Climate policies under wealth inequality

Siddharth Sahay
602334
ULB
siddharth.sahay@ulb.be

Salma Namouri
604449
ULB
salma.namouri@ulb.be

Jordi Ugarte
604645
ULB
jordi.ugarte.yaffar@ulb.be

ABSTRACT

Climate policies under wealthy inequality is a paper written by *Vítor V. Vasconcelos, Francisco C. Santos, Jorge M. Pacheco* and *Simon A. Levin* that studies the influence and impact of wealth inequality in society over climate change.

The paper highlights the many challenges of aligning rich and poor countries to implement policies that help mitigate the looming climate catastrophe. The paper frames this issue as a threshold public goods dilemma, where participants decide whether to cooperate or defect. This means players must balance potential future gains against present costs. The model reflects real-world disparities, with wealthy players contributing more on average but being a minority, while poorer players, though more numerous, contribute less. To maintain overall success, wealthy players often compensate for the shortfall in contributions from poorer players. [1]

One of the key factors influencing the results is homophily, the tendency of individuals to "copy" those similar to them. High levels of homophily can hinder behaviour change (within a population class and globally), reducing cooperation and essentially destabilizing collective efforts to combat climate change. Conversely, when homophily is minimized (i.e., players imitate each other indiscriminately across population class), broader cooperation is achievable and there is an increase in the likelihood of achieving public goods threshold. This study provides valuable insights into the dynamics of wealth inequality and its implications for global climate policy. It emphasizes the importance of fostering inclusivity and cross-group influence to address the collective risks posed by climate change.

1 INTRODUCTION

The tragedy of the commons is a theme that reappears constantly in public goods games (PGGs). The authors of this paper explore this theme in detail, with a focus on risk perception and wealth inequality. A participant who pays the cost of cooperation also provides a shared benefit to the group, whereas defectors enjoy the benefits without incurring any cost. Indeed, the rational choice is to "free-ride" on the benefits produced by others, and this dynamic often leads to suboptimal outcomes, where the collective good is "exploited". This is otherwise known as the aforementioned "tragedy of the commons".

In their paper, the authors introduce wealth inequality into the framework of a Threshold Public Goods Game (PGG), a special case of PGG, where the model requires surpassing a certain level of collective contributions for success. They investigate how disparities between "rich" and "poor" participants influence cooperation when the group faces the risk of collective failure if the threshold is not met. The model also adds a factor called homophily, or more simply, the preference of individuals to interact with those of similar wealth. This adds another layer of complexity to the social dynamics.

According to us, the most striking discovery in the paper is that wealth inequality can promote cooperation under certain conditions.

This is quite counterintuitive, given the assumption is usually that inequality would hinder collective action! The authors discovered that wealth inequality could actually enhance cooperation when homophily—the tendency of individuals to interact within their wealth class—is low or absent.

This finding underscores a fundamental insight into human behaviour. Numerous studies have been conducted on this very topic, even in fields such as economics: the dynamics of influence and imitation are not unidirectional. In mixed-wealth populations, not only do the poor imitate the behaviour of the rich, but the rich also imitate the poor more frequently. This mutual feedback loop suggests that cooperation can emerge and stabilize even in unequal settings, provided individuals are exposed to diverse social influences[1] and shared risk perceptions.

On the other hand, high levels of homophily lead to a collapse in cooperation, particularly among the poor, resulting in an overall failure to meet the threshold. When participants are not restricted by homophily (tendency to copy like participants), the wealthier individuals often compensate for the lower contributions of the less wealthy, increasing the likelihood of achieving the collective goal [1].

To analyze these dynamics, the authors developed an agent-based model. The population is divided into two wealth classes: "rich" agents (Z_R) and "poor" agents (Z_P), where $b_R > b_P$. Each individual adopts one of two strategies: cooperation (C), contributing a fraction of their endowment (c), or defection (D), contributing nothing. Success is achieved if the total contributions meet a threshold, $M \cdot c \cdot \bar{b}$, where $\bar{b}$ is the average endowment of the whole population, this has been kept at the value of 1 throughout (by adjusting b_R and b_P accordingly).

The two main parameters include:

- **Risk perception (r)**: This represents the probability of losing the endowment if the threshold is not met. The payoff for an individual in such a scenario is calculated as:

  $$\text{Payoff} = b \times (1 - r) - \text{Cost of Cooperation.}$$

  The cost of cooperation, of course, only applies to cooperators (or Cs). Defectors (Ds) do not pay this cost.
- **Homophily (h)**: This parameter determines the extent to which interactions occur within the same wealth class. A high homophily value ($h = 1$) indicates that interactions are restricted to individuals of the same wealth class, while

[1] Olson's The Logic of Collective Action: small groups with concentrated interests are often better at overcoming free-rider problems compared to larger, more dispersed groups.

a low homophily value ($h = 0$) allows interactions and copying behavior across different wealth classes.

There are other parameters that the paper tweaks as well. They include obstinate behaviour, which is where a part of the population is assumed not to change behaviour regardless of the fitness of other populations. In cases with obstinate cooperators, even if from the poor subpopulation, the average group achievement, η_G, is orders of magnitude higher; $\eta_G = 0.581$ instead of $\eta_G = 0.004$.[1]

Strategy evolution is governed by stochastic dynamics using a pairwise comparison rule, where individuals imitate others probabilistically based on payoff differences. Stationary distributions and gradients of selection visualize long-term configurations and evolutionary trajectories.

The results show that low homophily ($h = 0$) enhances cooperation as the rich compensate for the lower contributions of the poor, improving group success. In contrast, high homophily ($h = 1$) collapses cooperation among the poor, leading to widespread defection (which in turn makes groups lose the threshold PGGs they play across the board). This interplay between risk, homophily, and wealth inequality provides critical insights into fostering cooperation under unequal conditions.

2 RELATED WORK

The paper *Risk of collective failure provides an escape from the tragedy of the commons*[2] also studies public goods dilemmas through the lens of evolutionary game theory, focusing on the parameters of risk perception, group size, and cooperation thresholds. The study shows (much like this one) that higher perceived risk and smaller, localized groups foster cooperation more effectively than large-scale endeavours. Their methodology involves modelling the dynamics of cooperation and defection using replicator equations and exploring the impact of how different the composition of a group is.

Next, we looked at the paper *Social Diversity Promotes the Emergence of Cooperation in Public Goods Games*[3]. This paper studies how diversity in individual endowments and interaction networks impacts cooperation in public goods games. Using theoretical modelling, the authors show that diverse social structures promote higher cooperation levels due to the varying influence of individuals in heterogeneous networks. While both our paper and this study emphasize diversity, the earlier paper (also by Santos) explicitly models wealth inequality between rich and poor participants, linking it to real-world economic disparities. Also, the concept of homophily is still absent in this study, which is central to our paper and study. Our paper builds on top this by adding an understanding of how this social/economic parameter (homophily) might affect collective action.

