Threshold Public Good Games with Temporal Contribution Dynamics: Theory, Experiment, and Simulations*

Till Wicker†

October 3, 2025

*I am grateful to Eline van der Heijden, Daan van Soest, and Patricio Dalton for their useful suggestions on previous versions of the paper, and participants at the 16th RGS Conference, GSS Seminars, and TIBER 2023 Symposium for comments. I would also like to thank Maria Abou Chakra for her support. Funding is gratefully acknowledged from Tilburg University. IRB approval was obtained from Tilburg University (IRB EXE 2022-018), and the experiment was pre-registered *here*.

†Department of Economics, Tilburg University. Email: t.n.wicker@tilburguniversity.edu.

Abstract

Climate scientists are in agreement that the timing of greenhouse gas emissions plays a crucial role in climate change mitigation, with early reductions being more effective than equally sized emissions reductions in the future. This paper incorporates this scientific fact within Climate Protection Games, a multi-round collective risk public goods game that simulates countries' greenhouse gas emissions reductions over time, by introducing temporal dynamics that emphasize the importance of early action at mitigating climate change. A laboratory experiment with 300 participants compares two treatments: the status quo where emissions reductions are equally weighted across time (*Linear*), and another where earlier emissions reductions have a greater impact than equally sized later emissions reductions, reflecting the greater environmental benefits of early action (*Step*). Results indicate that while the likelihood of reaching the necessary threshold to avert disastrous climate change does not differ significantly between treatments, the *Step* treatment leads to higher early emissions reductions and increased individual payoffs, suggesting welfare improvements. Evolutionary Game Theory simulations corroborate these findings, showing that early emissions reductions increase individual payoffs. By highlighting the significance of early emissions reductions, this paper can contribute to the discourse on climate policy, suggesting that strategies incentivizing early emissions reductions can result in better environmental and welfare outcomes.

Keywords: Climate change mitigation, temporal dynamics, climate protection game, evolutionary game theory.

JEL codes: Q54, Q92, H41, C73.

"It's now or never, if we want to limit global warming to 1.5°C. Without immediate and deep emissions reductions across all sectors, it will be impossible." - Intergovernmental Panel on Climate Change Working Group III Co-Chair Jim Skea (2022).

1 Introduction

The provision of many public goods is not the result of contributions by individual actors at one point in time, but instead over several time periods. Important examples include donation and crowdfunding campaigns as well as climate change mitigation. A common element to these examples is that contributions at different time intervals differentially affect the likelihood of the level of the public good being provided. For example, on *Giving Tuesday* —traditionally the Tuesday after Thanksgiving where people are encouraged to give back in whatever way they can— NGOs and crowdfunding platforms offer matching pledges/donations, effectively increasing the marginal benefit of a donation. Relatedly, to effectively mitigate climate change by reducing greenhouse gas emissions, not only is the magnitude of emissions reductions important, but also the timing: earlier emissions reductions are more effective at mitigating climate change than equally sized reductions in the future, as the environmental damages of greenhouse gases are based on cumulative emissions (Shukla et al., 2022).

Little is understood about how the differing marginal returns to public goods contributions affect the intertemporal provision of public goods. On the one hand, a higher marginal benefit can induce greater contributions. However, this can come at the expense of future contributions (e.g., a dip in donations in the weeks after *Giving Tuesday*). Alternatively, if individuals want to contribute a fixed amount to the public good (akin to the daily income targeting of taxi drivers in Camerer et al. (1997)), they may wait for time periods when contributing results in a lower marginal cost (e.g., *Giving Tuesday*). While this benefits the individual (by reducing the marginal cost of their contribution), it does not affect the aggregate contribution to the public good.

In this paper, I experimentally investigate the impact of temporal contribution dynamics on the (voluntary) provision of a threshold public good. For this, I compare two cases: one in which the value of contributions to the public good does not depend on the timing on the contribution, and one in which it does. The value of contributions in the second case is a mean-preserving spread, with unchanged symmetric Nash Equilibria. I explore the impact of whether the timing of contributions matters or not on the timing of contributions, aggregate contributions to the public good, likelihood of reaching the threshold to ensure provision of the public good, and free-riding behavior. To guide the analysis, I develop a theoretical model that incorporates intertemporal contribution dynamics with the multiple rounds public goods game framework of Marx and Matthews (2000). In line with theoretical predictions, individual contributions to the public good are larger in rounds when the marginal cost is lower, and vice versa. Similarly, the likelihood of an individual free-riding is lower when the cost of doing so —in the form of foregone contributions to the public good— is higher. At a group level, introducing temporal contribution dynamics alters the contribution patterns, compared to the case where the timing of contributions does not matter. This has a marginally insignificant effect on the likelihood of reaching the threshold to ensure the public good is provided. Nevertheless, introducing temporal contribution dynamics improves individual's welfare, as measured by their experimental payments. Evolutionary Game Theory simulations that model how strategies evolve over generations in strategic, repeated interactions reproduce these results, thus providing further support for the theoretical model's predictions.

1.1 Contribution to the Literature

This paper contributes to the literature on multi-round public goods games, with a particular focus on threshold public goods games (see Croson and Marks (2000) for an overview). The laboratory experiment is an adaptation of the Climate Protection Game (Milinski et al., 2008), which has been highly cited and extended to include wealth inequality (Milinski et al., 2011; Burton-Chellew et al., 2013; Brown and Kroll, 2017; Tavoni et al., 2011), threshold uncertainty (Barrett and Dannenberg, 2013; Dannenberg et al., 2015; Brown and Kroll, 2017), and differing degrees of vulnerability to climate change damages (Burton-Chellew et al., 2013). However, no paper has explored the effect of temporal contribution dynamics in the CPG, despite the scientific evidence that earlier emissions reductions are more effective at mitigating climate change than future emissions reductions (IPCC, 2022), and the wide-ranging economic, environmental, and health-related benefits of earlier emissions reductions (Hamilton et al., 2017; Bosetti et al., 2012; Goulder, 2020). This paper addresses this gap in the literature, by comparing contribution dynamics across two treatments: one with intertemporal contribution dynamics, and one without. As such, this paper also contributes to other experimental games simulating climae change mitigation, including the Intergenerational Goods Game (Hauser et al., 2014), Climate Negotations (Barrett and Dannenberg, 2012), and the Persistent Carbon climate game (Calzolari et al., 2018).

The second literature this paper contributes to is on the theoretical representations and evolutionary game theoretic simulations of multi-round public goods games. The theoretical model builds on Marx and Matthews (2000) by incorporating temporal dynamics through a round-specific weighting function. Several studies have conducted evolutionary game theoretic simulations of the threshold public goods games (Abou Chakra and Traulsen, 2012; Hilbe et al., 2013; Santos et al., 2012). These simulations have been extended to evaluate how strategies respond to changes in uncertainty (Abou Chakra et al., 2018), risk and group sizes (Santos and Pacheco, 2011), wealth inequality (Vasconcelos et al., 2014; Abou Chakra and Traulsen, 2014), and changes in governance institutions (Vasconcelos et al., 2013). The EGT simulations in this paper extend the literature by evaluating the role of temporal dynamics in the timing of contributions and the evolution of strategies.

UPDATE: The structure of this paper is as follows. Section 3 describes the experimental design, while Section 4 presents the results, including those from the EGT simulations. Section 5 concludes. The Appendix contains a theoretical model that incorporates temporal dynamics within Marx and Matthews (2000), details on the evolutionary game theory simulations, and additional results.

2 Experimental Design and Theoretical Predictions

My experimental design allows me to study the role of intertemporal contribution dynamics on the provision of public goods in the presence of a threshold. Specifically, in each of the two treatment arms, five players are randomly paired to form a group. Across nine rounds, players can contribute part of their initial endowment to a public good. If after the nine rounds, aggregate contributions to the public good exceed a prespecified and communicated threshold, the public good was provided. If the threshold was not met, the public good was not provided.

- intro paragraph
    - many problems in ...
    - important examples

- a common element

- most studies don't take this into consideration, despite studies on the VCM showing that contributions vary based on payoffs
- In this paper, I experimentally...
- to guide our investigation...
- our experimental results. include EGT
- subsection: contribution to literature
- This paper proceeds as follows
- Experimental Design and Theoretical Predictions
  - our experimental design
  - payoff function of individuals
  - combine theory with description of experimental design
  - include predictions from theoretical model
  - leave most of the theoretical model in the appendix
- experimental procedure
  - Table 1, equivalent to the other paper
- experimental results
- EGT Simulations
  - Result 1, result 2, etc.
- Conclusion
  - generic
  - what i do
  - results, combined with theory
  - overall conclusion

Climate scientists are in agreement that in order to mitigate the risks of climate change, countries need to reduce greenhouse gas emissions. The latest Intergovernmental Panel on Climate Change (IPCC) report calculates that the carbon budget - reflecting the amount of CO_2 that can still be emitted while limiting the temperature rise to 1.5°C with 50% probability - is 500 Gt CO_2, before reaching net zero CO_2 emissions to ultimately halt global warming (Shukla et al., 2022).[1] The IPCC report also highlights that "more rapid near-term emissions reductions allow reaching net zero CO_2 at a later point in time". This underscores that earlier emissions reductions are more effective at mitigating climate change than equally sized reductions in the future, as the environmental damages of greenhouse gases are based on cumulative emissions (Shukla et al., 2022). Other papers have also emphasized the wider-reaching economic, environmental, and health-related benefits of earlier emissions reductions (Hamilton et al., 2017; Bosetti et al., 2012; Goulder, 2020).

