Common good before personal, an incentive impact on cooperative dilemmas for limited resources

Alexandre Chapelle, Nicolas Dardenne, Emma Dupuis and Simon Renard

Université Libre de Bruxelles
INFO-F409 - Learning Dynamics

Abstract

Collaboration between parties sometimes requires suitable incentives in long term situations. Society is currently facing several dilemmas that demand collective work instead of personal interest, such as climate change and more recently, the coronavirus pandemic. Our paper treats populations' behaviors regarding negative punishments and positive rewards as incentives. After analyzing populations under flexible and fixed motivations, it results that a local punishing scheme with absolute reward returns an optimal strategy.

Introduction

Since the dawn of humanity, humans have adapted and evolved in a hostile environment that is not conducive to their development. First problems were on a lower scale: finding food, being safe, and having a shelter. Since then, our tribes have turned into villages, our villages into cities, and now our cities have become societies that must make their way through time. Globalization, the exponential scientific progress and the lifestyle we have accommodated to lead to new modern dilemmas. Recently, the Sars-Cov-2 pandemic (also known simply as Covid-19) and the global warming are two of the most urgent and problematic dilemmas that societies have to face. Some of these issues cannot be solved by a few actors and therefore need the cooperation of parties. In the case of these dilemmas, both economic stakes and short-term rewards push us to be navel-gazing. These selfish actions are certainly rewarding in the short term for the entity who chooses them, but they have a very negative impact on others trying to choose the common good and even more for the defector entity itself in the long term.

Here comes the strategy of executors, where defectors are punished and cooperators rewarded. Our objective is to analyze the behavior of those populations according to the weights of rewards and punishments.

To illustrate this work, we will take a current problem affecting the entire population: the global warming. We will see how the executors (government, health agencies, ...) can influence cooperators and defectors (both are countries, populations, ...) through rewards and punishments to influence as many people as possible to get vaccinated.

To prevent a collective tragedy, the inevitable "tragedy of the commons" (Hardin, 1968) where a limited common resource is contested between actors, executors need to push both cooperators and defectors to see further than their self-interest.

Firstly, different cases with multiple parameters will be studied (fixed and flexible incentives, risk, mutation rate, ...). Furthermore, to push the thought further, a special case within a stochastic executors environment will also be analyzed. This corresponds to societies where the policy is not always perfectly executed (mistakes can be made). For example, a police officer can sometimes make mistakes by fining someone that respects every rule; how this person will react to this action?

Methods

The following results were obtained using a mixed population of Z individuals following the article Sun et al. (2021).

Public goods game

There are two types of individuals: Cooperators and defectors. From this population, N individuals were randomly chosen to take part in the game. Each player is given a capital b and has the ability to contribute a quantity c, knowing there is a minimum of M cooperators in the group $(0 < M \leq N)$ to reach the goal. If the goal contribution is achieved, all individuals will keep their endowments. Otherwise, there is a r probability $(0 \leq r \leq 1)$ to lose all their endowments.

The cooperators always contribute c to the group. Unlike them, defectors will be selfish and therefore keep their resources. The payoffs of defectors and cooperators are respectively Π'_D and Π'_C with j_C being the number of cooperators.

$$\begin{aligned}\Pi'_D(j_C) &= b\Theta(j_C - M) + (1-r)b(1-\Theta(j_C - M)) \\ \Pi'_C(j_C) &= \Pi'_D - c\end{aligned}$$

With:

$$\Theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{otherwise} \end{cases}$$

Collective-risk social dilemma game

A third type of population then enters the game, called executors, which contribute an amount c to the pool as cooperators, but also have the ability to contribute to a sanctioning pool. This pool can then be used to reward cooperators and punish defectors accordingly. If the number of executors reaches the threshold N_E, then a fraction α of the pool π will be used to reward cooperators while a fraction $1-\alpha$ will be used to punish defectors. α is the balance between rewarding and punishing incentives. In the special case where $\alpha = 0$, the executors will only punish while in the case where $\alpha = 1$, they will only reward the cooperators. Therefore, the payoffs of all three participants can be described as the following:

$$\begin{aligned} \Pi_D(j_C, j_E) &= \Pi'_D(j_C + j_E) - (1-\alpha)\pi_e\Delta \\ \Pi_C(j_C, j_E) &= \Pi'_C(j_C + j_E) + \alpha\pi_e\Delta \\ \Pi_E(j_C, j_E) &= \Pi'_C(j_C + j_E) + \alpha\pi_e\Delta - \pi_t \end{aligned}$$

Where j_C and j_E are respectively cooperators and executors. π_t and π_e are the cost and the sanctioning pool. These are fixed incentives. The Δ function depends on the scheme and indicates that the incentives are applied only if the number of executors reaches the threshold. For local schemes, it can be described as $\Delta = \Theta(j_E - n_E)$ while for global schemes, $\Delta = \Theta(i_E - n_E)$.