Lastly, we looked at another paper *Dynamic instability of cooperation due to diverse activity patterns in evolutionary social dilemmas*[4] which studies the impact of probabilistic participation. Here, players may abstain from interactions due to inactivity, leading to cascading failures of cooperation. In contrast, our paper focuses on how wealth inequality and homophily influence cooperative dynamics in global public goods games, such as climate change mitigation. Both these works highlight the instablity of cooperation but through different lenses: Xia et al. emphasise dynamic and immediate participation, while this paper centers on rich/poor interactions and homophily (in essence an economic factor). Together, these studies highlight the importance of structural factors, be they network-related effects or economic!

3 PRACTICAL ASPECTS OF THE REPRODUCTION

To reproduce this paper, we built a model from scratch implementing details of the climate threshold public game described in the paper by hand.

Game class parameters

The model simulates a population of Z individuals, split into two wealth classes: rich (Z_R) and poor (Z_P). Each individual adopts one of two strategies: cooperation (C) or defection (D). The parameters used in the simulation are:

- Z: Total population size, 200.
- Z_R, Z_P: Number of rich and poor individuals, with $Z_R + Z_P = Z$. Rich fraction is 20%, so $Z_P = 160$, $Z_R = 40$.
- b_R, b_P: Endowments for rich and poor individuals, respectively. $b_R = 2.5$, $b_P = 0.625$. The average endowment, $\bar{b}$, across the population is 1.
- c: Contribution factor, fraction of endowment contributed by cooperators, 10%.
- M: Threshold used to calculate success within groups, along with c and average($b_{R/P}$), $M = 3$.
- r: Risk perception, probability of loss if the group fails. Varies between $[0, 1]$.
- h: Homophily, the degree to which individuals interact within their wealth class. Varies between $[0, 1]$.
- μ: Mutation rate, the likelihood of random strategy changes. Mostly taken as $1/Z$, 0.005 or 0.01.
- β: Selection intensity, the strength of payoff-driven strategy adoption. Mostly taken as 5.

State Space Representation

The state of the whole system can be represented by (C_R, C_P), where C_R and C_P are the numbers of cooperators among the rich and poor. This compact representation allows the entire population configuration to be expressed as a simple tuple of (C_R, C_P) since defectors can be inferred as $D_R = Z_R - C_R$ and $D_P = Z_P - C_P$.

To efficiently manipulate states, a scalar representation s was defined using the mod scalar function:

$$s = C_R \cdot Z + C_P, \tag{1}$$

This enables us to project onto the x and y axes, it's particularly helpful to build the transition matrix (and map back to a stationary distribution).

Fitness Calculation

The fitness[2] of individuals quantifies their average success in groups, determined by contributions and payoffs. For example, the fitness

[2]We have re-written some of the equations from the main paper for better clarity.

of a rich cooperator (f_R^C) is:

$$f_R^C = \frac{1}{\binom{Z-1}{N-1}} \sum_{j_R} \sum_{j_P} \mathbb{P}(j_R, j_P) \Pi_R^C(j_R+1, j_P), \tag{2}$$

where:

- $\mathbb{P}(j_R, j_P)$ is the probability of selecting j_R rich and j_P poor cooperators, calculated using multivariate hypergeometric sampling!
- $\Pi_R^C(j_R+1, j_P)$ is the payoff for a rich cooperator in a group composed in this way.

Therefore, we take into account the probabilistic existence of group configurations in this way. That is, a group that is highly likely to exist will essentially count more towards the fitness than a group which is less probable. Similar calculations apply for f_R^D, f_P^C, and f_P^D.

Also, note that this fitness is for all individuals of a wealth class/-subpopulation, not for specific individuals!

Transition Matrix

The transition matrix W encodes probabilities of moving from one state (C_R, C_P) to another (C'_R, C'_P) [verbatim]:

$$T_k^{X\to Y} = \frac{i_k^X}{Z_k}\left[(1-\mu)\cdot\left(\frac{i_k^Y}{Z_k - 1 + (1-h)Z_l}\left(1+e^{\beta(f_k^X - f_k^Y)}\right)^{-1}\right.\right. \tag{3}$$

$$\left.\left. + \frac{(1-h)i_l^Y}{Z_k - 1 + (1-h)Z_l}\left(1+e^{\beta(f_k^X - f_l^Y)}\right)^{-1}\right) + \mu\right], \tag{4}$$

where:

- f_k^X and f_k^Y are the fitness values for strategies X and Y in "our" subpopulation k. So, f_R^C and f_R^D respectively.
- f_l^Y is the fitness of strategy Y in the "other" subpopulation l. If we were calculating the $T_R^{C\to D}$ then this would be f_P^D.

Notice how homophily, h, controls the influence of the other subpopulation l. If h is 1, i.e., the tendency to imitate individuals from the same population is high, then the entire second term with $(1-h)$ cancels out, and $T_k^{X\to Y}$ depends entirely on f_k^X and f_k^Y and there is no influence from f_l^Y (which is the fitness of the "other" subpopulation's opposing strategy). This idea is key to incorporating homophily into the mix!

Each entry $[i, j]$ of the transition matrix W represents the probability of changing state from j to i where both of these are represented by the scalar projection of a state configuration (C_R, C_P). As an example, the index $[0, 1]$ will represent the movement from state 1, which is $(0, 1)$, to state 0, which is $(0, 0)$. This entry can be written as $W_{1,0}$. Similarly, the probability to stay in the same state, say 0, would be $W_{0,0}$. [3]

Diagonal entries are adjusted to ensure each row sums to 1, since these are essentially probabilities!

[3]Reference: GeeksForGeeks.org - Transition Probability Matrix

Stationary Distribution

The stationary distribution π is computed by solving the eigenvalue problem:

$$W^T \pi = \pi, \tag{5}$$

where π represents the stationary distribution we are trying to compute. Once computed, it is has all the states in the matrix (C_R, C_P).

To compute this, we use *numpy* to get the transposed transition matrix, and solve for the eigenvector corresponding to eigenvalue 1[4] and then normalised that to obtain π. [5]

The stationary distribution, π, shows the "steady-state behaviour" of the population with the game configuration. Each entry in π represents the likelihood of the system being in a particular state (C_R, C_P) after an infinite number of transitions! This is because $W^T \cdot \pi$ is again π, i.e., the distribution remains unchanged under the dynamics of the Markov chain. As a quick aside, a Markov chain is a model that represents a system that transitions between states in a sequence and the probability of moving to the next state depends only on the current state. It's essentially "memory-less", or more formally, this property is called the Markov property.