Laboratory experiments, such as the Climate Protection Game (CPG, Milinski et al. (2008)), have simulated countries' reductions in CO_2 emissions over time as a collective risk threshold public goods game. While these experiments capture the role of the carbon budget in CO_2 emissions reductions by introducing a threshold that triggers disastrous climate change, they do not take into consideration *when* reductions in CO_2 emissions are made - despite the importance of timing. I address this limitation by introducing temporal dynamics within the Climate Protection Game (CPG) that more accurately reflect the role of the timing of emissions reductions in climate change mitigation.

I conduct a lab experiment with 300 university students, who are pooled into groups of five. Groups are randomized across two treatment arms, *Linear* and *Step*. The *Linear* treatment arm replicates the CPG of Milinski et al. (2008), where contributions to the public good (equivalent to reducing CO_2 emissions in the real world) are weighted equally across all rounds of the game. Extending this set-up to climate change mitigation implies that reductions in CO_2 emissions today are equally effective at mitigating disastrous climate change as equally-sized emissions reductions in the future, in contrast to scientific reports (Shukla et al., 2022). The *Step* treatment arm incorporates the intertemporal effects of emissions reductions, with contributions in earlier rounds yielding larger benefits than contributions in later rounds, reflecting the environmental

[1]The primary focus is placed on CO_2 as it represents the largest share of GHG emissions, has a long half-life, and plays a significant role in the trapping of heat in the Earth's atmosphere (Solomon et al., 2009).

benefits of earlier emissions reductions. This is done through a mean-preserving and decaying weighting factor that ensures symmetric pure Nash Equilibria are unchanged across the two treatment arms.

Results of the lab experiment show that while the likelihood of reaching the threshold required to avert disastrous climate change does not differ across the two treatments, the contribution patterns differ substantially. Compared with groups in the *Linear* arm, groups in the *Step* arm contribute more in earlier rounds of the Climate Protection Game, and subsequently reduce their contributions in later rounds. Payoffs are statistically significantly higher for groups in the *Step* arm, suggesting welfare improvements as a result of emphasizing earlier contributions to the public good.

The findings from the lab experiment are rationalized through a theoretical model that incorporates temporal dynamics within the model of Marx and Matthews (2000), and are replicated through Evolutionary Game Theoretic (EGT) simulations of the lab experiment that model how strategies evolve over generations in strategic, repeated interactions. Players in the *Step* EGT simulation contribute more in earlier rounds than players in the *Linear* EGT simulation, before reducing contributions afterwards. The likelihood of reaching the threshold is not statistically significantly different between the two simulated treatments, but payoffs are substantially higher in the *Step* EGT simulation - in line with the results from the lab experiment.

This paper contributes to the literature on threshold public goods games (see Croson and Marks (2000) for an overview), particularly those that simulate climate change mitigation including the Climate Protection Game developed by Milinski et al. (2008).[2] The CPG has been highly cited and extended to include wealth inequality (Milinski et al., 2011; Burton-Chellew et al., 2013; Brown and Kroll, 2017; Tavoni et al., 2011), threshold uncertainty (Barrett and Dannenberg, 2013; Dannenberg et al., 2015; Brown and Kroll, 2017), and differing degrees of vulnerability to climate change damages (Burton-Chellew et al., 2013). However, no paper has explored the effect of temporal contribution dynamics in the CPG, despite the scientific evidence that earlier emissions reductions are more effective at mitigating climate change than future emissions reductions (IPCC, 2022), and the wide-ranging economic, environmental, and health-related benefits of earlier emissions reductions (Hamilton et al., 2017; Bosetti et al., 2012; Goulder, 2020). This paper addresses this gap in the literature, by comparing

[2]Other examples include the Intergenerational Goods Game (Hauser et al., 2014), Climate Negotations (Barrett and Dannenberg, 2012), and the Persistent Carbon climate game (Calzolari et al., 2018).

the *Linear* and *Step* treatment arms. While coordinated early contributions in the *Step* treatment arm can improve efficiency and increase the likelihood of reaching the threshold, "last-ditch efforts" to mitigate climate change are less likely to be successful.

The second literature this paper contributes to is on the theoretical representations and evolutionary game theoretic simulations of multi-round public goods games. The theoretical model builds on Marx and Matthews (2000) by incorporating temporal dynamics through a round-specific weighting function $\beta(t)$. Several studies have conducted evolutionary game theoretic simulations of the Climate Protection Game (Abou Chakra and Traulsen, 2012; Hilbe et al., 2013; Santos et al., 2012). These simulations have been extended to evaluate how strategies respond to changes in uncertainty (Abou Chakra et al., 2018), risk and group sizes (Santos and Pacheco, 2011), wealth inequality (Vasconcelos et al., 2014; Abou Chakra and Traulsen, 2014), and changes in governance institutions (Vasconcelos et al., 2013). The EGT simulations in this paper extend the literature by evaluating the role of temporal dynamics in the timing of contributions and the evolution of strategies.

3 Experimental Design

I conduct a laboratory experiment of a multi-round threshold public goods game with and without temporal contribution dynamics. The experiment is based on the Climate Protection Game (Milinski et al., 2008), which simulates coordination between countries over time to reduce CO_2 emissions to prevent catastrophic climate change.

Across two treatments, there are five players per group (partner matching), each endowed with 36 Monetary Units (MU, MU1 = €0.50). Players can simultaneously contribute up to MU4 from their own endowment to the public good per round, and the game lasts 9 rounds.[3] Contributions made to the public good (called the *Climate Pot*) in round t are multiplied by $\beta(t)$ before being added to the *Climate Pot*.[4] After each round, players are shown the contributions of all group members, the total group contribution to the *Climate Pot*, and their remaining savings (equaling their endowment minus their total contributions).

The threshold group contribution at which disastrous climate change is averted is

[3] Contrary to Milinski et al. (2008), players can contribute MU$\{0, 1, 2, 3, 4\}$, and not only MU$\{0, 2, 4\}$.

[4] For example, if a player contributes MU4 in round t, it will count as MU$\big(4 \cdot \beta(t)\big)$ towards the *Climate Pot*.

MU90. If a group reaches or exceeds MU90, players keep their remaining savings.[5] If the group fails to reach the threshold, players face a 70% probability of losing their savings. In the remaining 30% of cases, players keep their savings.

Treatment 1: Linear

This treatment resembles the Milinski et al. (2008) CPG setup. Contributions to the *Climate Pot* are *Linear*, meaning they are not scaled ($\beta(t) = 1 \quad \forall t \in \{1, ..., 9\}$).

Treatment 2: Step

This treatment differs from the *Linear* design, with $\beta(t)$ decreasing over time:

$$\beta(t) = \begin{cases} 1.5 & t = \{1, 2, 3\} \\ 1.0 & t = \{4, 5, 6\} \\ 0.5 & t = \{7, 8, 9\} \end{cases}$$

Except for the decaying *Step* $\beta(t)$, the setup is identical to *Linear*. The decreasing $\beta(t)$ means that the returns to contributions decay: contributions in earlier rounds count for more in the provision of the public good, and reaching the threshold, than contributions in later rounds. This more accurately captures conclusions from scientific reports on climate change mitigation that document the greater impact of immediate CO_2 emissions reductions at mitigating climate change than equally sized future emissions reductions (Warren et al., 2013; Ciavarella et al., 2017; IPCC, 2022).

Nash Equilibria

There are two, Pareto-ranked pure symmetric Nash Equilibria for risk-neutral actors in both treatments: (i) players contribute 0 to the public good in all rounds, and fail to reach the threshold, or (ii) players contribute 2 to the public good in all rounds, and reach the threshold. The expected payoffs for the Nash Equilibria are MU10.8 and MU18, respectively. Additionally, there are infinitely many asymmetric equilibria, as discussed in Appendix A.3.1.

[5] Milinski et al. (2008) used 6 players over 10 rounds, with a threshold of 120 and endowment of 40. I deviate from this setup for budgetary reasons. However, for both my setup, and that of Milinski et al. (2008), in order to reach the threshold, the average contribution by each player is half their endowment.

Conditional on reaching the threshold, the most efficient contribution pattern for the *Step* treatment consists of each player contributing 4MU in each of the first 3 rounds, and 0MU otherwise.[6] For the *Linear* treatment the timing of contributions does not matter.

Experimental Procedure

The experiment was conducted in June 2022 and May 2023, at Tilburg University's CentERlab, and programmed using oTree (Chen et al., 2016). 300 participants were divided into 60 groups, evenly split across BOTH treatments. Participants could not identify or communicate with each other. 50% of the participants identified as female, and their average age was 21.89. The experiment lasted 30 minutes, and the average compensation was €11.99.

4 Results

Groups in the *Linear* treatment had a mean aggregate contribution of 80.80 $\pm$ 3.24 (mean $\pm$ SE), with 53% of the groups reaching the threshold of 90. In the *Step* treatment, the mean aggregate contribution was 89.87 $\pm$ 0.98, and 63% of groups reached the threshold (see Figure 1).[7] While the difference in the likelihood of reaching the threshold (p=0.26, n_1=30, n_2=30, Binomial Test) is not statistically significantly different, the difference in aggregate contributions (p=0.11, n_1=30, n_2=30, two-sided Mann-Whitney U) is marginally insignificant, offering suggestive evidence that the *Step* treatment increased the likelihood of groups reaching the threshold.

