



**HAL**  
open science

## Simulating collective risk management from experimental data

Omid Roozmand, Guillaume Deffuant, Giulia Andrighetto, Mario Paolucci

► **To cite this version:**

Omid Roozmand, Guillaume Deffuant, Giulia Andrighetto, Mario Paolucci. Simulating collective risk management from experimental data. Social Simulation Conference 2022, ESSA, Sep 2022, Milan (Italie), Italy. hal-04211468

**HAL Id: hal-04211468**

**<https://hal.science/hal-04211468>**

Submitted on 19 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Simulating collective risk management from experimental data

Omid Roozmand<sup>1</sup>, Guillaume Deffuant<sup>1,2</sup>, Giulia Andrighetto<sup>3</sup>, and Mario Paolucci<sup>3</sup>

<sup>1</sup> Université Clermont Auvergne, Inrae, UR LISC, France [guillaume.deffuant@inrae.fr](mailto:guillaume.deffuant@inrae.fr)

<sup>2</sup> Université Clermont Auvergne, LAPSCO, France.

<sup>3</sup> Istc, CNR, Italy

**Abstract.** We propose a new analysis of the data from a large-scale experiment about managing collective risk (Szekely et al. 2021). We distinguish three behaviours (free riding, default cooperation, and fluctuating contribution) which were mixed in previous analyses. We show that free riding and default cooperation strongly determine the outcome of the game. Social norm features (personal normative belief, empirical expectation, normative expectation) are significant predictors of these behaviours but the decision models derived with or without these variables perform very similarly on the data. However, we also run two agent based models, one with agents not using normative features (reactive agents) the other with agents using them (normative agents). In the average result of numerous simulations, the normative agents tend to keep a significantly lower frequency of free-riding and a significantly higher frequency of default cooperation. This confirms the previous analyses concluding that the normative variables play a key role in the emergence of a norm.

**Keywords:** collective risk, social norm, free-rider, agent-based model.

## 1 Introduction

Human cooperation is key for the success of human societies and is present in some form in almost every social relationship, from the family to the nation state to global organizations. Yet cooperation is often individually costly, making it inherently fragile. Many scholars have thus concentrated on understanding how to sustain it. Recently, researchers have proposed social norms as potential solutions to major cooperation problems, from the management of water resources to the reduction of smoking in public places, and the increase of vaccination [1]. Social norms can be broadly defined as informal and shared behavioural rules that, unlike legal norms, are not codified, but are learnt through social interaction. Those rules prescribe what individuals ought or ought not to do, and they are followed because of shared expectations of conformity to them and potentially social sanctions [2,3,4,5]. Although social norms are often cited as the cause of many social phenomena, there is still little experimental research about norm-based explanations [6]. A growing number of researchers claims that their results are consistent with social norms, but few of them measure norms directly and examine if they have causal effect on behavior [2,7,8,9]. This raises the question of whether social norms are indeed causal drivers of behaviour and can convincingly explain major cooperation-related regularities.

With the aim of assessing the role of social norms in promoting human cooperation, the long-term (30 days) experiment by Szekely et al. [10] is designed to test whether social norms of cooperation change due to risk and if they causally motivate behavior. Szekely et al. [10] use a collective risk

dilemma game similar to the setup of Milinski et al. [11] to assess under what conditions groups are able to socially coordinate to avoid the risk of a collective loss (see Section Experimental Design, for more details). Collective-risk social dilemmas are situations in which individuals have to collectively reach a target by contributing to a collective fund and avoid the risk of collective loss (a “disaster”). This experimental setting is coupled with an extensive set of measures [12] to detect the presence of social norms, their causal effect on behaviour, and their change over time. The analysis of the experimental results, mainly based on linear models of the contribution, suggest that a social norm is instituted and that the disaster probability modifies this norm.

In this paper, we revisit the data of this experiment and introduce the following new elements in this analysis:

- We distinguish between three different behaviours of the players: free riding (contribution 0), default cooperation (contribution 50) and a fluctuating contribution (mainly between 40 and 70). We show that free riding and default cooperation strongly determine the game outcome;
- With this distinction, we derive two decision models of the contribution in the game: (1) reactive model, based only on past data and (2) normative model, which is also based on normative features (normative beliefs, empirical and normative expectations). We find that the normative features are significant for predicting free riding and default cooperation, which is consistent with Bicchieri’s thesis [12]. However, we do not notice any significant difference between the two models in fitting the data;
- We developed reactive and normative agents using respectively the reactive and normative decision models. The normative model of the contribution is combined with models of the normative features, based on data. When we modify these models, making them slightly more or less trustful, the impact on the outcome of the game is spectacular.