Since each value in the distribution is the long-term probability of the system being in that state, let's say a state $(C_R = 36, C_P = 129)$ has a value of 0.9, then it means the system finds itself in this state 90% of the time. It reflects how strategies (cooperation or defection) converge or stabilise over time in each of the subpopulations. By weighting the configurations based on π, we can also calculate other aggregate metrics, such as the average group achievement (η_G).

Average Group Achievement (η_G)

The average group achievement, according to the paper, is calculated as a weighted sum of success probabilities across all configurations and the corresponding value in the stationary distribution (i.e., how likely it is for that state to occur):

$$\eta_G = \sum_{(C_R, C_P)} \pi(C_R, C_P) \cdot a_G(C_R, C_P), \tag{6}$$

where $a_G(C_R, C_P)$ is the fraction of successful groups in the state (C_R, C_P). and $\pi(C_R, C_P)$ is the value of the π matrix at the index $[C_R, C_P]$.

We also use multivariate hypergeometric sampling for this since in each state (i_R, i_P), there could exist many groups with the size of N with these many rich and poor cooperators.

Implementation design choices

We had to make several assumptions while implementing the model described in the original paper. Initially, while creating the model, it was hard to test until the entire game class was written, so there was a bit of wandering in the dark for a while as the class was being developed. The value of β in some cases is not made clear explicitly and has a noticeable impact on the results.

The calculation of η_G was also slightly unclear in the original paper, and so we made some assumptions about the calculation of $M.c.\bar{b}$.

[4]Reference: 3Blue1Brown's Eigenvectors and eigenvalues | Chapter 14, Essence of linear algebra

[5]Reference: Brilliant.org - Stationary Distributions

4 METHODS

Game model implementation

Initially while starting the implementation, we were exploring options to use EGTTools[5] to replicate the game model. The library provides great functionality to create models with built-in methods to calculate the transition and stationary distributions, which would come in handy for us. It has several modes like *AbstractNPlayerGame* and *MatrixNPlayerGameHolder* to simplify development. Unfortunately, it doesn't yet support multi-population games, and in this game model, we restrict movement from each population to the other; this goes without saying, but a Rich country cannot become a Poor one, and vice versa.

We could use this library for computing the stationary distribution, given the transition matrix, but decided to use *numpy* instead since it was already imported for our project and used almost ubiquitously.

To calculate the transition matrix, we used the concept of a tridiagonal matrix[6], where the only non-zero elements in the matrix are on the main diagonal, going from the top-left to the bottom-right. This is done because in our model, we only calculate transitions to adjacent states:

$$(C_{R_i}+1, C_{P_i}), (C_{R_i}-1, C_{P_i}), (C_{R_i}, C_{P_i}+1), (C_{R_i}, C_{P_i}-1)$$

Another small detail we initially missed was that this model is not being evolved, but rather we compute the fitness values, transition matrix, stationary distributions and gradient of selection in fixed system configurations. This is a very clever way to look at group dynamics from a global perspective in this scenario. It also gives us key insights about how the model would evolve from a certain configuration. Although we do evolve the model later on to study the games played within subpopulations, we do not have to evolve the model for the initial experiments.

Plotting

We used *Matplotlib*[7] throughout the *Jupyter Notebook*[8] to create our graphs and visualisations. The graphs ranged from simple line and scatter plots to more complex ones, like *Quiver* plots for showing vector fields, such as gradients of selection. These visualizations were crucial for understanding the model's outcomes, like how strategies evolve in a population or the stationary distribution. We spent a lot of time making the graphs clear and readable, carefully choosing colours, labels, and legends so they communicated the results effectively.

Use of large-language models in the writing of this report

Although the official documentation for *matplotlib* is extremely helpful, tweaking the plots just right takes time. For this purpose, we used ChatGPT's free tier large-language model (LLM) *4o* to help us correctly tune the graphs. The trickiest was the *Quiver* graph for the stationary distribution.

We also used the model to help us with some LaTeX syntax, especially for the formulas and equations. It made working with this a lot easier; formulas and equations can be tedious to type and format by hand! Additionally, we used the free tier of Grammarly to help write the report and catch small grammatical errors.

Our use of LLMs was mostly restricted to plots and equations in the report. It was also useful in understanding some esoteric concepts around eigenvectors and eigenvalues, as well as the idea of a transition matrix. For more concrete tasks though, we have found, throughout our time with these new LLMs that ultimately the results are often counter-productive. As an example, when it comes to debugging or fine-tuning complex code with many moving parts and subtle interplay of components, the suggestions are still frequently misleading, which end up wasting time and, at times, even introducing unwanted bugs!

In general, we've found that tools like *GitHub Copilot*[9], which work more line by line rather than spitting out entire chunks of code are better suited to streamline coding workflows, but in our case, since we used a Jupyter notebook, we had to more-or-less write everything from scratch.

Of course, we also understand that LLMs like ChatGPT, Claude.ai and the like are not a replacement for deeper understanding, knowledge, or problem-solving, so in general, we keep away from them for complex tasks.

5 RESULTS

The results from our experiments are presented under various headings as below. Our model takes as input many of these parameters which dictate how these visualisations are created. The primary focus of our study will be the risk, r, and homophily, h. We will later also alter the spread of the total population, Z_R and Z_P, and the endowments for the rich (b_R) and poor (b_P) and visualise how these parameters affect overall cooperation.

Another interesting idea is freezing evolution in one subpopulation and studying how the other subpopulation changes. This highlights the games that are being played within each subpopulation.

5.1 Average Group Achievement, η_G

The gray line shows the average group achievement in the case of no wealth inequality; that is, all individuals have an initial endowment $b = \bar{b} = 1$ and the cost of cooperation is $0.1b$. The blue line shows results for wealth inequality with the homophily parameter $h = 0$, whereas the red line shows results for $h = 1$. We split the population of $Z = 160$ individuals into $Z_R = 32$ rich (20%) and $Z_P = 128$ poor (80%); initial endowments are $b_R = 2.5$ and $b_P = 0.625$, ensuring that the average endowment $\bar{b}$ remains $\bar{b} = 1$ (used to generate the gray line); the cost of cooperation also remains, on average, $0.1\bar{b}$, which means $c_R = 0.1b_R$ and $c_P = 0.1b_P$.