Despite statistically insignificantly different aggregate contributions, the contribution paths differ substantially between groups in the *Linear* and *Step* treatments (n_1 = 270, n_2 = 270, p<0.01, K-S test).[8] Figure 2a depicts the average cumulative contributions across the nine rounds for the *Linear* and *Step* treatments. The contribution path in the *Linear* treatment – in line with Milinski et al. (2008) and other papers – increases linearly. The cumulative contribution path of the *Step* treatment, on the

[6]This results in savings of 24MU per player.

[7]These findings are slightly higher than those in Milinski et al. (2008), who find that only 50% and 10% of groups reach the threshold when the probability of losing their savings was 90% and 50%, respectively. An explanation is the smaller group size (5 vs. 6 members per group) in my study.

[8]$n_1 = n_2 = 270$: 30 groups per treatment, nine rounds per group.

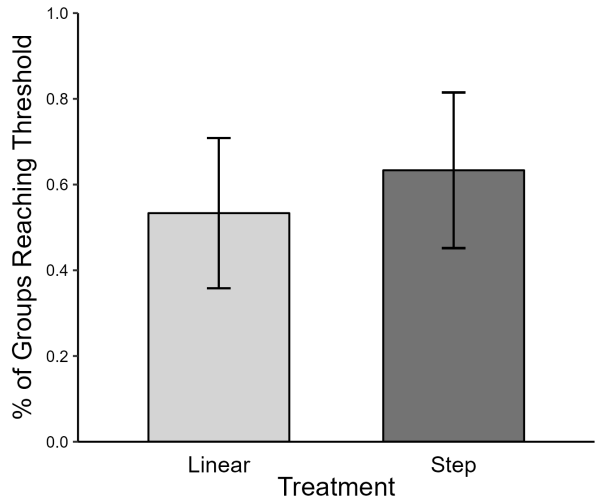


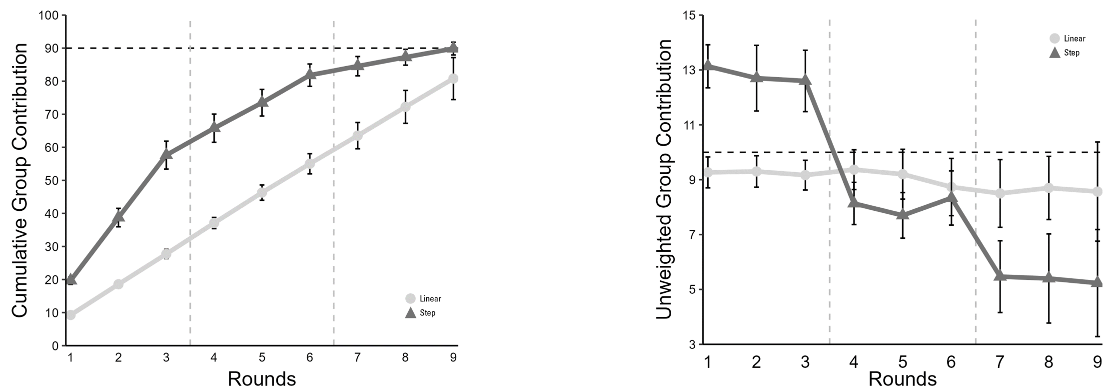
Figure 1. Likelihood of Reaching Threshold, by Treatment

other hand, is concave – characterized by high contributions in earlier rounds that subsequently decay.

The difference in cumulative contribution paths between *Linear* and *Step* can be due to two competing explanations. The first explanation is that groups in both treatments contributed similar amounts to the *Climate Pot* across all rounds, however these contributions were weighted by $\beta = \{1.5, 1.0, 0.5\}$ in the *Step* treatment. The decaying weights will then naturally lead to the observed concave cumulative contribution path. The competing explanation is that groups in the *Step* treatment contributed more than groups in the *Linear* treatment in rounds where the weighting factor was greater than 1, and less in rounds where the weighting factor was less than 1.

Figure 2b plots the average unweighted group contribution per round for *Linear* and *Step* treatments.[9] In line with the second explanation, groups in the *Step* treatment contribute statistically significantly more in the first three rounds, when contributions are weighted by 1.5 (38.43 ± 1.44 vs. 27.73 ± 0.71; p<0.01, n_1=30, n_2=30, two-sided MWU), and statistically significantly less in the last three rounds, where contributions are weighted by 0.5 (16.10 ± 2.33 vs. 25.77 ± 1.95; p<0.01, n_1=30, n_2=30, two-sided MWU).

[9] Unweighting entails dividing the individual's contribution by $\beta(t)$. This involves no transformation for *Linear*, while for *Step*, contributions are divided by $\{1.5, 1.0, 0.5\}$, respectively.



(a) Cumulative Contribution Paths

(b) Unweighted Group Contributions

Figure 2. Group Contributions by Treatment

Further support comes from looking at individual player contributions in the first round. While contributions in rounds 2-9 can be influenced by other group members' actions, contributions in the first round are more likely to reflect players' preferences.[10] Compared to the *Linear* treatment, individuals in the *Step* treatment were less likely to contribute nothing to the public good (0.02 ± 0.01 vs. 0.09 ± 0.02; p=0.01, n_1=150, n_2=150, two-sided t-test), and their unweighted contributions were higher (2.63 ± 0.09 vs. 1.85 ± 0.06; p<0.01, n_1=150, n_2=150, two-sided t-test).[11]

While both treatments had identical pure symmetric Nash Equilibria, the resulting differing contribution patterns led to different payoffs: The average payoff for groups in the *Step* treatment was MU14.91 $\pm$ 1.48, compared with MU13.07 $\pm$ 1.08 for the *Linear* treatment (p=0.02, n_1=30, n_2=30, two-sided MWU).[12] As such, rewarding earlier contributions to the *Climate Pot* resulted in a welfare-improving reallocation of contributions, without altering the likelihood of reaching the threshold.

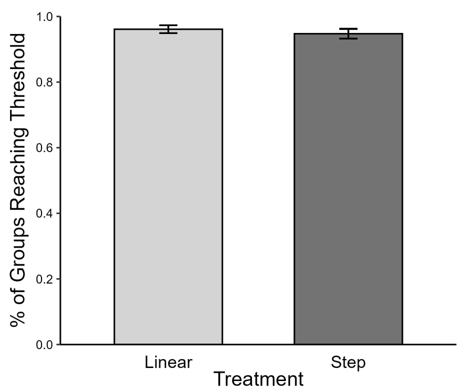
[10] An alternative explanation could be confusion. However, there is no correlation between first-round contributions and performance on the comprehension test ($\rho = -0.016, p = 0.79$).

[11] Conditional on cooperating, first round unweighted contributions are statistically significantly larger for *Step* than *Linear* (2.68 ± 0.08 vs. 2.03 ± 0.05; p<0.01, n_1=150, n_2=150, two-sided t-test).

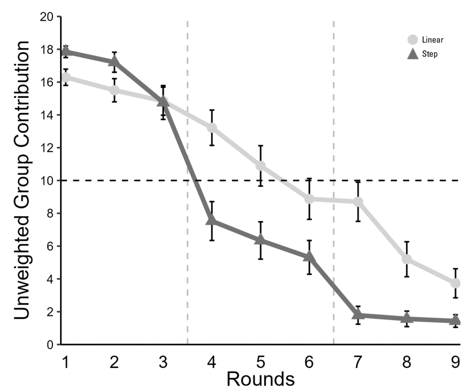
[12] The expected payoff - equal to ones remaining savings in case the threshold was met, and equal to 70% of the remaining savings otherwise, was also statistically significantly different across *Step* and *Linear* groups: MU18.08 $\pm$ 0.66 vs. MU16.71 $\pm$ 0.39, p=0.06, n_1=30, n_2=30, two-sided MWU).

4.1 Evolutionary Game Theory Simulations

I use Evolutionary Game Theory (EGT) to study the role of temporal dynamics within Climate Protection Games in the evolution of strategies (Weibull, 1997; Hofbauer and Sigmund, 1998; Hilbe et al., 2013). For both treatments, I run 150 iterations of 100,000 generations of the game. Per generation, each player participates in up to 1000 Climate Protection Games, randomly paired with other players that follow different strategies. After each generation, a player's average payoffs are calculated, with strategies that result in higher payoffs having a higher probability of being passed on to the next generation following a Wright-Fisher process. Appendix B outlines the parameters of the EGT simulations in more detail.



(a) Likelihood of Reaching Threshold



(b) Unweighted Group Contributions

Figure 3. Evolutionary Game Theory Simulations

Figure 3 reproduces Tables 1 and 2b for the EGT simulations. As Figure 3a illustrates, in the EGT simulations the groups are far more likely to reach the threshold than in the lab experiment. The likelihood of reaching the threshold is statistically insignificantly different across the two treatments (0.96 ± 0.01 vs. 0.95 ± 0.01; p=0.16, n_1=150, n_2=150, two-sided t-test). Figure 3b presents the unweighted group contributions averaged across all iterations and generations, for the EGT simulations of both the *Linear* and *Step* treatments. The pattern of the *Step* treatment is remarkably similar to that of the lab participants, characterized by high initial contributions that decay after round 3. While the pattern of contributions in the *Linear* simulation does not reflect those among lab participants, the statistically significant differences in contributions across rounds between the EGT simulations of *Linear* and *Step* support the findings of the lab experiment.