Flexible incentives depend on the amount of executors and will have effect on π_e. The reward pool is directly proportional to the number of executors. The enhancement factor δ gives the weight of the rewards and the punishments.

$$\begin{aligned} \Pi_D(j_C, j_E) &= \Pi'_D(j_C + j_E) \\ &\quad -(1-\alpha)\frac{\pi_t \delta j_E}{N - j_C - j_E}\Delta \\ \Pi_C(j_C, j_E) &= \Pi'_C(j_C + j_E) + \alpha\frac{\pi_t \delta j_E}{j_C + j_E}\Delta \\ \Pi_E(j_C, j_E) &= \Pi'_C(j_C + j_E) + \alpha\frac{\pi_t \delta j_E}{j_C + j_E}\Delta - \pi_t \end{aligned}$$

Average payoff

The average payoff of the three strategies in a specific population $i = (i_C, i_E, i_D)$ can be computed:

$$\begin{aligned} f_D(i) &= \sum_{j_C=0}^{N-1}\sum_{j_E=0}^{N-1-j_C} \frac{\binom{i_C}{j_C}\binom{i_E}{j_E}\binom{i_D-1}{j_D-1}}{\binom{Z-1}{N-1}}\Pi_D(j_C, j_E) \\ f_C(i) &= \sum_{j_C=0}^{N-1}\sum_{j_E=0}^{N-1-j_C} \frac{\binom{i_C-1}{j_C}\binom{i_E}{j_E}\binom{i_D}{j_D-1}}{\binom{Z-1}{N-1}}\Pi_C(j_C+1, j_E) \\ f_E(i) &= \sum_{j_C=0}^{N-1}\sum_{j_E=0}^{N-1-j_C} \frac{\binom{i_C}{j_C}\binom{i_E-1}{j_E}\binom{i_D}{j_D-1}}{\binom{Z-1}{N-1}}\Pi_E(j_C, j_E+1) \end{aligned}$$

At each evolution stage, random participants using the L strategy will adopt another participant's R strategy. This adoption has a probability $1/1[1 + e^{-\beta(f_R - f_L)}]$ of happening, β representing the intensity of selection. The closer $\beta \to 0$, the more random the drift will be while the more $\beta \to \infty$, the more the imitation process will prefer high payoff strategies. β was set to 5 in our case, which will generally make the best performing players strategies an example for others. Yet, this does not exclude that some players might adopt less efficient strategies. We also introduce μ which is the mutation rate allowing a participant to adopt a random different strategy with μ probability.The transition probability can be described as the following:

$$T_{L\to R} = (1-\mu)\frac{i_L}{Z}\frac{i_R}{Z-1}[1 + e^{-\beta(f_R - f_L)}]^{-1} + \mu\frac{i_L}{2Z}$$

where i_L and i_R represent the number of individuals using the strategies L and R respectively. The probability of gaining or losing one participant in the cooperating population is as follows:

$$T_i^{C\pm} = T_{(i_C\pm1, i_E\mp1, i_D)} + T_{(i_C\pm1, i_E, i_D\mp1)}$$

The same goes for executors.

$T_{ii'}$ is the transition probability from i' to i per τ unit time. With the matrix $T = [T_{ij}]^T$, we can obtain the stationary distribution $\bar{p_i}(t)$ by computing its eigenvector with the eigenvalue 1. This distribution gives information on the fraction of time the population spent in every configuration of the finite population.