The observed differences between our reproduced model and the results in the original paper can be attributed to the following factors:

Differences observed in our graph vs. the authors' results:

- **Population Size (Z)**: The original paper uses $Z = 200$, while our implementation (only in this case) uses $Z = 160$,

[6] https://en.wikipedia.org/wiki/Tridiagonal_matrix
[7] matplotlib: Official Docs
[8] Jupyter notebook for the reproduction of this model can be found on GitHub.
[9] https://github.com/features/copilot

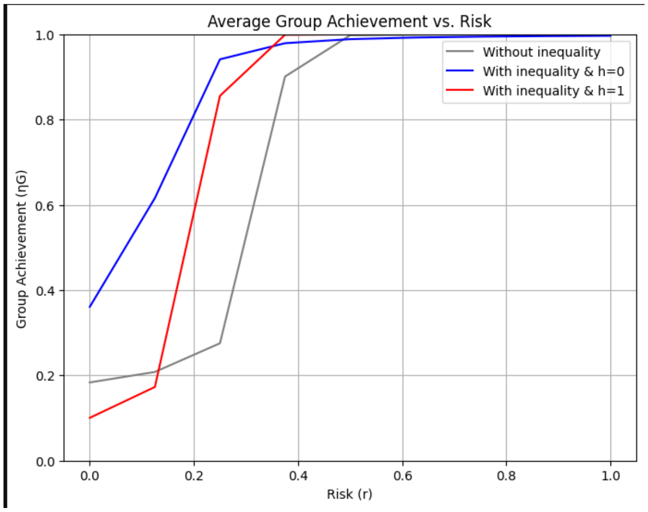

Figure A: Average group achievement as a function of risk

calculating the values for this graph was extremely computationally heavy[10]. Smaller populations of course amplify fluctuations, particularly impacting low-risk scenarios. We also kept the group size ($N = 6$) to be the same, which also had an impact.

- **Calculation of η_G**: In our model, η_G is calculated as the weighted sum of average fraction of groups reaching the threshold with the stationary distribution of that configuration of (C_R, C_P). Even at low risk ($r = 0$), some groups achieve the threshold, leading to a non-zero η_G. We don't fully understand how the original paper reports $\eta_G = 0$ for low r, possibly due to differences in how success is defined or calculated (apart from reaching the threshold of $Mc\bar{b}$ per group via multivariate hypergeometric sampling). Unfortunately, the exact implementation details of η_G were not discussed, even in the SI Text[1].
- **Threshold Success Dynamics**: Differences in the method of calculating how many groups cross the threshold (e.g., summation of contributions) likely contribute to the variation, especially for edge cases. If we had a smaller group size, in relation to the smaller population, we might have replicated the results of the authors more accurately.

Our results suggest that even under low-risk conditions, nonzero successes contribute to collective achievement! As a silver lining, this offers a more optimistic perspective, though it may be misinterpreting the author's definition of threshold success.

Interpretation of the results: Since our results don't exactly replicate the result of the experiment by the authors, we will analyse *Figure 1* from the original paper.

The graph in the original paper illustrates the impact of wealth inequality and homophily (h) on group achievement (η_G) as a function of risk (r). It show that wealth inequality enhances group success when homophily is low ($h = 0$), as seen in the blue curve. In this scenario, cross-group influence allows contributions from the rich to compensate for the poor, fostering cooperation and achieving higher η_G for all values of r.

In contrast, when homophily is high ($h = 1$), represented by the red curve, η_G is reduced greatly. This shows that wealth inequality, combined with class division, discourages cooperation.

The gray curve, which represents the absence of wealth inequality, performs at an "intermediate" success level, highlighting that cross-population influence is critical for overall cooperation.

At a more zoom-out perspective, η_G increases with risk (r) across all scenarios, with the most substantial seen observed in the low homophily case ($h = 0$).

5.2 Stationary Distribution and Gradient of Selection under various values of risk (r) and homophily (h)

Stationary distribution and gradient of selection for different values of risk r and of the homophily parameter h. (A–F) Each panel contains all possible configurations of the population (in total $Z_R \times Z_P$), each specified by the number of rich (i_R) and poor (i_P) it contains, represented by a gray-colored dot. Darker dots represent those configurations in which the population spends more time, thus providing a contour representation of the stationary distribution. The curved arrows show the so-called gradient of selection (∇), which provides the most likely direction of evolution from a given configuration. We use a color code in which red lines are associated with higher speeds of transitions.

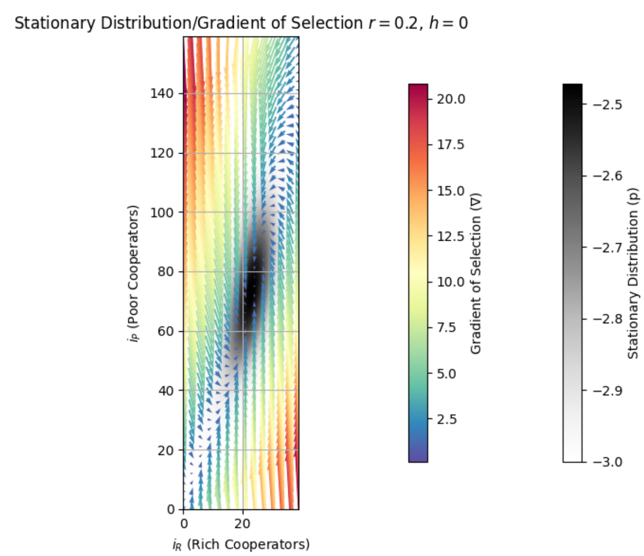

Figure B: Stationary distribution and gradient of selection for $r = 0.2$ and $h = 0$.

Our observations. The stationary distribution and gradient of selection analysis (Figure B-G) provide interesting insights into the dynamics of cooperation under wealth inequality, risk perception, and homophily, all of which vary in these plots.

[10]It took almost 30 minutes to get all the values of this graph with our current hardware.

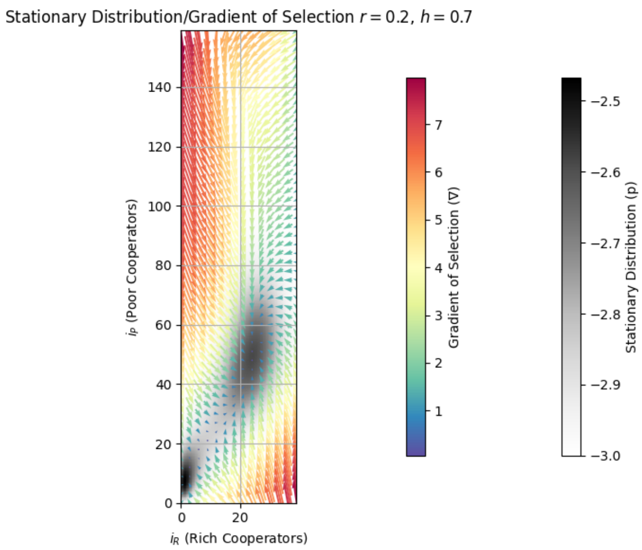

Figure C: Stationary distribution and gradient of selection for $r = 0.2$ and $h = 0.7$.