The paper is organised in a classical manner. The next section describes the data and methods, including a description of the experiment, the used data and the modelling choices for decision models and their use in the agent based simulations. Then the following section describes the results of the decision models and the agent based simulation. The final section discusses these results in the perspective of the emergence of norms.

## 2 Material and Methods

In this section we describe the experiment, the resulting data and methods applied to design the reactive and normative decision models and then we describe the agent-based models.

### 2.1 Experiment on collective risk social dilemma

The experiment is described in details in [10]. We only summarise here its main features. Participants (n=296) are matched in groups of six people and interact on-line according the collective risk social dilemma [11] answering questions on their empirical and normative expectations and deciding their actual contribution and their conditional contributions for 28 consecutive days.

Each day/round, participants are assigned a constant initial money (100 points) and are grouped randomly with other 5 participants in a way that they do not know each other. Subjects can avoid the risk of collective loss by investing money from their private account into a public project that is able to protect from the loss only if a minimal threshold is reached (300 points). If the threshold is

reached, the disaster is averted with certainty and subjects earn the amount that was kept in their private account; otherwise, subjects risk losing their earnings for that round with probability  $p$ . At the end of every round, subjects are informed about the contribution of their group members, the outcome, if a disaster occurred, and their individual payoff for that round.

The objective of (Szekely et al. 2021) is to study the dynamics of social norms under collective risk, whether social norms causally motivate behavior and how this affects the ability of groups to solve cooperation problems. Following Bicchieri (2006), Szekely et al. (2021) defines social norms as informal behavioral rules that individuals follow conditionally on their believing that: (i) a sufficiently large number of people in their community conform to the rule (empirical expectations), and (ii) a sufficiently large number of people in their community think that they ought to conform to the rule and may be willing to sanction transgressions (normative expectations). In this view, a social norm exists when there is a set of individuals who are disposed to follow a behavioral rule because they believe that both these conditions are fulfilled.

First, to test how risk affects social norms and cooperation, the disaster probability (high: 0.9 vs low: 0.6) and the order in which subjects face different risks are manipulated. Participants play in two different sequences of disaster probability. In the High-Low sequence, subjects first experience high probability of disaster ( $p=0.9$ ) (rounds 1-14) followed by low probability ( $p=0.6$ ) (rounds 15-28); while in Low-High sequence, this order is reversed. Second, to assess the existence of social norms, empirical expectations and normative expectations are elicited in each of the 28 game-days to detect the basic conditions for norms to motivate behavior (Bicchieri et al. 2014). To ensure truth-telling, the questions for eliciting and measuring social norms are incentivised such that participants receive a payment for each correct answer. Finally, to identify the causal effects of empirical and normative expectations on behavior, social expectations are manipulated in a “conditional contribution” phase (i.e., with the strategy method) on a subset of the game rounds. Results from the experiment show greater cooperation and stronger social norms for higher risk and slower behavior change after an exogenous change in risk when social norms are stronger. Moreover, social norms are shown to predict cooperation and causally affected behavior.

## 2.2 Data and data treatment

Table 1 breaks down the main data from the experiment that we use in our analysis. Some of these variables are collected once for each participant (psychological traits) others are collected at each round (contribution, group id,...). We added the variable called "responsible", whose value is 1 if there is a disaster and the participant contributed less than 50 and 0 otherwise. Indeed, this variable turned out to be significant in predicting the contribution.

**Distinguishing free-riding, default cooperation and fluctuating contribution.** Looking at the whole distribution of contributions (Figure 1), it appears that contributions 0 and 50 are particularly frequent with respect to the frequency in the neighbourhood of these values. Indeed, the participants contributed 297 times zero during the whole game and this appears as an isolated peak in the distribution. Similarly, participants contributed 50 for 3539 times overall (over 8008 in total, so about 44%). The rest of the contributions vary in majority between 40 and 70, with more frequent contributions above 50.