5 Conclusion

The Climate Protection Game has been used to simulate the ability of countries to coordinate to reduce greenhouse gas emissions over time in order to mitigate disastrous climate change. Following the initial paper by Milinski et al. (2008), studies have extended the CPG to measure the effect of differing initial conditions of actors and proposed policies. However, these studies overlooked a significant fact of the science of climate change mitigation: earlier CO_2 emissions reductions are more effective at mitigating climate change than equally-sized reductions in the future because greenhouse gas emissions are a stock pollutant (Shukla et al., 2022).

I address this limitation by introducing temporal dynamics within the CPG. Compared with the traditional CPG setup, actors in the *Step* treatment arm contribute more in earlier rounds when they get a bigger bang for their buck. These earlier contributions decay in later rounds, and do not affect the level of overall contributions. While the likelihood of reaching the public good threshold required to avert disastrous climate change is unchanged, payoffs - a proxy for welfare - are higher.

Given the importance of early contributions in climate change mitigation, an interesting avenue for further research is to better understand how to foster early stage coordination and contributions. It is unclear whether the insights gained from studies building on the Climate Protection Game by Milinski et al. (2008) can directly translate to a dynamic setting with temporal dynamics that more accurately reflects the scientific knowledge underlying climate change mitigation. Furthermore, dynamic CPGs can be extended to evaluate further climate change mitigation policies, such as the rachet-up mechanism (Alt et al., 2023), or coalitions (Bosetti et al., 2017).

References

Abou Chakra, M., Bumann, S., Schenk, H., and Milinski, M. (2018). Immediate action is the best strategy when facing uncertain climate change. *Nature Communications*, 9:2566.

Abou Chakra, M. and Traulsen, A. (2012). Evolutionary dynamics of strategic behavior in a collective-risk dilemma. *PLOS Computational Biology*, 8(8):1–7.

Abou Chakra, M. and Traulsen, A. (2014). Under high stakes and uncertainty the rich should lend the poor a helping hand. *Journal of Theoretical Biology*, 341:123–130.

Alt, M., Gallier, C., Kesternich, M., and Sturm, B. (2023). Collective minimum contributions to counteract the ratchet effect in the voluntary provision of public goods. *Journal of Environmental Economics and Management*, 122:102895.

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

Barrett, S. and Dannenberg, A. (2013). Sensitivity of collective action to uncertainty about climate tipping points. *Nature Climate Change*, 4:36–39.

Bosetti, V., Carraro, C., and Tavoni, M. (2012). Timing of mitigation and technology availability in achieving a low-carbon world. *Environmental and Resource Economics*, 51(3):353–369.

Bosetti, V., Heugues, M., and Tavoni, A. (2017). Luring others into climate action: coalition formation games with threshold and spillover effects. *Oxford Economic Papers*, 69(2):410–431.

Brown, T. and Kroll, S. (2017). Avoiding an uncertain catastrophe: climate change mitigation under risk and wealth heterogeneity. *Climatic Change*, 141(2):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(4):815–830.

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

Camerer, C., Babcock, L., Loewenstein, G., and Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time*. *The Quarterly Journal of Economics*, 112(2):407–441.

Chen, D., Schonger, M., and Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Ciavarella, A., Stott, P., and Lowe, J. (2017). Early benefits of mitigation in risk of regional climate extremes. *Nature Climate Change*, 7:326–330.

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

Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., and Tavoni, A. (2015). On the Provision of Public Goods with Probabilistic and Ambiguous Thresholds. *Environmental & Resource Economics*, 61(3):365–383.

Goulder, L. (2020). Timing Is Everything: How Economists Can Better Address the Urgency of Stronger Climate Policy. *Review of Environmental Economics and Policy*, 14(1):143–156.

Hamilton, K., Brahmbhatt, M., and Liu, J. (2017). Multiple benefits from climate change mitigation: Assessing the evidence. Technical report, Grantham Research Institute on Climate Change and the Environment.

Hauser, O., Rand, D. G., Peysakhovich, A., and Nowak, M. A. (2014). Cooperating with the future. *Nature*, 511(7508):220–223.

Hilbe, C., Abou Chakra, M., Altrock, P., and Traulsen, A. (2013). The evolution of strategic timing in collective-risk dilemmas. *PLOS ONE*, 8(6):1–7.

Hofbauer, J. and Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge University Press.

IPCC (2022). Climate change 2022: Impacts, adaptation, and vulnerability. contribution of working group ii to the sixth assessment report of the intergovernmental panel on climate change. Report, Intergovernmental Panel on Climate Change.

Marx, L. and Matthews, S. (2000). Dynamic voluntary contribution to a public project. *Review of Economic Studies*, 67(2):327–358.

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., Sommerfeld, R., Krambeck, H., Reed, F., and Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, 105(7):2291–2294.

Santos, F. C. and Pacheco, J. M. (2011). Risk of collective failure provides an escape from the tragedy of the commons. *Proceedings of the National Academy of Sciences*, 108(26):10421–10425.

Santos, F. C., Vasconcelos, V. V., Santos, M. D., Neves, P. N. B., and Pacheco, J. M. (2012). Evolutionary dynamics of climate change under collective-risk dilemmas. *Mathematical Models and Methods in Applied Sciences*, 22(supp01):1140004.

Shukla, P. R., Skea, J., Slade, R., Al Khourdajie, A., van Diemen, R. P., McCollum, D., Pathak, M., Some, S., Vyas, P., Fradera, R., Belkacemi, M., Hasija, A., Lisboa, G., Luz, S., and Malley, J., editors (2022). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, UK and New York, NY, USA.

Solomon, S., Plattner, G.-K., Knutti, R., and Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences*, 106(6):1704–1709.

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*, 108(29):11825–11829.

Vasconcelos, V., Santos, F., and Pacheco, J. (2013). A bottom-up institutional approach to cooperative governance of risky commons. *Nature Climate Change*, 3:797–801.

Vasconcelos, V. V., Santos, F. C., Pacheco, J. M., and Levin, S. A. (2014). Climate policies under wealth inequality. *Proceedings of the National Academy of Sciences*, 111(6):2212–2216.

Warren, R., Lowe, J., Arnell, N., Hope, C., Berry, P., Brown, S., Gambhir, A., Gosling, S., Nicholls, R., O'Hanley, J., Osborn, T., Osborne, T., Price, J., Raper, S., Rose, G., and Vanderwal, J. (2013). The avoid programme's new simulations of the global benefits of stringent climate change mitigation. *Climatic Change*, 120:55–70.

Weibull, J. (1997). *Evolutionary Game Theory.* MIT Press. MIT Press.

Online Appendix:

A Theoretical Framework: Threshold PGGs with Weights

The game has a set of players $N = \{1, ..., n\}$, with $n \geq 2$. Each player is endowed with private good E_i, which can be used to contribute $z_i(t)$ to a public good in each round $t \geq 1$. The contribution horizon is $[1, \bar{T}]$, where $\bar{T} \leq \infty$. Contributions $z_i(t)$ are non-refundable and multiplied by a round-specific weight $\beta(t) > 0$, capturing the temporal dynamics of contributions.[13]

The round t contribution vector is defined as $z(t) = \big(z_1(t), ..., z_n(t)\big)$, with the entire contribution sequence being $\{z\} = \{z(t)\}_{t=0}^{\bar{T}}$. Individuals observe their own past contributions and aggregate of the others' contributions, let $Z(t) = \sum_{j \in N} z_j(t)$, and $Z_{-i}(t) = Z(t) - z_i(t)$. Player i's personal history at the start of round t then is: $h_i^{t-1} = \big(z_i(\tau), Z_{-i}(\tau)\big)_{\tau=0}^{t-1}$.[14]

Individual payoffs are determined by individual and aggregate contributions at the end of the contribution horizon, $\bar{T}$. Player i's individual unweighted cumulative contribution at the end of round t is $\hat{z}_i(t) = \sum_{\tau \leq t} z_i(\tau)$ and $\hat{z}_i(\bar{T}) = \hat{Z}_i$, while the weighted cumulative contribution is $\tilde{z}_i(t) = \sum_{\tau \leq t} \beta(\tau) z_i(\tau)$ and $\tilde{z}_i(\bar{T}) = \tilde{Z}_i$. The aggregate group unweighted and weighted cumulative contribution, henceforth referred to as the *unweighted* and *weighted cumulation*, are $\hat{Z}(t) = \sum_{j \in N} \hat{z}_j(t) = \sum_{t=0}^{\bar{T}} z(t)$ and $\tilde{Z}(t) = \sum_{j \in N} \tilde{z}_j(t) = \sum_{t=0}^{\bar{T}} \beta(t) z(t)$, respectively.

The benefit Player i receives after $\bar{T}$ rounds is $f_i\big(\tilde{Z}(\bar{T})\big)$, which depends on the weighted group cumulation at the end of the contribution horizon. Contribution costs enter the utility function quasilinearly: $U_i(\{z\}) = f_i\big(\tilde{Z}(\bar{T})\big) - \hat{z}_i(\bar{T})$.