In the aim of plotting the two-dimensional simplex, the gradient of selection ∇_i can be computed. This gradient represents the average evolution direction of the population from an initial population state i. ∇_i is computed as follows with u_C and u_E, two units vector:

$$\nabla_i = (T_i^{C+} - T_i^{C-})u_C + (T_i^{E+} - T_i^{E-})u_E$$

To get a better understanding of sanctioning policies, $a_G(i)$ computes the portion of groups who reached the collective goal threshold for a given population configuration i.This can be done through multivariate hypergeometric distribution:

$$a_G(i) = \binom{Z}{N}^{-1} \sum_{j_C=0}^{N} \sum_{j_E=0}^{N-j_C} \binom{i_C}{j_C}\binom{i_E}{j_E}\binom{Z-i_C-i_E}{N-j_C-j_E} \theta(j_C+j_E-M)$$

By using the latter, the average group achievement η_G is a mean of every $a_G(i)$ weighted by the stationary distribution. More precisely, $\eta_G = \sum_i \bar{p}_i a_G(i)$.

Similarly, another relevant metric is the portion of groups $a_I(i)$ that reach n_E executors for local or global institutions for a configuration i.

$$a_I(i) = \binom{Z}{N}^{-1} \sum_{j_C=0}^{N} \sum_{j_E=0}^{N-j_C} \binom{i_C}{j_C}\binom{i_E}{j_E}\binom{Z-i_C-i_E}{N-j_C-j_E} \theta(j_C+j_E-M)\Delta$$

We can also mean $a_I(i)$ over every population configuration i to get the average institution prevalence $\eta_I = \sum_i \bar{p}_i a_I(i)$.

Stochastic Executors

In order to push the reflection further, it is possible to consider the executors are stochastic. Each time an executor punishes or rewards a member of the group, there is an error rate ϵ for a cooperator to be punished and for a defector to be rewarded. This can be seen as an error done by the executor. The corresponding payoffs become the followings:

$$\begin{aligned}
\Pi_D(j_C, j_E) &= \Pi'_D\Theta(j_C+j_E) \\
&+ \Delta\pi_e\left(-(1-\epsilon)(1-\alpha)+\epsilon\alpha\right) \\
\Pi_C(j_C, j_E) &= \Pi'_C\Theta(j_C+j_E) \\
&+ \pi_e\Delta\left((1-\epsilon)\alpha-\epsilon(1-\alpha)\right) \\
\Pi_E(j_C, j_E) &= \Pi'_C\Theta(j_C+j_E) \\
&+ \pi_e\Delta\left((1-\epsilon)\alpha-\epsilon(1-\alpha)\right)-\pi_t
\end{aligned}$$

This formula allows the possibility for the executors to make mistakes.

Results

The code for all results is available on github (Dupuis et al., 2021).

In the following results, we compute the matrix $T_{ii'}$ and the corresponding stationary distribution for different parameters and strategies. We compute the average group achievement in order to compare the efficiency of these parameters. We also compute the gradient of selection to plot the two-dimensional simplex that shows the evolution of a population.

Fixed incentive

The following results were obtained using fixed incentives.

In Figure 1, where a pure reward strategy is executed (so when $\alpha = 1$), the high amount of arrows flowing to the middle edge of $C - E$ suggests that the population evolves to a cooperation dominant state. We can also observe that the defector state is almost empty.

In Figure 2, where a pure punishment strategy is executed, the populations will spend a lot of time in configurations where many individuals are defectors. The best outcomes are achieved with the pure reward strategy. Indeed, the population is far from the D vertex (defectors). As displayed below, the two simplex, η_G, which is the average group achievement, is much higher in the pure reward strategy (99.87%) than in the pure punishment strategy (56.56%).

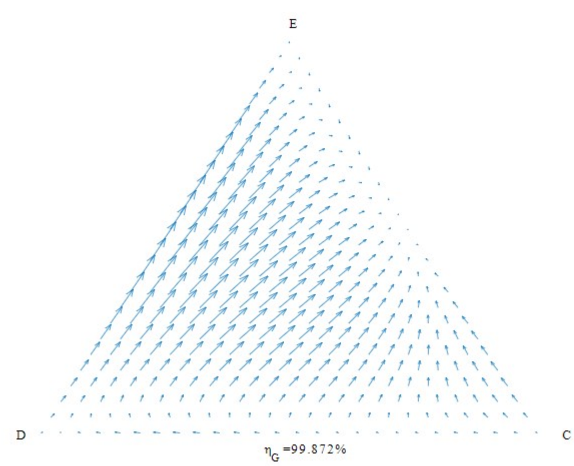

Figure 1: Reward ($\alpha = 1$)