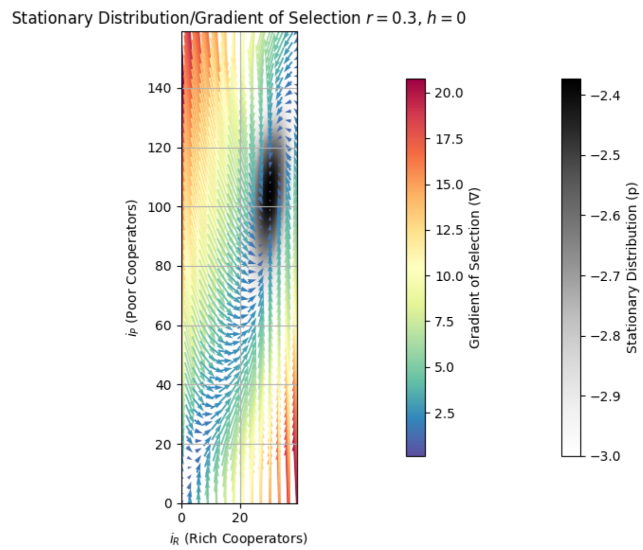

Figure E: Stationary distribution and gradient of selection for $r = 0.3$ and $h = 0$.

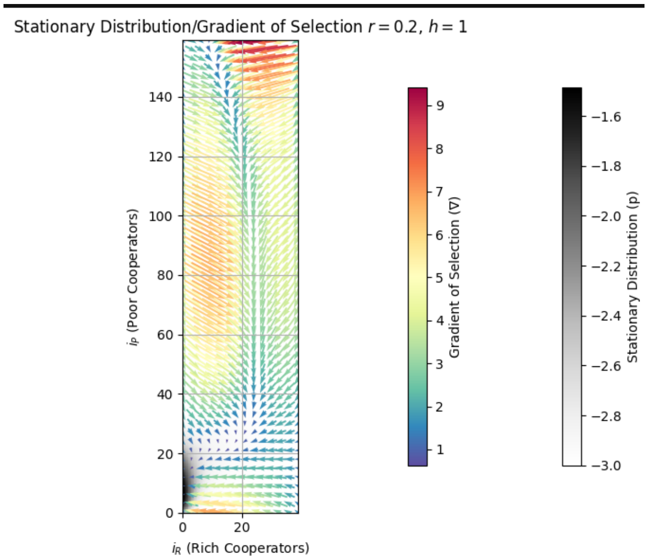

Figure D: Stationary distribution and gradient of selection for $r = 0.2$ and $h = 1$.

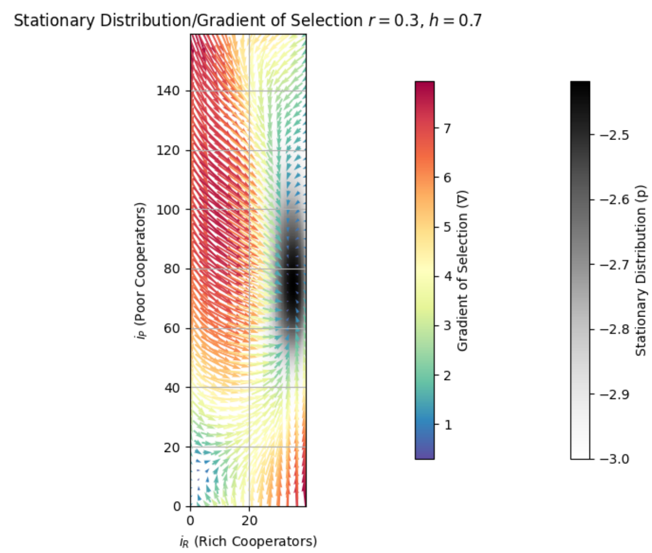

Figure F: Stationary distribution and gradient of selection for $r = 0.3$ and $h = 0.7$.

Our results replicate the results of the authors almost exactly, and they show some interesting patterns about what drives cooperation in populations with different levels of risk perception and homophily. First, higher risk perception ($r = 0.3$) clearly encourages both the rich and poor to contribute more. Even a small increase of risk is able to foster much greater cooperation.

One key comparison is from the graphs of $r = 0.2$, $h = 0.7$ versus $r = 0.3$, $h = 0.7$: here homophily remains the same, but a small increase in risk (just 10%) pushes both subpopulations to cooperate much more, we see that nearly 50% of P and almost all of R already prefer C as their strategy.

On the other hand, high homophily ($h = 1$), creates barriers to cooperation. As homophily increases, people interact less across groups, and this breaks the feedback loop that fosters mutual support. When homophily is very high, cooperation among the poor almost completely falls apart, and the contributions of the rich alone are not enough to sustain group success. This will be seen later with the average group achievement. This highlights how

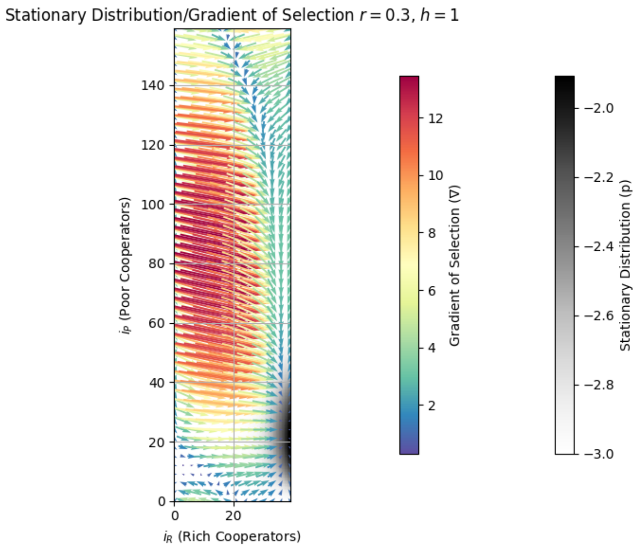

Figure G: Stationary distribution and gradient of selection for $r = 0.3$ and $h = 1$.

important it is to encourage interactions between different groups to bring people together and align their goals!

Interestingly, under higher risk, the rich step up more than the poor to make up for the poor's lack of cooperation, showing that collective success often depends heavily on wealthier participants. This is something the authors of the paper also mentioned[1]. This dynamic mirrors real-world situations, like climate change, where wealthier countries are expected to take on more responsibility. However, this approach isn't always fair or sustainable unless everyone contributes in some way. [11]

Conversely, imitation of strategies across subpopulations also means more cooperation. When there's little homophily ($h = 0$), the rich and poor influence each other very positively. This back-and-forth creates a healthy cycle of cooperation.