$f_i(\tilde{Z})$ is a *step benefit function* where $f_i(\tilde{Z}) = Y_i = g(E_i, \hat{z}_i, p)$ until $\tilde{Z}$ exceeds a threshold value Z^*, beyond which individual benefits jump to $f_i(\tilde{Z}) = V_i = h(E_i, \hat{z}_i)$. Contributions $Z_i(t)$ are non-refundable.

$$f_i(\tilde{Z}) = \begin{cases} Y_i = g(E_i, \hat{z}_i, p) & \text{for } \tilde{Z} < Z^* \\ V_i = h(E_i, \hat{z}_i) & \text{for } \tilde{Z} \geq Z^* \end{cases}$$

[13] Note that setting $\beta(t) = 1 \quad \forall t$ results in no temporal dynamics, and simplifies to the setup of Marx and Matthews (2000).

[14] The framework can be extended to games where individual contributions are publicly observed, such that $h_i^{t-1} = \big(z_i(\tau), z_{j \in N, j \neq i}(\tau)\big)_{\tau=0}^{t-1}$, see Marx and Matthews (2000). Note it does not matter whether players observe weighted or unweighted contributions, as long as $\beta(\tau)$ is known to all players for all $\tau \leq t-1$.

Where $g'(\hat{z}_i) \leq 0; g'(E_i) \geq 0; h'(\hat{z}_i) \leq 0; h'(E_i) \geq 0;\ h(\hat{z}_i|E_i) > g(\hat{z}_i|E_i) > 0 \quad \forall \hat{z}_i < E_i$; and $h(\hat{z}_i) = g(\hat{z}_i) = 0$ for $\hat{z}_i = E_i$.[15] Hence the benefit jump from reaching the threshold is $V_i - Y_i \geq 0$.[16] Similar to Marx and Matthews (2000), I focus on cases where free-riding is an issue, meaning that no player has the incentive to complete the public goods project alone, but aggregate utility is greater if the project is completed: $V_i < \dfrac{Z^*}{\max_t\{\beta(t)\}} < \sum_{j=1}^{N} V_j \quad \forall i \in N, \quad \forall t$, equivalent to equation (2.7) in Marx and Matthews (2000).

A.1 One-Round Game

Dropping time notation, the n players each contribute z_i of their endowment E_i to the public good, yielding a contribution vector z, aggregate contributions $\tilde{Z}(= \beta \cdot z)$, and payoffs $f_i(\tilde{Z}) - z_i$.[17] If the other group members have contributed $\tilde{Z}_{-i} \geq Z^*$ and thus met the public good's threshold, player i's best response is to contribute 0.[18] However if the others' contributions are such that $\tilde{Z}_{-i} < Z^*$, player i's best response is to either contribute zero, or $z_i = \dfrac{1}{\beta}(Z^* - \tilde{Z}_{-i})$. Player i will contribute $z_i > 0$ to reach the threshold if the marginal benefit of doing so exceeds the marginal cost:

$$V_i(z_i) - Y_i(0) > \frac{1}{\beta}(Z^* - \tilde{Z}_{-i})$$
$$\beta \cdot \big(V_i(z_i) - Y_i(0)\big) > Z^* - \tilde{Z}_{-i}$$

Hence player i will only contribute $z_i > 0$ and ensure the provision of the public good if the contribution is less than the player's critical contribution, $c_i^* = \beta \cdot \big(V_i(z_i) - Y_i(0)\big) = \beta\big(h(z_i) - g(0)\big)$.[19]

Hence, for $\tilde{Z}_{-i} < Z^*$, player i's best response function is:

$$z_i(Z_{-i}) = \begin{cases} \dfrac{1}{\beta}(Z^* - \tilde{Z}_{-i}) & \text{if } Z^* - \tilde{Z}_{-i} \leq c_i^* \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

[15] Note, $f_i(\tilde{Z})$ can be extended such that Y_i and V_i both depend on cumulative weighted group contributions $\tilde{Z}$, so long as $h(\tilde{Z}) > g(\tilde{Z}) \quad \forall E_i, \hat{z}_i$ and hence $\lim_{\tilde{Z}\to Z^{*+}} h(\tilde{Z}) > \lim_{\tilde{Z}\to Z^{*-}} g(\tilde{Z})$

[16] For simplification, E_i will be dropped from notation. The endowment may be important when players have differing initial endowments (like in Milinski et al. (2011)), however this is not the case in this paper. Furthermore, p, the probability of keeping ones savings if the threshold is not reached, will also be dropped from notation. However, as Milinski et al. (2008) illustrates, the probability of keeping ones savings impacts individual's contributions and hence likelihood of reaching the threshold.

[17] Note that in a one-round game, $z_i = \hat{z}_i$ by definition.

[18] Where $\tilde{Z}_{-i} = \sum_{j\in N, j\neq i} \beta \cdot z_j$, the weighted aggregate of the others' contributions.

[19] Note that the critical contribution c_i^* refers to the weighted contributions $\tilde{z}_i = \beta \cdot z_i$.

Nash Equilibria

Given $V_i < \dfrac{Z^*}{\beta} < \sum_{j=1}^{N} V_j$ for all $i \in N$, one Nash Equilibrium is for every player to contribute 0, and the threshold is not met.

> **Proposition 1:** One Nash Equilibrium of the One-Round Game is for each player to contribute 0, $z = (0, ..., 0)$.

Proof: One kind of deviation must be considered: a player unilaterally decides to contribute $0 < z_{i0} \leq E_i$ such that $\beta \cdot z_{i0} = Z^*$. Their payoff is then $U_i(z_{i0}) = V_i - z_{i0}$. Given $V_i(z_{i0}) < \dfrac{Z^*}{\beta} \Rightarrow V_i(z_{i0}) < z_{i0}$, the individual's payoff would be negative, and hence they are better off not deviating from the Nash Equilibrium of contributing 0. Intermediate values of z_i ($0 < z_i < z_{i0}$) are inferior as $g(\hat{z}_i) < h(\hat{z}_i) \quad \forall \hat{z}_i < E_i$, and are hence not a Nash Equilibrium.

Additional Nash Equilibria with positive aggregate contributions exist as long as two conditions are met:

> **Proposition 2:** The One-Round Static Game has further Nash Equilibria of contributions $(z_1, ..., z_n)$ that satisfy: 1) $\beta \cdot \sum_{i=1}^{n} z_i = Z^*$, and 2) $0 \leq z_i \leq \dfrac{c_i^*}{\beta}$.

Proof: Four kinds of deviations need to be considered.

Firstly, consider a player i that contributes less than the Nash Equilibrium (z_i^{NE}), such that $\beta \cdot \sum_{i=1}^{n} z_i < Z^*$ and $0 \leq z_i \leq \dfrac{c_i^*}{\beta}$. The player contributes $z_{i1} < z_i^{NE}$, resulting in a benefit function $f_i(\tilde{Z}) = g(z_{i1})$, and hence payoff $U_i(z_{i1}) = g(z_{i1}) - z_{i1}$. As $g'(z_i) \leq 0$, $Y_i(0) > Y_i(z_{i1})$. Therefore, the player is better off contributing 0 to the public good. If the player contributes $z_{i2} = 0$ to the public good, $U_i(z_{i2}) = g(0) - 0$, which is greater than $U_i(z_{i1})$.

Secondly, consider a player i that contributes more than the Nash Equilibrium, such that $\beta \cdot \sum_{i=1}^{n} z_i > Z^*$, but $0 \leq z_i \leq \dfrac{c_i^*}{\beta}$. Player i contributes $z_{i3} > z_i^{NE}$, the Nash Equilibrium level of contribution. For both z_{i3} and z_i^{NE}, the threshold is met, and hence $f_i(\tilde{Z}) = V_i(\hat{z}_i)$. The corresponding payoff functions are: $U_i(z_{i3}) = V_i(z_{i3}) - z_{i3}$ and $U_i(z_i^{NE}) = V_i(z_i^{NE}) - z_i^{NE}$. As $z_{i3} > z_i^{NE}$ and $V_i'(z_i) \leq 0$, $U_i(z_{i3}) < U_i(z_i^{NE})$, and hence contributing $z_{i3} > z_i^{NE}$ is not a profitable deviation.

Thirdly, consider a player i that contributes more than the Nash Equilibrium ($z_{i4} > z_i^{NE}$), such that $\beta \cdot \sum_{i=1}^{n} z_i = Z^*$, but $0 \leq z_{i4} \geq \dfrac{c_i^*}{\beta}$. Player i has contributed more than the benefit jump from reaching the threshold, $V_i(z_{i4}) - Y_i(0)$. Hence, $U_i(z_{i4}) = V_i(z_{i4}) - z_{i4} < Y_i(0) = U_i(z_i = 0)$, and so the player is better off contributing 0 to the public good.

Fourthly, consider a player i that contributes more than the Nash Equilibrium ($z_{i5} > z_i^{NE}$), such that $\beta \cdot \sum_{i=1}^{n} z_i > Z^*$, but $0 \leq z_{i5} \geq \dfrac{c_i^*}{\beta}$. This is a combination of the second and third deviation case, and hence does not increase the players' utility, compared with the case where they contribute z_i^{NE}.