Contributing 0 is the most non-cooperative behaviour in the game. Indeed, if there is no disaster or if there is a disaster and all the others contributed enough then the participant keeps her 100

NO	Parameter's Name	Description	Value
1	disaster sequence	general categories of agents	1, 2
2	round	the day of the game	[1, 28]
3	id	the id of each participant which remains unchanged during the game	[1, 286]
4	group_id	the id assigns to each group which changes every day	random
5	disaster_probability	the risk of disaster	0.6, 0.9
6	contribution	amount of money that participant contribute at reach round	[0, 100]
7	total_contribution	total sume of contributed money by all members of a group at each round	[0, 600]
8	pnb	personal normative belief	[0, 100]
9	ee	empirical expectations	[0, 100]
10	ne	normative expectations	[0, 100]
11	unconditional_payoff	the amount of money that participant earns at each round (100-contribution)	[0, 100]
12	disaster_outcome	the occurrence of disaster for each group at each round (1: true, 0: false)	0, 1
13	svo	social value orientation	[-11.68, 61.19]
14	asq	autism spectrum quotient	[6, 44]
15	risk	the personal trait related to risk acceptance of participant	[0, 5]
16	extraversion	personality trait of participant	[8, 40]
15	agreeableness	personality trait of participant	[17, 45]
17	conscientiousness	personality trait of participant	[17, 45]
18	neuroticism	personality trait of participant	[9, 40]
19	openness	personality trait of participant	[20, 50]
20	female	if the participant is female it is 1 otherwise 0	0, 1
21	threshold	An indicator if disaster occurs (if total_contribution < threshold then disaster occurs)	300

**Table 1.** Key variables of the game which is used in this paper

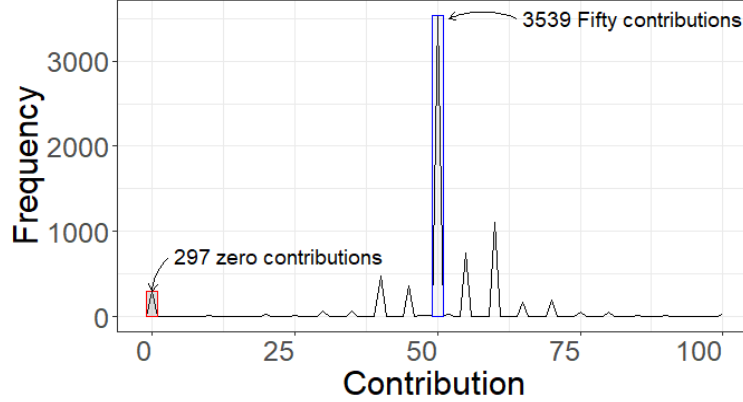
tokens. Hence the participant does not participate in the common effort against the risk and takes advantage from the contributions of the others. Therefore we call this behaviour "free-riding".

Contributing 50 is assuming that the burden of the risk is evenly supported by all participants. Therefore, we call this behaviour "default cooperation".

Finally, the other contributions appear to fluctuate with circumstance and the psychological profile of the participant. Therefore, we call this behaviour "fluctuating contribution".

In the following, we present the method that we used to model if the participants contribute 0 or 50 and if not, which other value they contribute.

**Decision trees of step-wise linear regressions.** To choose a model for predicting 0 or 50 or other contributions, we tested different methods including simple linear regression model, step-wise regression, decision tree, and finally we chose binary decision trees of linear models generated by step-wise regression. The potentially splitting variables of the tree are the ones that have only two modalities including disaster\_probability, responsible (t-i), disaster\_outcome (t-i), female. A split is chosen when the splitting variable is selected in a step-wise linear regression applied to the considered data set (or a subset defined by a splitting variable). In this way, the model can take into account changes of behaviour which depend for instance on the probability of disaster (which has two modalities). The step-wise regression is executed recursively on the subset of data of each



**Fig. 1.** Distribution of contribution. Zero and 50 appear as more frequently used than the values in their neighbourhood.

branch until no more splitting variable is chosen or the number of records of the subset gets lower than minimum value (`Min_split`).

The reason for this choice is that the values of binary variables can put the participant in different situations which could lead to different decisions. For instance if the participant feels responsible for a disaster or not, or if the participant plays in a high risk situation (`disaster_probability = 0.9`) or low risk situation (`disaster_probability = 0.6`), then the decision making may change. The regression trees automatically detect the relevance of such distinctions.

For the tree predicting the fluctuating contribution, we do not take as inputs the previous contributions. Indeed, the tree is trained only on the specific cases where the output is not 0 nor 50, whereas afterwards, in simulations, the model is used for other data on which it has not been trained. It turned out that the model was not robust enough for this when taking as inputs the previous contributions.

**Kernel estimation on regression tree leaves for probability estimation.** Each leaf of the tree includes a linear model derived from a step-wise regression. However, the linear model does not guarantee to output a number between 0 and 1 when trained on the presence of contribution 0 or 50. Therefore, on each leaf of the tree, we use a kernel estimation of the probability, based on the output of the linear model. We chose a simple sliding window, characterised by its width  $w$ , as kernel. Consider a leaf of the tree characterised by the data set  $D$  and the linear model  $L$ . The output by the kernel estimation for the input data vector  $x$  is:

$$K(x) = \frac{\sum_{y \in D} (L(y) \in [L(x) - w, L(x) + w]) \cdot (c(y) = v)}{\sum_{y \in D} (L(y) \in [L(x) - w, L(x) + w])} \quad (1)$$

where  $c(y)$  is the contribution at  $t$  in data vector  $y$  and  $v$  is 0 or 50.