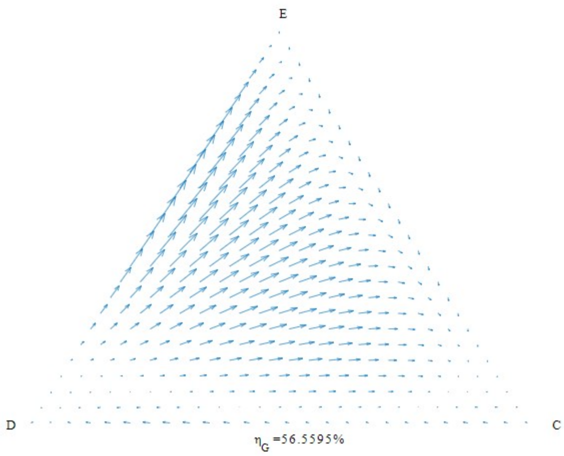

Figure 2: Punishment ($\alpha = 0$)

These two figures let us understand that both cooperators and defectors would be more eager to cooperate if executors use reward strategy instead of punishment strategy.

The previous two simplex were computed using the following parameters: A local scheme is triggered if the amount of executors is more than 25% of the population. $r = 0.2$, $M = 75\%$ of group size. $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, $\pi_e = 0.3$

Figure 3 shows the average group achievement η_G according to α for different values of r. We can observe that the average group achievement grows fast, then slowly stabilizes. This simulation was run with $c/b = 0.1$ but these curves are still valid for other ratios. We can therefore affirm that pure reward strategies perform better than other strategies.

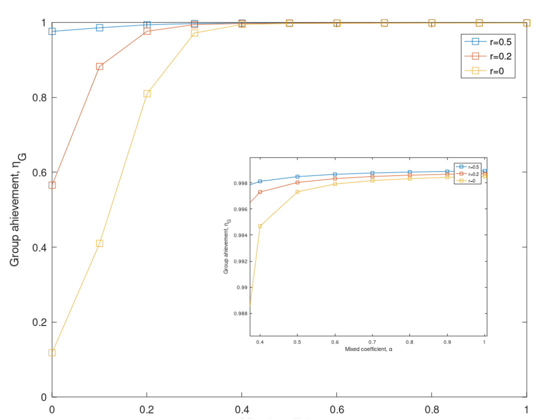

Figure 3: Group achievement depending on the mixed coefficient

In Figure 4, we show η_G as a function r for values of α. We can observe that the collective target is reached faster as r is increased.

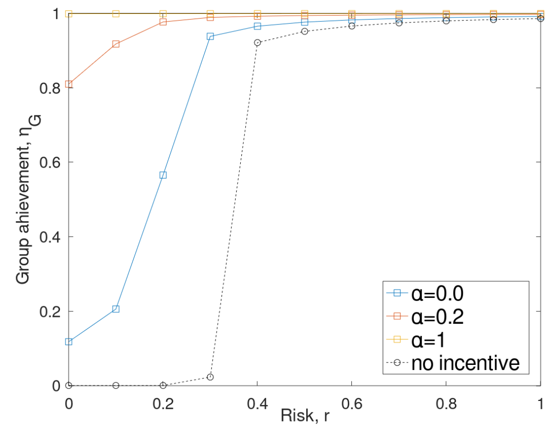

Figure 4: Group achievement depending on risk

The mutation rate, as mentioned before, is an important factor that can alter the strategies adopted by a player. As seen on Figure 5, at the very beginning of the mutation rate growth, the group achievement is at its highest. The more the mutation rate increases, the more the group achievement tends to reach a specific value that does not seem to fluctuate with α. Therefore, our results are not very impacted by the mutation rate. The best curve is when $\alpha = 1$, which confirms our previous results.

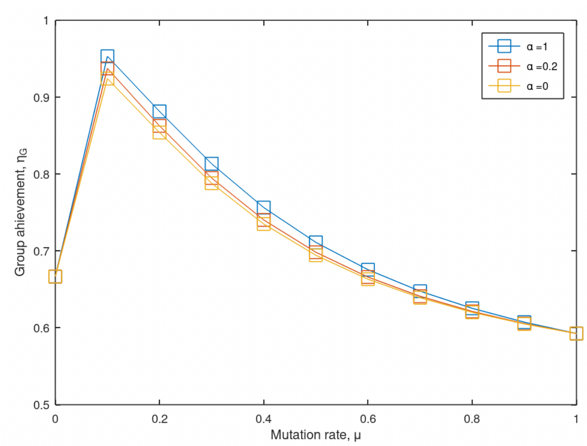

Figure 5: Group achievement depending on mutation rate

Let's now have a look at the institution efficiency regarding the success of the pure reward strategy. As shown in Figure 6, the institution prevalence growths as α is increased.