A.2 Multi-Round Game

Re-introducing time notations, $\{z\} = \{(z_1(t), ..., z_n(t))\}_{t=0}^{\bar{T}}$ is a sequence of non-negative contributions by $i = 1, ..., n$ players across $t = [0, \bar{T}]$ rounds. Similar to the one-round game, players must decide every round whether and how much to contribute to the threshold public good. These decisions are made weighing off the marginal benefits and costs, which in turn are influenced by previous contributions to the public good - both by other players and the player themselves. In the final round $\bar{T}$, previous weighted contributions by the group are $\sum_{\tau=0}^{\bar{T}-1} \sum_{i \in N} \beta(\tau) \cdot z_i(\tau)$, and the other players contributed $\beta(\bar{T}) \cdot Z_{-i}(\bar{T})$ in round $\bar{T}$. If these cumulative weighted contributions are less than Z^*, individual i must decide whether to contribute sufficiently to ensure the provision of the public good or not. The marginal cost of doing so is the total unweighted contributions outstanding to reach the threshold, divided by the weighting factor in round $\bar{T}$, $\beta(\bar{T})$:

$$\frac{1}{\beta(\bar{T})}\big(Z^* - \sum_{\tau=0}^{\bar{T}-1} \sum_{i \in N} \beta(\tau) \cdot z_i(\tau) - \beta(\bar{T}) \cdot Z_{-i}(\bar{T})\big)$$

The marginal benefit is the individual payoff of reaching the threshold as a result of covering the outstanding contributions required to reach the threshold, minus the individual payoff if the player contributes nothing to the public good in round $\bar{T}$, and the group falls short of the threshold:

$$V_i\big(\sum_{\tau=0}^{\bar{T}-1} z_i(\tau) + (Z^* - \sum_{\tau=0}^{\bar{T}-1} \sum_{i \in N} \beta(\tau) \cdot z_i(\tau) - \beta(\bar{T}) \cdot Z_{-i}(\bar{T}))\big) - Y_i\big(\sum_{\tau=0}^{\bar{T}-1} z_i(\tau)\big)$$

Using backward induction, the decisions of individuals in rounds $t < \bar{T}$ can be determined following the same evaluation of marginal costs and benefits.

Nash Equilibria

One Nash Equilibrium is that every player contributes zero in each of the $\bar{T}$ rounds of the game. In this case, aggregate contributions are zero, the public good is not provided, and an individual's payoff is $Y_i(0)$. Further Nash Equilibrium outcomes are ones where the outcome is wasteless, such that $\sum_{t=0}^{\bar{T}} \sum_{i \in N} \beta(t) \cdot z_i(t) = Z^*$, and no player wishes to deviate unilaterally, if deviation is punished with the *grim-g strategy profile* (Marx and Matthews, 2000). Under this strategy profile, each player plays g in every round, where $\{g\} = \{(g_1(t), ..., g_n(t))\}_{t=0}^{\bar{T}}$

unless $Z(t) \neq G(t)$, after which each player maximally punishes the deviation and contributes $z_i(\tau) = 0$ for $\tau > t$. For notation purposes, $G_{-i}(t) = G(t) - g_i(t)$.

Marx and Matthews (2000) demonstrate why longer contribution periods can result in larger contributions to the public good, as lump-sum contributions can be spread over multiple rounds and made contingent on past contributions. For a multi-round threshold public goods game with a benefit function $f(\tilde{Z})$, a player will not deviate from the grim-g profile in the first round if

$$Y_i(0) < V_i(g_i) - \sum_{t=0}^{\bar{T}} g_i(t) \tag{2}$$

The left hand side of Equation 2 captures the player's benefit from deviating from the grim-g strategy in the first round. If they contribute 0 instead of $g_i(0)$, the other players will contribute 0 in all future rounds, and hence aggregate contributions are $G_{-i}(0)$, falling short of the threshold Z^*.[20] The right hand side is the expected payoff of adhering to the grim-g strategy, as a result of which the threshold Z^* is reached. The generalization of Equation (2) to round $t < \bar{T}$ is:

$$\begin{aligned} Y_i\Big(\sum_{\tau=0}^{t-1} g_i(\tau)\Big) - \sum_{\tau=0}^{t-1} g_i(\tau) &\leq V_i(g_i) - \sum_{\tau=t}^{\bar{T}} g_i(\tau) - \sum_{\tau=0}^{t-1} g_i(\tau) \\ Y_i\Big(\sum_{\tau=0}^{t-1} g_i(\tau)\Big) &\leq V_i(g_i) - \sum_{\tau=t}^{\bar{T}} g_i(\tau) \\ \sum_{\tau=t}^{\bar{T}} g_i(\tau) &\leq V_i(g_i) - Y_i\Big(\sum_{\tau=0}^{t-1} g_i(\tau)\Big) \end{aligned} \tag{3}$$

where g_i is the defined sequence of player i's non-negative contributions to the public good in line with the grim-g strategy, and hence $\sum_{\tau=t}^{\bar{T}} g_i(\tau)$ refers to player i's contributions in rounds $[t, \bar{T}]$ in line with the strategy. As such, the left hand side of Equation (3) captures the continuation cost of following the grim-g strategy, and is monotonically decreasing in t. The right hand side is the continuation payoff of not deviating from the grim-g strategy. It is monotonically increasing in t: while $V_i(g_i)$ is time-invariant, $Y_i(\sum_{\tau=0}^{t-1} g_i(\tau))$ is monotonically decreasing in t, as $Y_i'(g_i(t)) < 0$, and $\sum_{\tau=0}^{t-1} g_i(\tau)$ is the aggregate of contributions in rounds $[0, t-1]$.[21]

As the left-hand side of Equation (3) is monotonically decreasing in t, and the right-hand

[20] Grim-g strategies are only a Nash Equilibrium if they are wasteless, that is, if they exactly meet the threshold: $Z^* = \sum_{t=0}^{\bar{T}} \beta(t)G(t)$. Deviations from the strategy by contributing less will result in aggregate contributions that fall short of the threshold.

[21] The right hand side of Equation (3) is strictly increasing in t if $g_i(\tau) > 0 \quad \forall \tau > 0$. This is in line with simulations in Hilbe et al. (2013), who show that dominant strategies are characterized by delayed contributions, with positive contributions in every round.

side is monotonically increasing in t, the incentive to deviate from the grim-g strategy is monotonically decreasing in t, and hence highest at $t = 0$. This is further supported by Marx and Matthews (2000), who argue that the grim-g strategy profile favours small initial contributions, with the majority of contributions made in the future, and hence contingent on other player's past adherence to the grim-g strategy profile.

Equation (3) specifies the *under-contributing constraint*, however Marx and Matthews (2000) show that a strategy profile g must also satisfy the *over-contributing constraint* in order to be a Nash equilibrium, enforced by the grim-g strategy's maximal feasible punishment. The over-contributing constraint guarantees that player i does not have the incentive to prematurely ensure the public good's threshold is met in round $t \leq \bar{T}$:

$$V_i\Big(\sum_{\tau=0}^{t} g_i(\tau) + \frac{1}{\beta(t)}\Big[Z^* - \sum_{\tau=0}^{t}\beta(\tau)G(\tau)\Big]\Big) - \frac{1}{\beta(t)}\Big(Z^* - \sum_{\tau=0}^{t}\beta(\tau)G(\tau)\Big) - \sum_{\tau=0}^{t-1} g_i(\tau) \leq V_i(g_i) - \sum_{\tau=t}^{\bar{T}} g_i(\tau) - \sum_{\tau=0}^{t-1} g_i(\tau)$$

$$V_i\Big(\sum_{\tau=0}^{t} g_i(\tau) + \frac{1}{\beta(t)}\Big[Z^* - \sum_{\tau=0}^{t}\beta(\tau)G(\tau)\Big]\Big) - \frac{1}{\beta(t)}\Big(Z^* - \sum_{\tau=0}^{t}\beta(\tau)G(\tau)\Big) \leq V_i(g_i) - \sum_{\tau=t}^{\bar{T}} g_i(\tau) \quad (4)$$

The right hand side again captures the continuation payoff of not deviating from the grim-g strategy, and is monotonically increasing in t. The left hand side captures the payoff from unilaterally ensuring the threshold is met in round $t < \bar{T}$. This captures the utility from reaching the threshold V_i as a result of the larger individual contribution (recall that $V_i'(z_i) \leq 0$). This inequality depends on $\beta(t)$: for a $\beta(t) >> 1$, an individual may have the incentive to unilaterally ensure the public good is provided. Similarly, for a $0 < \beta(t) < 1$, an individual has less incentive to unilaterally ensure the threshold is met.

While the role of the round specific multiplication factor $\beta(t)$ is not directly obvious from the *under- and over-contributing constraints*, it plays an important role. For example, the grim-g strategy profiles described by Marx and Matthews (2000), consisting of small initial contributions and larger, later contributions may be cut short by one player unilaterally ensuring the provision of the public good in round $t < \bar{T}$ if $\beta(t)$ is sufficiently large such that the inequality in Equation (4) no longer holds. Another example is the "fair share commitment" Nash Equilibrium of Hilbe et al. (2013) where players contribute nothing in the first half of the game's rounds, and maximally contribute in the second half of the game.

If $\beta(t) < 1 \quad \forall t > \frac{\bar{T}}{2}$, the group will no longer reach the threshold.

Perfect Bayesian Equilibria

Marx and Matthews (2000) show that the Nash Equilibria outcomes defined above are also Perfect Bayesian Equilibria (PBE) outcomes, if either of the following two conditions are met:

1. $c_1^* = c_i^*$ for all $i \in N$, or
2. $c_1^* = c_2^*$, and both $g_1(t) \geq g_i(t)$ and $g_2(t) \geq g_i(t)$ for all $t \geq 0$ and $i = 3, ..., n$.