For each leaf of the tree, we determine the width  $w$  of the sliding window that minimises the cross validation error.

**Reactive and normative decision models.** The reactive and normative decision models include three regression trees:  $T_0$  predicting probability that the participant contributes 0,  $T_{50}$  predicting the probability that the participants contributes 50 and  $T_c$  predicting the participant’s fluctuating contribution.

For the reactive model, each regression tree is built on data from the following variables, for  $t$  the round varying from 4 to 28 and for  $i = 1, 2, 3$  the distance in the past: Contribution at round  $t - i$ , Unconditional Payoff at round  $t - i$ , average contribution of other members of the group at time  $t - i$  of others, the disaster outcome at time  $t - i$ , responsible  $t - i$  the disaster probability, female, 8 psychological variables.

For the normative decision model, the data include also the normative variables: personal normative belief at rounds  $t - i$ , empirical expectation at rounds  $t - i$ , normative expectation at rounds  $t - i$ .

For each of the first three rounds, we build specific decision models (reactive and normative) in which we only take into account the variables that do not depend on  $t$  for round 1, we add the variables at  $t - 1$  for round 2, and the variables at  $t - 1$  and  $t - 2$  for round 3.

### 2.3 Agent Models replicating the game

We are interested in comparing the game played by reactive or normative agents and the data. Indeed, running it with a large number of agents, drawn at random from the participants profiles, provides indications about the robust outcomes. Moreover, we can test the sensitivity of the outcome of the game to some changes in the decision models.

**Reactive Agent.** The decision process of reactive agents uses the reactive decision models for the first 3 rounds and then the one for rounds 4 to 28 presented previously, in a non-deterministic way.

Indeed, the decision process uses a uniformly drawn value  $\theta \in [0, 1]$ , a value  $\epsilon$  drawn from a Gaussian distribution  $N(0, \sigma)$  of mean 0 and standard deviation  $\sigma$ . The output  $M_{\theta, \epsilon}(x)$  given to input  $x$  is determined as follows:

$$M_{\theta, \epsilon}(x) = \begin{cases} 0, & \text{if } \theta \in [0, T_0(x)]; \\ 50, & \text{if } \theta \in (T_0(x), T_0(x) + T_{50}(x)]; \\ T_c(x) + \epsilon, & \text{if } \theta \in (T_0(x) + T_{50}(x), 1]. \end{cases} \quad (2)$$

The value of  $\sigma$  is calibrated so that the standard deviation of  $M_{\theta, \epsilon}$  is close to the standard deviation of the non-0 non-50 contributions in the data.

**Normative Agents.** The normative agents are based on the normative decision process that takes the normative variables at  $t - i$  as additional inputs. The values of these variables vary over time and with the game circumstances for each agent. Therefore, we developed decision models based on regression trees that predict the values of these variables. The models are similar to those predicting the contributions except that, since there are very few contributions 0 in the normative variables, we only develop a tree predicting the contribution 50 and a tree for the fluctuating values. These two trees are derived for each of the three cognitive variables, for the three first rounds and then for round 4 to 28.

Moreover, to each model of a normative variable, we add a Gaussian noise to  $T_c(x)$ . The standard deviations of these noises are more difficult to determine than the one of the reactive model because these noises depend on the noise of the contribution and reciprocally. Indeed, the models of the normative variables take the contributions (at  $t - i$ ) as input and reciprocally, the contribution model takes the normative variables as input (at  $t - i$ ).

We solve this problem by an iterative calibration process. At the first iteration, we fix the standard deviation of the contribution noise at the value determined for the reactive model and we calibrate the standard deviations for the noise of the normative variables (by comparing with the standard deviation in the data on the non-50 values of these variables). Then we fix the noises of the normative variables and we calibrate the noise of the contribution. We iterate this process until getting good values for both the normative variables and the contribution.

**Replicating the game.** We initialise 300 agents, 150 for each sequence of disaster probability, drawing their psychological profile at random among the participants of this sequence. At each round, we randomly define 50 groups of 6 agents for each sequence and they play the game, using their decision model (with specific models for the three first rounds and then a model for rounds 4 to 28). We repeat 100 runs like this and average the results over these runs.