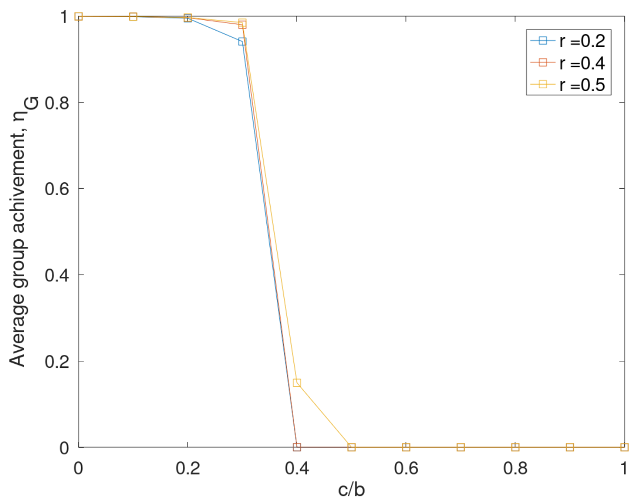

Figure 7: Group achievement depending on $\frac{c}{b}$ ratio

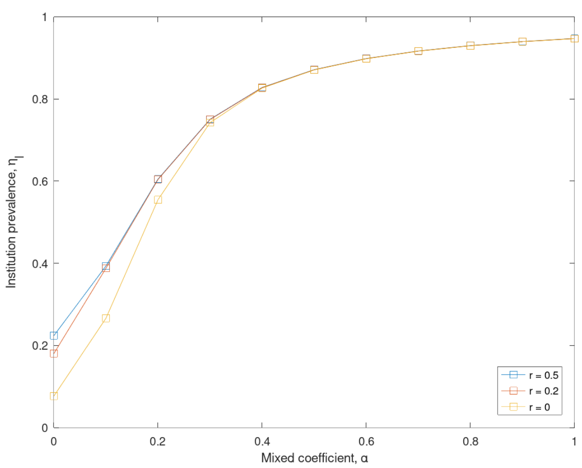

Figure 6: Institution prevalence depending on the mixed coefficient α

Figure 7 illustrates the average group achievement in terms of the ratio $\frac{c}{b}$. We observe that the bigger the portion a cooperator must invest in the common pool, the less the group is likely to cooperate. In particular, when the ratio $\frac{c}{b} \geq 0.5$, the average group achievement becomes almost null.

The previous graphs were plotted using these parameters: $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$ and $\pi_e = 0.3$ in Figures 3, 4, 6 and 7. $r = 0.3$ in Figure 5.

Flexible incentives

Another way to apply incentives is to make their value flexible according to the number of executors players. For that purpose, we introduce δ, an enhancement factor that will modify the reward and punishment values accordingly. In Figures 8 to 11, we can observe the gradient of selection according to multiple enhancement factor values. We observe that, whatever the value of δ is, the average group achievement is higher when a pure reward ($\alpha = 1$) is applied. A higher value of δ increases this group achievement. The arrows that go toward the vertex D are smaller when δ is bigger. Thus, the distribution of the population around this vertex is lower. As there are fewer defectors, the group achievement is higher.