Condition 1 means that all individuals have the same critical contribution c_i^*, while condition 2 specifies that the two largest critical contributions need to be the same, and belong to the two individuals who have contributed the most in every round. Under condition 1, the equilibrium outcomes are also a Subgame Perfect Equilibrium (SPE), as shown by Marx and Matthews (2000).

A.3 Theoretical Framework: Climate Protection Games

The CPG is a multiple-round threshold public goods game. The game's set of players is $N = \{1, ..., n\}$, with $n \geq 2$. Each player is endowed with private good E_i, which can be used to contribute $z_i(t)$ to a public good (the *Climate Pot*) in each round $t \geq 1$. The contribution horizon is $[1, \bar{T}]$, where $\bar{T} \leq \infty$. Contributions $z_i(t)$ are non-refundable and multiplied by a round-specific $\beta(t) > 0$ that captures the temporal dynamics of contributions.[22]

Individual i's investment in the *Climate Pot*, $z_i(t)$, must satisfy $0 \leq z_i(t) \leq \alpha E_i$, where $0 < \alpha < 1$ limits the amount individual i can invest per round. Furthermore, $\sum_{t=1}^{\bar{T}} z_i(t) \leq E_i$ (implying $\alpha \cdot \bar{T} \leq 1$).

$\tilde{Z}(t) = \beta(t) \sum_{j \in N} z_j(t)$ is the sum of contributions to the *Climate Pot* by the set of individuals N in round t. $\beta(t) > 0$ is a scaling factor whereby contributions made to the *Climate Pot* can be scaled up ($\beta(t) > 1$), down ($\beta(t) < 1$), or not scaled ($\beta(t) = 1$).

Individual monetary payoffs depend on whether the sum of scaled group-level contributions to the public good across $\bar{T}$ rounds reach a threshold level, Z^*: $\sum_{t=1}^{\bar{T}} \tilde{Z}(t) \lesseqgtr Z^*$. The threshold is set such that $\sum_{t=1}^{\bar{T}} \beta(t) z_i(t) < Z^* < \sum_{i=1}^{n} E_i$, meaning that the threshold cannot be reached by an individual's contributions alone, however it is less than the total initial endowments of all players in a group.[23] There are two cases:

[22]Note that setting $\beta(t) = 1 \quad \forall t \in [1, \bar{T}]$ results in no temporal dynamics, and simplifies to the setup of Marx and Matthews (2000), Milinski et al. (2008), and the *Linear* treatment.

[23]This implies that $n > \max_t\{\beta(t)\}$.

Case 1: $\sum_{t=1}^{\bar{T}} \tilde{Z}(t) \geq Z^*$

In this scenario, the set of individuals N's weighted cumulative contributions across rounds $[1, \bar{T}]$ to the *Climate Pot* are equal to or exceed the threshold level, and thus disastrous climate change is averted. All individuals $i \in N$ keep their savings, $E_i - \sum_{t=1}^{\bar{T}} z_i(t)$. Individual $i's$ monetary payoff is:

$$\pi_i = E_i - \sum_{t=1}^{\bar{T}} z_i(t)$$

Case 2: $\sum_{t=1}^{\bar{T}} \tilde{Z}(t) < Z^*$

In this scenario, the set of individuals N's cumulative contributions across rounds $[1, \bar{T}]$ to the *Climate Pot* will fall short of the threshold level. Thus disastrous climate change is not averted. There is a probability p that individual $i \in N$ keeps their savings, $E_i - \sum_{t=1}^{\bar{T}} z_i(t)$, and probability $(1 - p)$ that they lose their savings and end up with 0.[24] Individual $i's$ monetary payoff is:

$$\pi_i = \begin{cases} E_i - \sum_{t=1}^{\bar{T}} z_i(t) & \text{with probability } p \\ 0 & \text{with probability } (1-p) \end{cases}$$

Hence, $\mathbb{E}[\pi_i | \sum_{t=1}^{\bar{T}} \tilde{Z}(t) < Z^*] = p \cdot [E_i - \sum_{t=1}^{\bar{T}} z_i(t)]$. Figure 4 depicts individual i's payoffs as a function of the aggregate weighted group contributions $\sum_{t=1}^{\bar{T}} \tilde{Z}(t)$, and b_i reflects the benefit jump from marginally reaching the public good's threshold.

A.3.1 Representation of the Lab Experiment

Across all treatments, there are five individuals per group, so $n = 5$ and $N = \{1, ..., 5\}$. Each individual is endowed with $E_i = 36$, and can contribute a maximum of 4 to the *Climate Pot* per round. The CPG runs for a total of $\bar{T} = 9$ rounds. The threshold level Z^* is 90, and the probability with which an individual keeps their savings if the threshold is not met, is $p = 30\%$.

Linear **vs.** ***Step*** **CPG**

In the *Linear* set-up of the game, contributions to the *Climate Pot* are neither amplified nor reduced across the nine rounds, hence $\beta(t) = 1$ for $t = \{1, ..., 9\}$. On the contrary, in the

[24] This probability is individual-specific, and thus if a group fails to reach the threshold, some group members can get a positive payoff while others may not. This differs from Milinski et al. (2008) where the probability is group-specific. The probability being individual-specific captures the idea that disastrous climate change may have varying levels of impact on different countries.

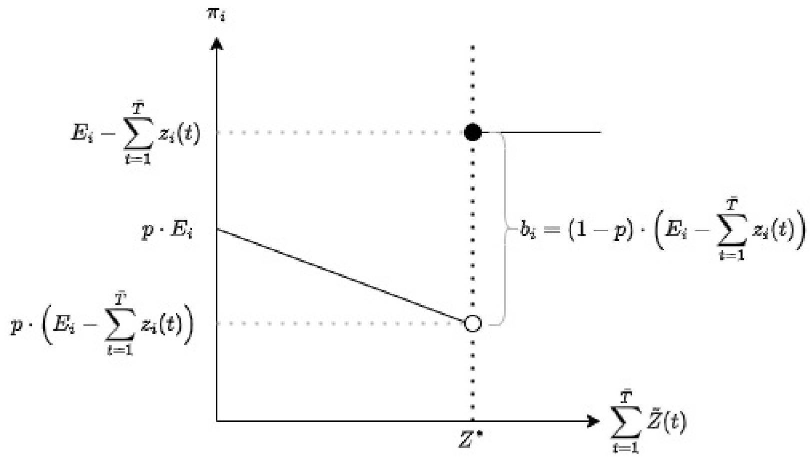


Figure 4. Individual Payoff Function

(more) realistic set-up, the β_t is a *Step* function that varies according to the round of the game. Thus the returns to contributions decay, as $\beta(t) = 1.5$ for $t = \{1, 2, 3\}$, $\beta(t) = 1.0$ for $t = \{4, 5, 6\}$, and $\beta(t) = 0.5$ for $t = \{7, 8, 9\}$.

Nash Equilibria

The Climate Protection Game, both in its *Linear* and *Step* version, presents two, Pareto-ranked symmetric pure strategy Nash Equilibria (NE) for risk-neutral players.[25] Either all players contribute 0 per round to the *Climate Pot*, or they contribute 2 per round. The first Nash Equilibrium, in which all players decide not to contribute, results in expected earnings of $0.30 \cdot 36 + 0.70 \cdot 0 = 10.80$. This equilibrium exists due to the *No Deviation Constraint*, meaning that players cannot unilaterally reach the threshold (Croson and Marks, 2000). The second, Pareto dominant, symmetric pure strategy NE ensures that players contribute 2 per round, thus coordinating such that the 90 threshold level is met in the final round, and all players receive earnings of 18.

Both the *Linear* and *Step* CPG have numerous asymmetric NE where individuals contribute varying amounts (across rounds) to the *Climate Pot*, depending on the likelihood of meeting the threshold, thus ensuring that past investments were not in vain. Each path of contributions where the threshold of 90 is exactly met, and no individual contributes more than 25.2 constitutes a NE as players do not have a profitable deviation (Abou Chakra and Traulsen, 2012).[26]

[25] Symmetric means across players and rounds. Hence, all players make the same contribution to the *Climate Pot* in every round of the game.

[26] $25.2 = 36 \cdot 0.7$, the expected losses from contributing zero to the public good and the threshold

The *Step* CPG offers many additional NE, with the most efficient one consisting of each player contributing 4 in rounds $t = \{1, 2, 3\}$ and contributing 0 in the remaining rounds, thus reaching the threshold with 24 left in savings and hence earnings. On the flip side, strategies that emerge as dominant strategies in Evolutionary Game Theory simulations of the *Linear* CPG by Hilbe et al. (2013) do not get passed along in the *Step* CPG - for example a strategy where players contribute 0 in each round of the first half of the CPG, and the maximum amount (4) in the remaining rounds. This strategy acts as a self-commitment, as well as a credible signal to the other players that one won't contribute more than the "fair share" (Hilbe et al., 2013). However due to $\beta_t < 1$ in the later rounds of the *Step* CPG, groups consisting of individuals deploying this strategy fall short of the threshold.