### 3 Results

#### 3.1 Decision models.

As an example, Figure 2 shows the regression tree  $T_0$  (contribution 0) of the reactive decision model for rounds 4-28. The variable `cont0_1` indicating whether the agent has played zero at  $t-1$  is the first splitting variable which is chosen. Then the data is divided to two subsets, e.g. subset of participants who have not played zero at round( $t-1$ ) (`cont0(t-1) = 0`), and subset of data which includes the participants who have played zero at round ( $t-1$ ) (`cont0(t-1) = 1`). The leaves are figured in blue and each leaf includes a linear model, result of a step-wise regression on the data set defined by the leaf. The process is recursively applied on each subset to create the whole tree as shown on figure 2.

Figure 3 shows the results of reactive and normative models tested against data for  $T_0$ ,  $T_{50}$  and  $T_c$ , distinguishing high-low from low-high disaster probability.

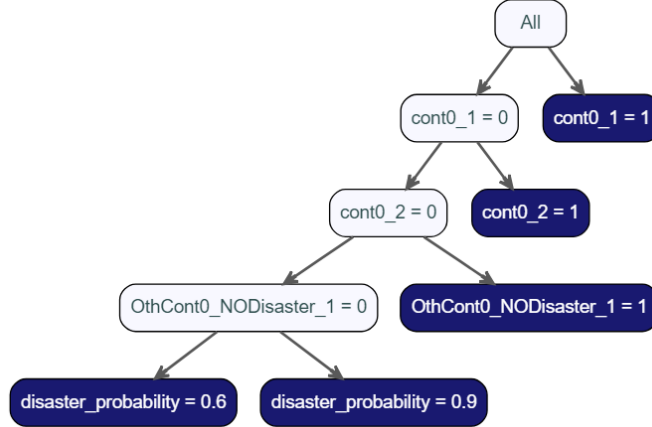
The probability to contribute 0 is 0 when the disaster probability is high (0.9), and it is increasing over time up to about 0.1, when the disaster probability is low (0.6). Not surprisingly, the free-riding behaviour is encouraged when the risk is lower. Moreover, a contagion process of the free-riding behaviour seems very likely, as the variable coding for the presence of other contributing zero in previous rounds is selected by the step-wise regression in several leaves of regression tree  $T_0$  (both for reactive and normative models).

The probability to contribute 50 (default cooperation behaviour) is increasing over time for the low-High sequence of disaster probability, while it is increasing the slightly decreasing and increasing again fir the High-Low sequence. Hence the high disaster probability seems to favour the default behaviour and a contagion process is also very likely as the variable coding for the presence of other contributing 50 is selected by the step-wise regression in several leaves of  $T_{50}$  (both for reactive and normative models).

The models  $T_0$  and  $T_{50}$  show very similar results and fit the data quite satisfactorily.

The results of trees  $T_c$  (right panel) show that the average value of the non-0 non-50 contributions fluctuates over the rounds without showing a clear influence of the disaster probability. This average





**Fig. 2.** Example of regressions tree.  $T_0$  for reactive decision model, for rounds 4-28. The nodes indicate the value of a binary variable corresponding to a split. The dark blue nodes (leaves) indicate a final linear regression model.

is higher than 50 for all rounds. The results of the reactive and normative models are very similar the sequences Low-High and High-Low disaster probabilities do not seem to impact the values.

Table 2 shows the r-squared of reactive and normative models for  $T_0$ ,  $T_{50}$  and  $T_c$ . It shows that the  $R^2$  of  $T_0$  and  $T_c$  are significantly better for the normative decision model. However, this difference is not noticeable when comparing the average results per round of game as shown on figure 3.

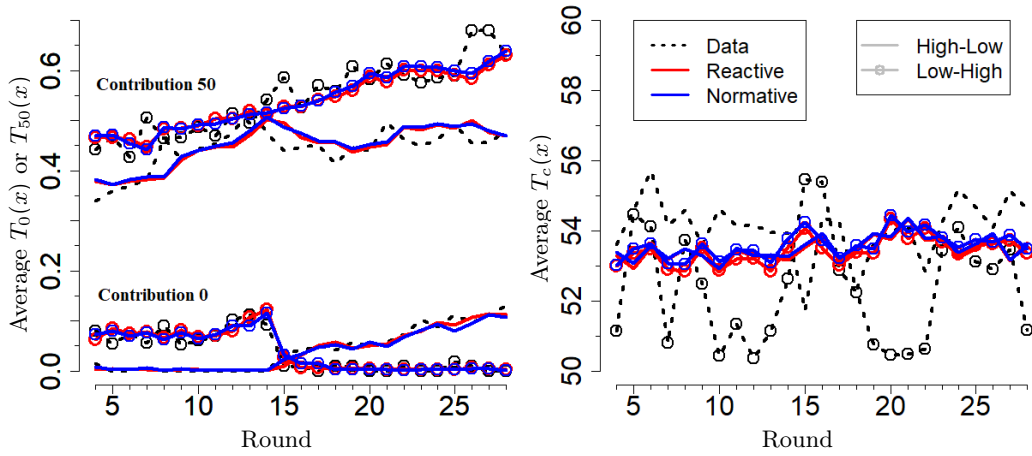
	$T_0$	$T_{50}$	$T_c$
Reactive model	0.4649	0.4807	0.0977
Normative Model	0.5395	0.4884	0.0936