These findings offer some useful ideas for real-world policies. For example, making risks feel more immediate can inspire action. This shows why it is important to make the apparent risk of climate change clear if no action is taken. A higher perceived risk already means more cooperation. It's also important to ensure responsibilities are shared fairly, so no single group feels overburdened.

These results are in line with the results that the authors found in their experiments.

5.3 Doomsday scenario: High risk, $r = 0.9$

While this specific scenario was not explored in the original paper, we chose to investigate it as an interesting edge case. Here, the risk of climate change is extremely high ($r = 0.9$), reflecting a world where catastrophic consequences are almost certain without collective action (the end is nigh!)

[11]We have recently seen many talks about implementing a Carbon Tax which hopes to encourage this cooperation from the rich world, among other policies!

5.3.1 Low homophily, $h = 0.1$. First, the homophily parameter is kept low ($h = 0.1$), representing a hopeful scenario that countries in this situation will be eager to put aside differences. This combination is one of the outcomes we wanted to test to understand how populations would behave under extreme risk and minimal social segregation. A true doomsday scenario!

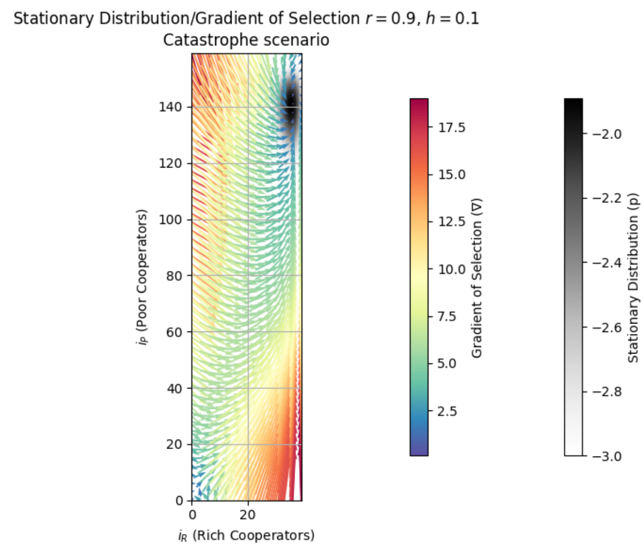

Figure H: Stationary distribution and gradient of selection for $r = 0.9$ and $h = 0.1$.

Interestingly, instead of descending into chaos or defection as one might anticipate under extreme risk, the model predicts a significant increase in cooperation. A striking result emerges: individuals from the Poor subpopulation show a higher preference for cooperation (C). This suggests that under overwhelming existential threats, even poor individuals are motivated to contribute to the public good, especially when homophily is small.[6]

These findings show a hopeful perspective! They indicate that extreme risk and open social structures can foster a collective effort to mitigate our climate crisis. This underscores the importance of reducing segregation and promoting cross-group influence in addressing global challenges like climate change. Yet more food for thought!

5.3.2 High homophily, $h = 0.9$. We also looked at a high risk and high homophily scenario, which is slightly more concerning. Here we can see that while the number of Cooperators in the Rich subpopulation is high, the Cs in the poor subpopulation varies greatly on the starting configuration. This paints a troubling "doomsday" scenario where social divides limit mutual influence and cooperation.

This highlights the fragility of cooperative dynamics under social segregation! High homophily isolates the Poor from the motivational influence of the Rich, leading to unequal contributions and reduced collective resilience. These results highlight the importance of fostering cross-group interactions and equitable resource sharing to ensure inclusive and sustainable cooperation under extreme risk.

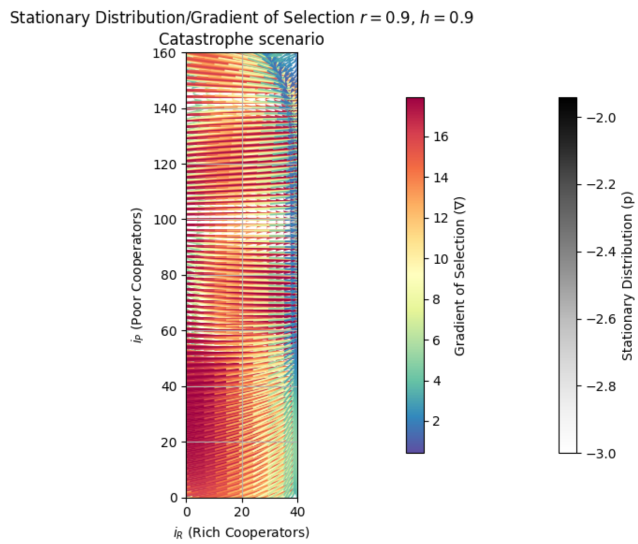

Figure I: Stationary distribution and gradient of selection for $r = 0.9$ and $h = 0.9$.

One can only hope that, in such a high-risk scenario, our world would choose to be the inclusive system of cooperation seen in the first scenario rather than the divisive and fragile dynamics of the second!

5.4 Timescale Separation: Games Among the Rich and Among the Poor

To assess what games the rich play in the presence of the poor (Cooperators (Cs) and Defectors (Ds)) and the poor play in the presence of the rich (Cs and Ds), we let each subpopulation evolve assuming that the rate of evolution of the other subpopulation is zero.

The results are shown in Fig. J-M, where we compute the gradient of selection (∇) that governs the evolutionary dynamics of the rich in the presence of frozen, mixed configurations of the poor (Fig. J and K) and vice versa (Fig. L and M). We consider the cases in which the population is subdivided into subpopulations of equal size ($Z_P = Z_R$, Fig. S1 A and C) or not ($Z_P = 4Z_R$, as in the main text, Fig. S1 B and D). $r = 0.3$ for these figures.

Figures J-M highlight the interplay between cooperation dynamics of the rich and poor under different configurations.

It's clear that the rich become more cooperative as the wealth gap increases, whereas cooperation among the poor remains largely unaffected by this disparity. The poor engage in a coexistence game[1], with overall cooperation decreasing as rich cooperation declines. Meanwhile, the rich participate in an N-player stag-hunt game[12], where their behavior depends on the relative size of the poor subpopulation and the fixed fraction of poor cooperators. These variations in parameters (fraction of C_P) can change the rich individuals' dynamics into either a coordination game, coexistence game, or defection-dominance dilemma[1].

[12]The rich can "hunt the stag" (high cooperation payoff) if there is enough cooperation among the poor; otherwise, they default to defecting!