B Evolutionary Game Theory Simluations

The EGT simulations build on the simulations by Abou Chakra and Traulsen (2012).[27] Similar to Abou Chakra and Traulsen (2012), each individual has a strategy per round t that consists of (i) a threshold τ_i, (ii) a contribution j_i made if the aggregate group contributions exceed the threshold τ_i, and (iii) a contribution k_i made if the aggregate group contributions do not exceed the threshold τ_i. At the start of the simulation, every player is assigned a random strategy (in terms of $\{\tau_i, j_i, k_i\}$) for each of the nine rounds of the CPG.

In each generation, 1000 games are played. For each game, five random players are drawn to participate. Per generation, each player's payoff is the sum of total payoffs per game they participated in (based on the payoff rules described in Section 3), divided by the number of games they participated in. At the end of each generation, a player's payoff is translated into their fitness value: $f_i = exp(\delta \cdot \pi_i)$, where π_i is the player's payoff, and δ is the selection intensity.

In the transition from one generation to the next, the next generation's strategies are selected based on a Wright-Fisher process, where an individual's fitness in the previous generation determines the probability that their strategy is passed on to the next generation. Strategies with a higher fitness are more likely to be passed on. Errors are incorporated into this process: with a probability μ, Gaussian noise with standard deviation σ are added to τ_i, j_i, k_i of each round, independently. Averages are computed from 100,000 generations from 150 iterations per treatment arm.

not being met, which applies to both *Linear* and *Step* settings. This is required to satisfy the Individual Rationality condition, ensuring that no individual contributes more than their own valuation of meeting the threshold.

[27] The code is written in Python, rather than C++, as is the case with Abou Chakra and Traulsen (2012).

The specifications of the EGT simulation are as follows:

Variable	Value
δ	1.0
μ	0.03
σ	0.15

Table 1: Values of EGT Simulation Parameters

Table 2: Overview of Results from EGT Simulations

	Linear EGT	Step EGT	p-value
Contributions Rounds 1-3	46.63	49.81	0.0013
	(0.71)	(0.68)	
Contributions Rounds 4-6	32.98	19.19	0.0000
	(0.91)	(1.01)	
Contributions Rounds 7-9	17.64	4.78	0.0000
	(0.84)	(0.54)	
Total Unweighted Contributions	97.25	73.78	0.0000
	(0.48)	(0.62)	
Total Weighted Contributions	97.25	96.29	0.2052
	(0.48)	(0.58)	
Prob. Reaching Threshold	0.95	0.96	0.1612
	(0.48)	(0.58)	
Hypothetical Payoffs	16.04	20.39	0.0000
	(0.03)	(0.09)	

Notes: p-values are based on a two-sided t-test with $n_1 = 150$ and $n_2 = 150$. Standard errors are reported in parentheses.

Figure 3 illustrates that the *Linear* contributions in the EGT simulation differ from the *Linear* contributions observed in the lab experiment. Nevertheless, the *Linear* contributions in the EGT simulation follow a similar pattern to the *Step* contributions (both in the lab and EGT simulations). Comparing the two EGT treatments, contributions to the public good are statistically significantly lower in the *Linear* treatment in the first three rounds, as reported in Table 2. In rounds 4-6 and 7-9, contributions in the *Linear* treatment are substantially higher, resulting in greater total unweighted contributions to the public good. This in turn leads to significantly lower (hypothetical) expected payoffs for the *Linear* arm, whose average payoff is 16.04 ± 0.03, compared with 20.39 ± 0.09 for the *Step* arm. Nevertheless, the findings of the EGT simulations closely mirror those of the lab experiment.

C Regressions - Clustered SE at Group Level

Table 3: Past Individual Contributions on Free-riding

	Free-riding		
	(1)	(2)	(3)
	OLS	Probit	Logit
Player Past Cum. Contrs	-0.0092*	-0.0294*	-0.0589*
	(0.0040)	(0.0136)	(0.0251)
N	2235	2235	2235

Notes: Columns (1), (2), and (3) show the results of a regression of Player's individual cumulative past contributions, on their likelihood to free-ride in the current round. Control variables include Other Group Members' cumulative past contributions, the treatment status, a dummy for the year of the experiment, the $\beta(t)$ scaling factor, the player's age and gender, their risk and loss aversion, and their experiment comprehension score. Standard errors (reported in parentheses) are clustered at the group level. Column (1) reports results from an OLS regression, (2) a probit regression, and (3) a logit regression. I exclude observations from the first round, and observations where the group's aggregate contribution already reached or exceeded 90 in the previous round. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 4: Past Group Contributions on Free-riding

	Free-riding		
	(1)	(2)	(3)
	OLS	Probit	Logit
Group Past Cum. Contrs	0.0008	0.0047	0.0082
	(0.0006)	(0.0027)	(0.0047)
N	2235	2235	2235

Notes: Columns (1), (2), and (3) show the results of a regression of a player's Group cumulative past contributions, on their likelihood to free-ride in the current round. Control variables include the treatment status, a dummy for the year of the experiment, the $\beta(t)$ scaling factor, the player's age and gender, their risk and loss aversion, and their experiment comprehension score. Standard errors (reported in parentheses) are clustered at the group level. Column (1) reports results from an OLS regression, (2) a probit regression, and (3) a logit regression. I exclude observations from the first round, and observations where the group's aggregate contribution already reached or exceeded 90 in the previous round. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 5: Game Round on Contributions and Free-Riding

	Below Nash Equilibrium Contribution			Free-riding		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Probit	Logit	OLS	Probit	Logit
Game Round	0.0577***	0.1627***	0.2698***0	0.0665***	0.2210***	0.4154***
	(0.0158)	(0.0452)	(0.0794)	(0.0156)	(0.0565)	(0.1120)
N	2535	2535	2535	2535	2535	2535

Notes: Columns (1), (2), and (3) show the results of a regression of the game's round, on a player's likelihood to contribute less than 2 (a symmetric pure Nash Equilibrium) in the current round. Columns (4), (5), and (6) show the results of a regression of the game's round, on a player's likelihood to free-ride in the current round. Control variables include the Group Past Cumulative Contributions, treatment status, a dummy for the year of the experiment, the $\beta(t)$ scaling factor, the player's age and gender, their risk and loss aversion, and their experiment comprehension score. Standard errors (reported in parentheses) are clustered at the group level. Columns (1) and (4) reports results from an OLS regression, (2) and (5) a probit regression, and (3) and (6) a logit regression. The number of observations is larger than in previous tables, as observations are included from the first round. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

D Regressions - Clustered SE at Individual Level

Table 6: Past Individual Contributions on Free-riding

	Free-riding		
	(1)	(2)	(3)
	OLS	Probit	Logit
Player Past Cum. Contrs	-0.00917	-0.0294	-0.0589
	(0.00486)	(0.0163)	(0.0302)
N	2235	2235	2235

Notes: Columns (1), (2), and (3) show the results of a regression of Player's individual cumulative past contributions, on their likelihood to free-ride in the current round. Control variables include Other Group Members' cumulative past contributions, the treatment status, a dummy for the year of the experiment, the $\beta(t)$ scaling factor, the player's age and gender, their risk and loss aversion, their experiment comprehension score, and we include group-level fixed effects. Standard errors (reported in parentheses) are clustered at the individual level. Column (1) reports results from an OLS regression, (2) a probit regression, and (3) a logit regression. I exclude observations from the first round, and observations where the group's aggregate contribution already reached or exceeded 90 in the previous round. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 7: Past Group Contributions on Free-riding

	Free-riding		
	(1)	(2)	(3)
	OLS	Probit	Logit
Group Past Cum. Contrs	0.0008	0.0047	0.0082
	(0.0009)	(0.0044)	(0.0078)
N	2235	2235	2235

Notes: Columns (1), (2), and (3) show the results of a regression of a player's Group cumulative past contributions, on their likelihood to free-ride in the current round. Control variables include the treatment status, a dummy for the year of the experiment, the $\beta(t)$ scaling factor, the player's age and gender, their risk and loss aversion, their experiment comprehension score, and we include group-level fixed effects. Standard errors (reported in parentheses) are clustered at the individual level. Column (1) reports results from an OLS regression, (2) a probit regression, and (3) a logit regression. I exclude observations from the first round, and observations where the group's aggregate contribution already reached or exceeded 90 in the previous round. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.

Table 8: Game Round on Contributions and Free-Riding

	Below Nash Equilibrium Contribution			Free-riding		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Probit	Logit	OLS	Probit	Logit
Game Round	0.0577	0.1627*	0.2698	0.0665*	0.2210*	0.4154*
	(0.0289)	(0.0804)	(0.1418)	(0.0288)	(0.0969)	(0.1935)
N	2535	2535	2535	2535	2535	2535

Notes: Columns (1), (2), and (3) show the results of a regression of the game's round, on a player's likelihood to contribute less than 2 (a symmetric pure Nash Equilibrium) in the current round. Columns (4), (5), and (6) show the results of a regression of the game's round, on a player's likelihood to free-ride in the current round. Control variables include the Group Past Cumulative Contributions, treatment status, a dummy for the year of the experiment, the $\beta(t)$ scaling factor, the player's age and gender, their risk and loss aversion, their experiment comprehension score, and we include group-level fixed effects. Standard errors (reported in parentheses) are clustered at the individual level. Columns (1) and (4) reports results from an OLS regression, (2) and (5) a probit regression, and (3) and (6) a logit regression. The number of observations is larger than in previous tables, as observations are included from the first round. ***, ** and * represent significant differences at the 1, 5 and 10% level, respectively.