**Table 2.** R-Squared of regression trees  $T_0$ ,  $T_{50}$  and  $T_c$  in the reactive and normative decision models.

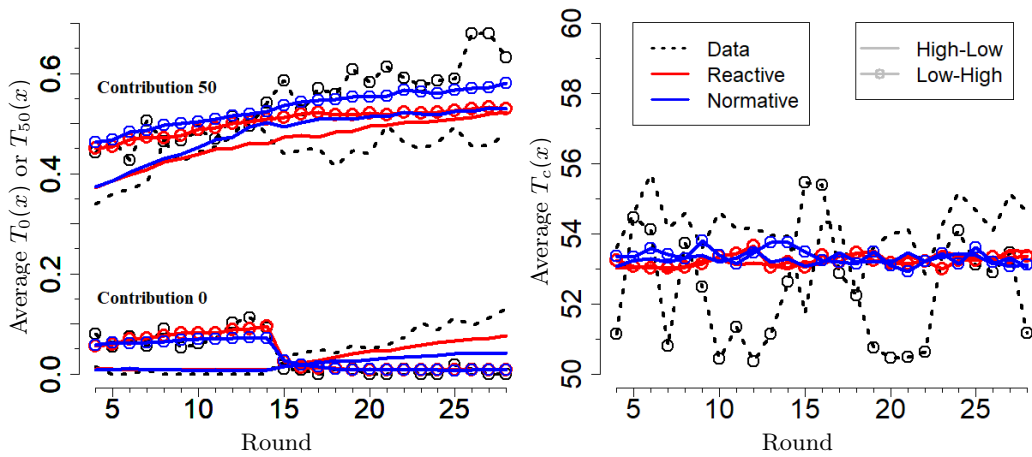
### 3.2 Agent based simulations

Figure 4 shows the results of the agent based simulations (average over 100 replicas of the game with 300 agents). Whereas the reactive and normative decision models tested on the data provide very similar results, several interesting differences appear between the results of reactive and normative agent simulations, and with the results of the decision models.

A striking difference between the results of the decision models on data (Figure 3) is that the number of 0 contributions in agent simulations during rounds 15-28 (second half of the game) is significantly smaller for the High-Low sequence of disaster probability. This is true for the reactive agents (red bottom curve of Fig. 4, left panel) but the effect is more pronounced for the normative agents (blue bottom curve of Fig. 4, left panel). Indeed, at round 28, the probability to contribute 0 is around 0.12 in the data, about 0.8 for the reactive agents and about 0.4 for the normative agents.



**Fig. 3.** Decision models against data. Left panel: contribution 0 and 50 from data and from trees  $T_0$  (bottom curves) and  $T_{50}$  (top curves), for reactive and normative decision models. Right panel average fluctuating contribution per round for data and  $T_c$  for reactive and normative decisions models. High-Low and Low-High refer to the disaster probability sequence in the experiment.



**Fig. 4.** Agent simulations. Left panel: contribution 0 and 50 from data and agents (reactive and normative) over the rounds. Right panel average fluctuating contribution per round for data and agents (reactive and normative). High-Low and Low-High refer to the disaster probability sequence in the experiment.

This means that there is a much more significant impact of the sequence of disaster probability on the probability of free riding in the agent models, and particularly in the normative agent model. This indicates the tendency for the normative agent to adopt less frequently the free-riding behaviour, when starting by a period where this behaviour is not present.

Another important difference, the probability to contribute 50 for the High-Low sequence (top curves without points in the left panel) is higher than the data for both the reactive and normative agents and more so for the normative agents (blue curve). On the contrary for the Low-High sequence, the probability to contribute 50 (top curves with points in the left panel). The number of 50 contributions is higher for the normative agents than for the reactive agents though, indicating a tendency for normative agents to increase rapidly the default cooperation behaviour when the risk is high and then to keep or still increase slowly this frequency when the risk decreases.

