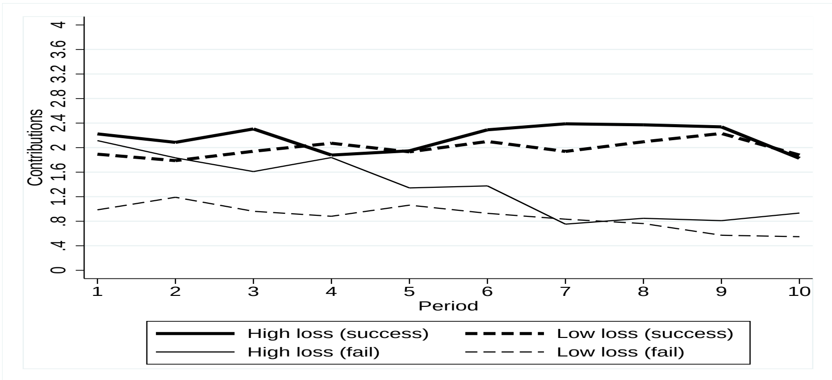
Table B.3: Random-effect estimations of contribution decisions over time

	High type	Low type
Group accumulated ($t-1$)	0.02***	0.01***
	(0.00)	(0.00)
L/H-ratio ($t-1$)	-0.36***	0.06
	(0.05)	(0.06)
AL	-0.68***	-0.78***
	(0.23)	(0.20)
AEL	-0.49***	0.22**
	(0.16)	(0.10)
L/H-ratio ($t-1$) × AL	0.61***	1.36***
	(0.22)	(0.23)
L/H-ratio ($t-1$) × AEL	0.30***	-0.13**
	(0.05)	(0.06)
Round	-0.10	-0.03
	(0.09)	(0.07)
Round squared	-0.02***	-0.01**
	(0.01)	(0.00)
Constant	2.50***	1.44***
	(0.17)	(0.16)
Observations	1728.00	1728.00
R^2 (overall)	0.2521	0.2202
Wald-χ^2	426.76***	1127.80***

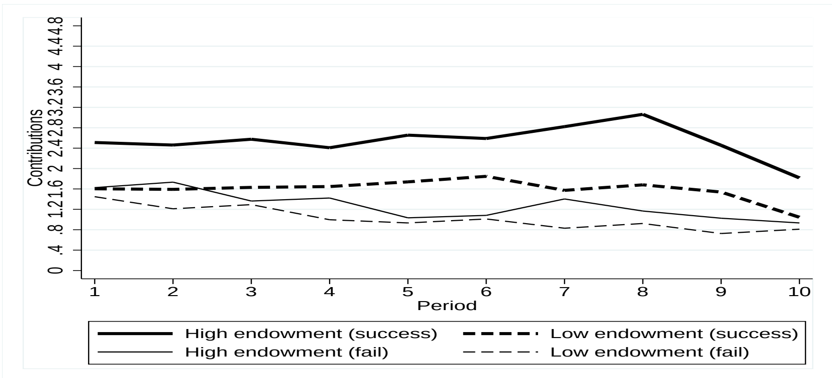
The dependent variable is contribution to the prevention account in round t. "Group accumulated ($t-1$)" denotes the amount in the prevention account after round t-1. "GINI ($t-1$)" denotes the inequality in accumulated contributions after round $t-1$. Similarly, L/H-ratio ($t-1$) denotes the contribution ratio in accumulated contribution between low- and high-type subjects after round $t-1$. Robust standard errors are clustered across groups. High types are defined as high-loss or high-endowment members in the AL and AE treatments, and high-endowment & low-loss members in the AEL treatment. Finally, *, **, and *** denote significance levels between 5% and 10%, between 1% and 5%, and at 1% or less, respectively.



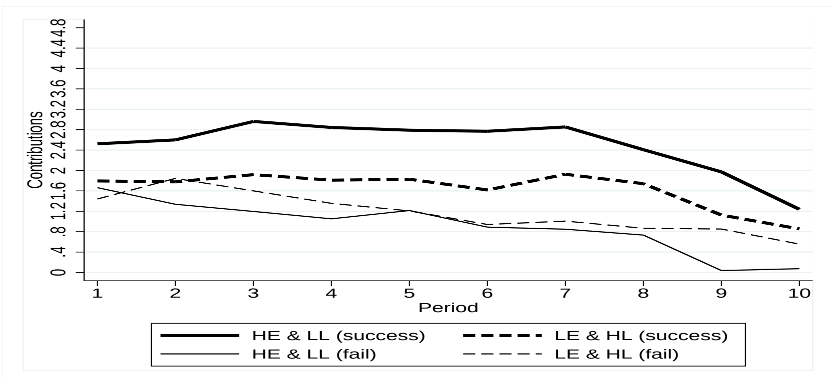
(a) SY treatment



(b) AL treatment



(c) AE treatment



(d) AEL treatment

Figure B.1: Evolution of average contribution decisions of group-member types in successful and unsuccessful groups