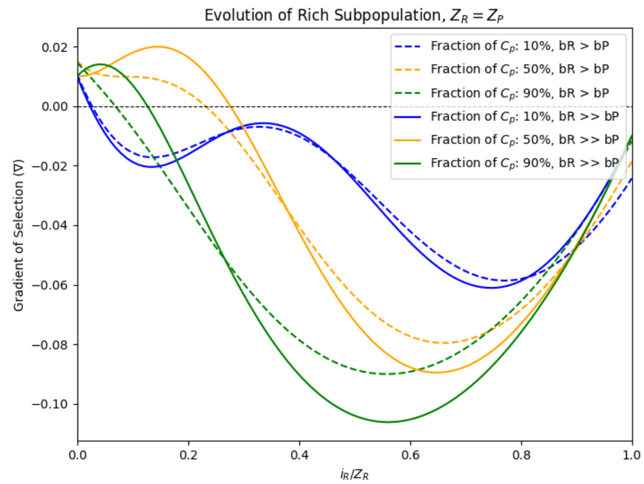

Figure J: Gradient of selection for R vs. i_R/Z_R

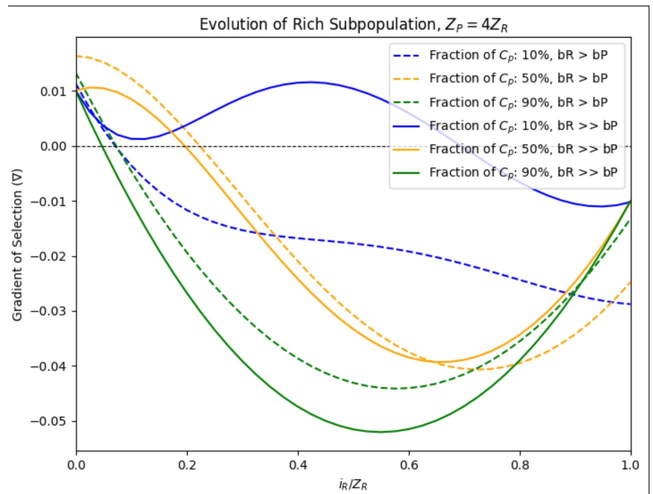

Figure K: Gradient of selection for R vs. i_R/Z_R

These results suggest that fostering cooperation among the poor can indirectly stabilize cooperation among the rich! So policies that increase the cooperation just for the poor, would have cascading effects on the rich, and the overall system and group dynamics. This implies that by encouraging smaller groups of cooperators in poor subpopulations we can foster global cooperation, including from the rich!

It is also noteworthy that the gradients of the rich are 10 times smaller than those of the poor. This means the rate of response of the rich to changes is (on average) slower than that of the poor. In practice, the poor will adjust their behavior more rapidly to changes in the configuration of the rich, thus quickly shifting between the corresponding levels of co-existence between poor Cs and Ds.[1]

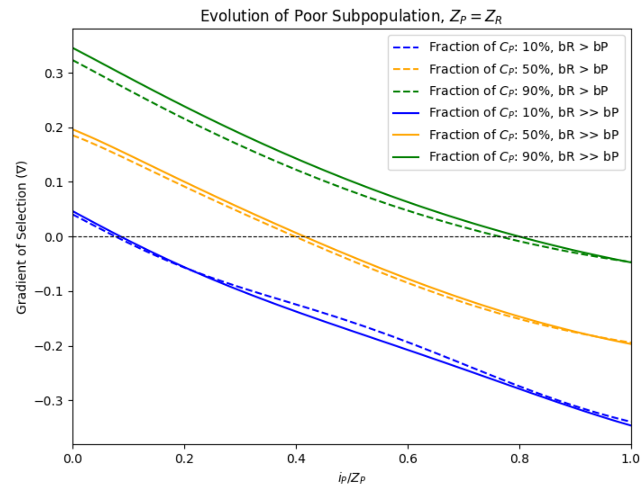

Figure L: Gradient of selection for R vs. i_R/Z_R

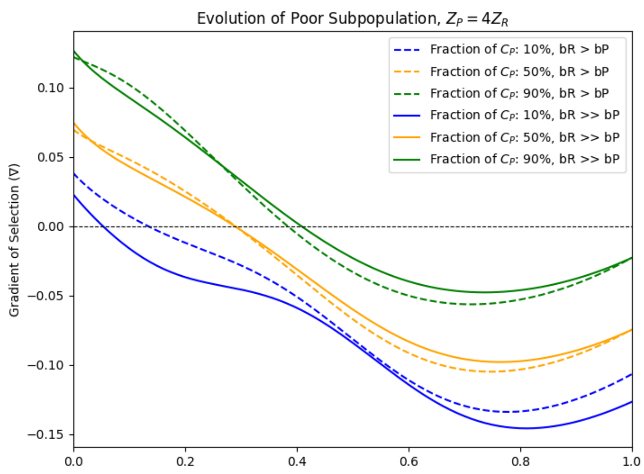

Figure M: Gradient of selection for R vs. i_R/Z_R

We have also analysed the behaviour of Rich and Poor subpopulations when $r = 0.9$ but due to space constraints will not delve into it here.[13]

5.5 Dynamics for equal fraction of rich and poor, $Z_R = Z_P$

Though we have replicated results for varying values of h and r in this scenario with equal rich and poor populations, due to space constraints, we will discuss only a couple of interesting cases.

For low risk, when $Z_R = Z_P$, the model suggests there are more pessimistic prospects regarding overall cooperation[1]. When we compare this to asymmetric subpopulations $4Z_R = Z_P$, we do not achieve the same levels of cooperation. Particularly, the rich cooperate much less and do not compensate for the lack of cooperation from the poor. In these cases, b_P and b_R are closer together because

[13]Jupyter notebook for the reproduction of this model can be found on GitHub (link).

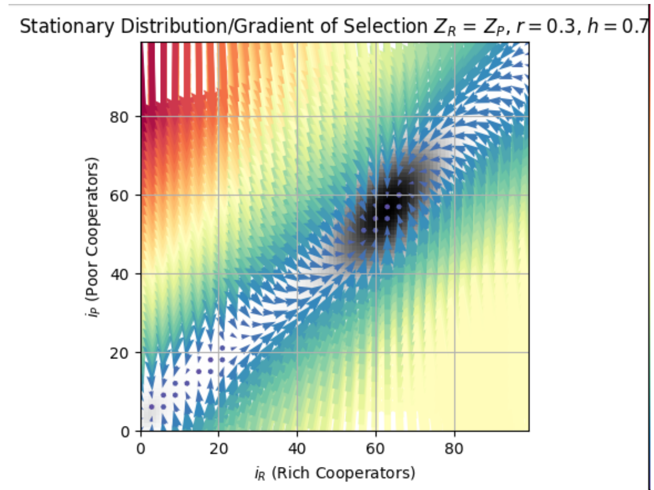

Figure N: Stationary Distribution/Gradient of Selection, $Z_R = Z_P$, $r = 0.3$, $h = 0.7$

we still want to maintain that $\frac{b_R Z_R + b_P Z_P}{Z} = \bar{b} = 1$, so this may impact the fitness of C_R in a detrimental way. So again, paradoxically, to effectively reach consensus to mitigate climate change, we need a more sizeable poor population to have higher levels of global cooperation!