The right panel shows that the fluctuating contribution is almost flat at an average contribution of around 53.5 and no significant difference appear between reactive and normative agents or the sequence of disaster probability. This can be explained by the averaging effect of the 100 replicas and it suggests that the fluctuations in the data and in the decision models shown on the right panel of Figure 3 are not significant.

Overall, the sequence of disaster probability has a stronger impact on the agent simulations and more particularly on the normative agents. When the risk is high in the first half of the game, the normative agents tend to stick to the corresponding behaviour, with a low frequency of free-riding and a high frequency of default cooperation.

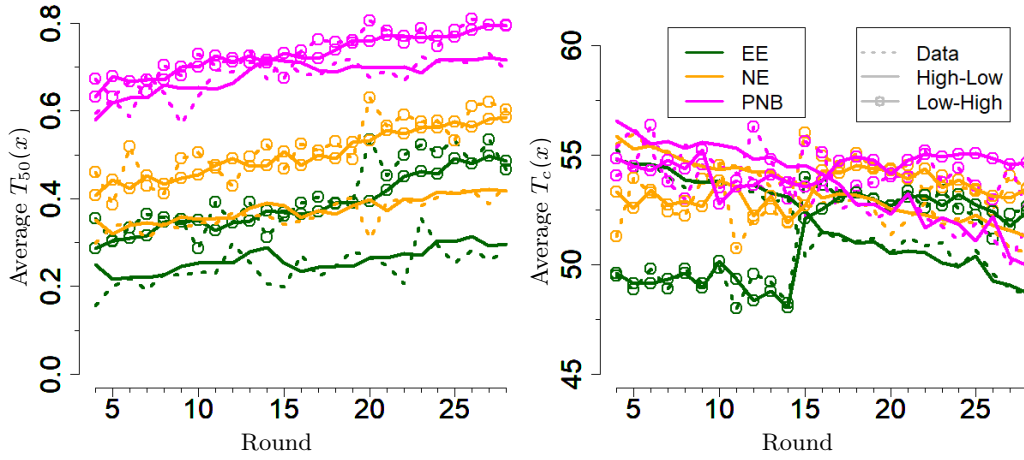
### 3.3 Analysis of the effect of normative variables

The difference between the reactive and normative agents comes from the effect of the normative variables: empirical expectation ( $ee$ ), normative expectation ( $ne$ ) and personal normative belief ( $pnb$ ).

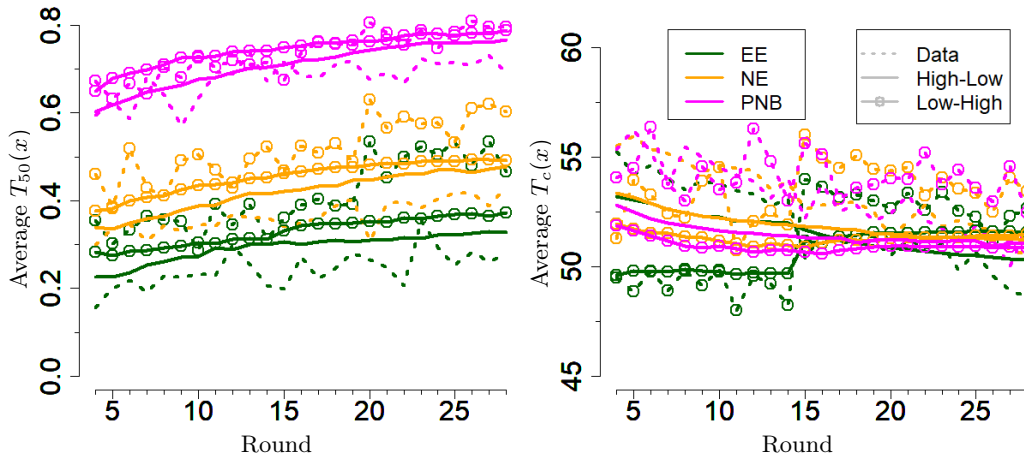
Figure 5 shows the results of the decision models predicting the normative variables when taking the data as inputs. Like previously, the fitting performance of frequency of contribution 50 or default cooperation (left panel) is rather satisfactory for the three variables. It is noticeable that all three variables are growing in both High-Low and Low-High sequences of disaster probabilities. The empirical expectation is surprisingly under-evaluating the actual frequency of default cooperation (shown on the left panel of Fig. 3) while the frequency of default cooperation given by personal normative belief is significantly higher than the actual frequency and may appear as an aim to reach.