There are other such comparisons with the case of asymmetric subpopulations, they are available in the *Jupyter Notebook*[14] for this report.

6 DISCUSSION

The results from the original paper highlight the impact of wealth inequality on collective action in climate change mitigation. The paper highlights that wealth inequality, paradoxically, can promote cooperation (when homophily is low) as wealthier participants compensate for the contributions of less wealthy ones. This outcome, for us, challenges traditional assumptions that inequality inherently undermines cooperation. However, the results also show that high homophily leads to a collapse in contributions from the poor, which in turn destabilises overall cooperation.

Although we did not replicate this part of the model in our reproduction, obstinate behaviour[15] has a sizeable impact on increasing cooperation. This is seen especially with obstinate poor cooperators. We also see another issue: when the contribution of the poor is widespread, the rich refrain from contributing!

Despite this, the main conclusions drawn from the experiments are, in general, that as long as risk perception is high, we partition countries into small groups agreeing on local, short-term targets, and where cooperation is fostered by imitating more successful peers, the prospects to face the looming climate crisis are not that grim![1]

[14]Jupyter notebook for the reproduction of this model can be found on GitHub (link).
[15]Obstinate behaviour means that a certain part of the population does not change their behaviour regardless of their groups/peer's fitness

Assumptions made by the paper

The paper makes some assumptions about success. We have calculated η_G as $Mc\bar{b}$, where $M = 3$, which implies we need just 3 cooperators (on average!) in a group to be "successful". This is a very loose definition of the word.

We noticed this while implementing the η_G (Figure A). We could easily overestimate group achievement, and so, by focusing solely on meeting this loosely defined numerical threshold without incorporating qualitative or some kind of long-term cooperation criterion, it's hard to say how the results shown here replicate the real world. The idea with game theory models of course is to simplify certain aspects of our world and be able to quantitatively study the effects of varying parameters, so this assumption is very fair, but one does wonder how accurate each such "simplistic" model imitates our inherently complex world. Indeed, one could even argue that there is some nuance that many such models miss even though we might look at them in relation to one another. Though this is a philosophical tangent we will refrain from engaging in in this report!

Future work

It would be interesting (and perhaps valuable) to extend and refine the model in several directions. Firstly, we could explore more nuanced forms of wealth inequality, moving beyond a simple rich/poor system to include a more continuous distribution of wealth and the possibility of wealth changing over time. Though we don't truly evolve the model in this paper, it might be useful to do that even in the current configuration as well to see if noteworthy patterns emerge (even though the stationary distribution gives us a fair idea as to what states the model would settle into).

With a model with continuous distribution of wealth and simple exchanges of payments, we would also be able to study the "returns" of investments in climate change mitigation. In this model, we essentially lose the endowment contributed towards the PGG, but in reality investments in green energy are seldom black hole investments - delayed returns and their potential amplification through economies of scale would provide a valuable area for future research.

Another interesting area might be enhancing the social structure underlying the total population in our PGG. In this model, we take homophily as the only such social constraint, but there are other factors that could be used to change group dynamics. Extending the model to include more complex network, perhaps representing world alliances (such as trade, cultural, or political alliances) could offer deeper insights. Doing this, of course, is yet another challenge, drawing up political alliances formally requires inputs from other fields such as economics and political science!

In short, the next steps and future work could involve broadening the model's economic and social dimensions. This would be all with the aim of better understanding the conditions that foster collective solutions to global climate challenges and hopefully, tackle the imminent crisis effectively!

7 CONCLUSIONS

In this report, we explored the dynamics of climate policies under wealth inequality, focusing on the critical parameters of risk perception and homophily within the framework of a threshold public goods game. Our reproduction reveals insightful patterns, particularly the dual role of wealth inequality! While inequality may hinder cooperation in the presence of high homophily, it can paradoxically promote collective action when cross-group interactions are fostered, enabling rich individuals to compensate for the shortfalls of poor ones.

The findings suggest a "cautious" yet optimistic outlook for addressing climate challenges, provided certain conditions are met. First, a high perception of risk is crucial to encourage collective action. Second, partitioning climate negotiations into smaller, localized groups with achievable short-term targets can foster more effective collaboration, as it simplifies coordination. Lastly, success depends on fostering an environment where individuals imitate their more successful peers, regardless of wealth class (i.e., low homophily), and of course, some degree of error in strategy adoption allows for exploration and adaptation (μ), both in our model and in real life!

Experience reproducing the model and its results

Reproducing the model from the paper was a very rewarding experience! It was quite exciting to dive into concepts like transition matrices and stationary distributions. We had touched on these briefly in class, but actually implementing them gave us a much deeper understanding. Figuring out details, like how to convert between scalar values and 2D tuples, wasn't straightforward at first, but it was a crucial piece of the puzzle and a great learning moment.

Of course, there were a few bumps along the way: some parts of the paper and its details, as well as other references we checked, were a bit ambiguous. It would be helpful if researchers made their code more accessible. It could save time for others trying to reproduce their results and might also make peer review and validation a lot easier.

Overall, this project gave us a real sense of what it takes to produce research of this quality. Seeing our results match up (almost fully!) with the original paper was incredibly satisfying. One thing that really stuck with us was wondering how the authors arrived at some of the more intricate equations, especially for transition probabilities. It's easy to take them for granted when you see them written down, but there are so many choices that could have sent the model in a completely different direction. Reproducing this work gave us a real appreciation for both the complexity and the creativity involved in research like this!

REFERENCES

[1] J. M. P. Vítor V. Vasconcelos, Francisco C. Santos and S. A. Levin, “Climate policies under wealth inequality,” *PNAS*, 2014.
[2] F. C. Santos and J. M. Pacheco, “Risk of collective failure provides an escape from the tragedy of the commons,” *PNAS*, 2010.
[3] M. D. S. . J. M. P. Francisco C. Santos, “Social diversity promotes the emergence of cooperation in public goods games,” *Nature*, 2008.
[4] M. P. Cheng-Yi Xia, Sandro Meloni and Y. Moreno, “Dynamic instability of cooperation due to diverse activity patterns in evolutionary social dilemmas,” *EPL*, 2015.
[5] E. Fernández Domingos, “Egttools: Toolbox for evolutionary game theory,” https://github.com/Socrats/EGTTools, 2020.
[6] J. Sell and T. P. Love, “Common fate, crisis, and cooperation in social dilemmas,” *Advances in Group Processes*, 2009.