The fitting performance of the fluctuating behaviour is also satisfactory (right panel). There is a significant difference between High-Low and Low-High sequences for the empirical expectation. Note that it tends to be lower than the actual average contributions (shown on the right panel of Fig. 3) for high disaster probability and higher for low disaster probability. Both normative expectation and personal normative belief tend to be higher than the actual average fluctuating behaviour.

Figure 6 shows the average frequency of default cooperation and fluctuating contributions for the normative variables averaged over 100 replicas of the game by agent based simulation. For all the variables, the results for the High-Low and Low-High sequences are closer than when tested on the data. Hence, for all the variables, frequency of default cooperation increases more over time in the sequence High-Low than when testing on the data. This explains the lower frequency of free-riding in the agent based simulations than on the data in this sequence because the free-riding is favoured by low values of the normative variables.



**Fig. 5.** Results on data for models of normative features: empirical expectation (*ee*), normative expectation (*ne*), personal normative belief (*pnb*). Left panel: contribution 50 from data and models over the rounds. Right panel average fluctuating contribution per round for data and agents (reactive and normative). High-Low and Low-High refer to the disaster probability sequence in the experiment.



**Fig. 6.** Results from simulations for models of normative features: empirical expectation (*ee*), normative expectation (*ne*), personal normative belief (*pnb*). Left panel: contribution 50 from data and models over the rounds. Right panel average fluctuating contribution per round for data and agents (reactive and normative). High-Low and Low-High refer to the disaster probability sequence in the experiment.

## 4 Discussion

Identifying three behaviours (free-riding, default cooperation and fluctuating contribution) in the game provides new keys for analysing the experimental data. Indeed, the decision models derived from the data suggest that free-riding and default cooperation tend to propagate by imitation. They also show that the normative variables (empirical expectation, normative expectation and personal normative belief) show an increasing frequency of default cooperation over time, in both High-Low and Low-High sequences of disaster probabilities.

Moreover, introducing these behaviours in agent based models and repeating the game many times, provides robust average results. Indeed virtually repeating the experiment many times erases some of its contingencies. Moreover, the simulations with reactive and normative agents show very significant differences, that do not appear when simply comparing the reactive and normative decision models on the data. In particular, the normative agents show a significantly lower frequency of free-riding than the reactive agents do for a high disaster probability in the Low-High sequence. The normative agents also show a higher frequency of default collaboration than the reactive agents do for both High-Low and Low-High sequences. Therefore the agent based simulations suggest that the experiment could have shown a much stronger difference between the High-Low and Low-High sequences of disaster probabilities.

Overall these results reinforce the interpretation proposed by Szekely et al. [10], that a norm favouring default cooperation and disfavouring free-riding is propagated during the game and that the normative variables play a key role in this propagation.

## 5 Acknowledgement

This work has been partly supported by the Agence Nationale de la Recherche through the FuturICT2.0 project (Flagera call) and the ToRealSim project (ORA call).

## References

1. K. Nyborg, J.M. Anderies, A. Denneberg, T. Lindah, C. Schill, M. Schutter, N. Adger, K.J. Arrow, S. Barrett, A. de Zeeuw, *Science* **354**, 42 (2016)
2. C. Bicchieri, *The Grammar of Society: The Nature and Dynamics of Social Norms*. (Cambridge University Press, 2006)
3. R. Cialdini, R. Reno, C. Kaligren, *Journal of Personality and Social Psychology*, **58**(6), 1015 (1990)
4. J. Coleman, *Foundations of Social Theory* (Belknap Press of Harvard University Press, 1990)
5. E. Ostrom, *Journal of Economic Behavior and Organization* **61**(2), 149 (2006)
6. D. Bell, M. Cox, *Journal of Family Theory and Review* **7**(1), 28 (2015)
7. E. Krupka, R. Weber, *Journal of the European Economic Association* **11**(3), 495 (2013)
8. E. Fehr, I. Shurtengerger, *Nature Human Behaviour* **2**, 458 (2018)
9. E. Kimbrough, A. Vostroknutov, *Economics Letters* **168**(C), 147 (2018)
10. A. Szekely, F. Lipari, A. Antonioni, M. Paolucci, L. Tummolini, G. Andighetto, *Nature Communications* **12**(5452) (2021)
11. M. Milinski, R. Sommerfeld, H.J. Krambeck, F. Reed, J. Marotzke, *Proceedings of National Academy of Sciences USA* **105**(7), 2291 (2008)
12. C. Bicchieri, J.W. Lindemans, T. Jiang, *Frontiers en Psychology* **5**(1418) (2014)