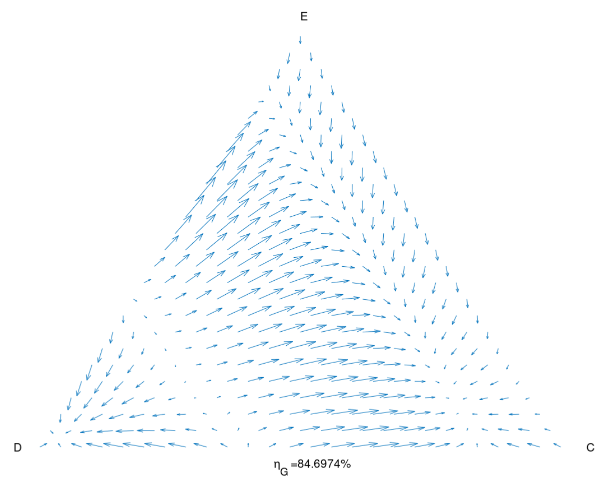

Figure 8: $\alpha = 1$, $\delta = 1.4$

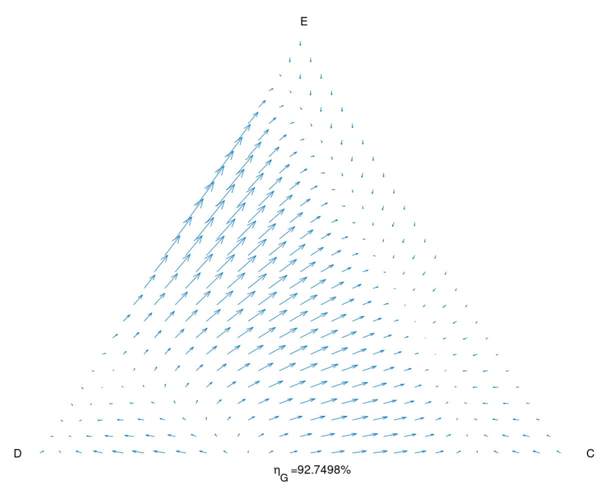

Figure 9: $\alpha = 1$, $\delta = 3$

Local and global scheme

The Figure 12 shows the difference between global and local schemes. The parameters used are the same as those for

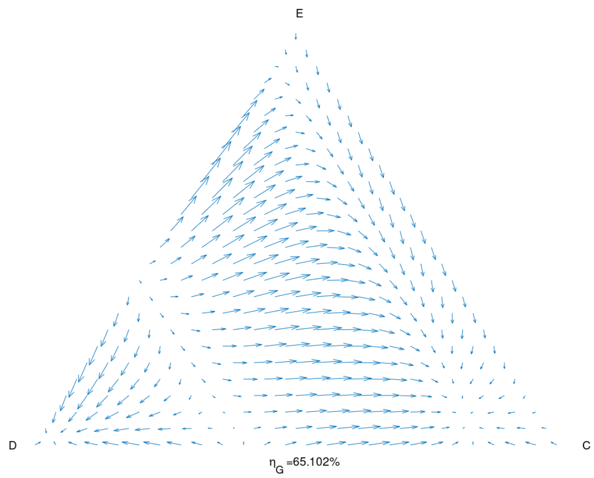

Figure 10: $\alpha = 0$, $\delta = 1.4$

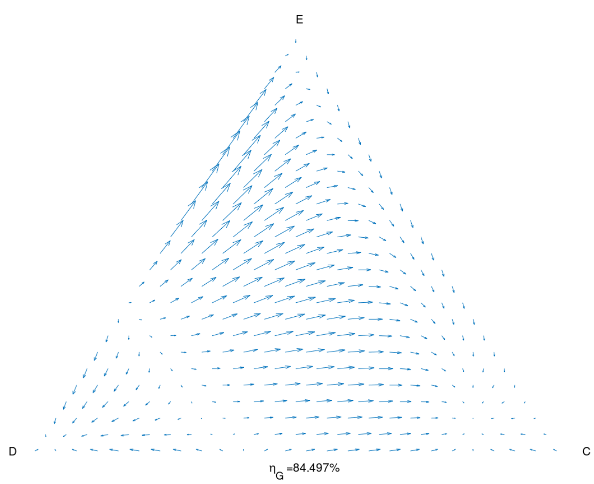

Figure 11: $\alpha = 0$, $\delta = 3$

Figure 1. We observe that when the global scheme is applied, the average group achievement is constant whatever the mixed coefficient α is. When the local scheme is applied, as we already know, the average group achievement grows initially fast with the mixed coefficient α, then its growth slow down. We see that when $\alpha = 0.4$ the local scheme overtakes the global one. Thus, the best results are obtained with the local scheme with a pure reward strategy ($\alpha = 1$).

Stochastic executors

In this section, we took the best results we had analyzed before. Therefore, we only will work in local schemes to directly provide the best and most interesting results. Finally, fixed incentive has been used to provide constant results.

Figure 13 highlights the evolution of the average group achievement η_G depending on the error rate ϵ for different values of α. We observe that whatever the value of α, when the error rate reaches a certain value, the average group achievement drops to nil. However, once again, the pure reward strategy ($\alpha = 1$) is the most robust to the increase of the error rate.

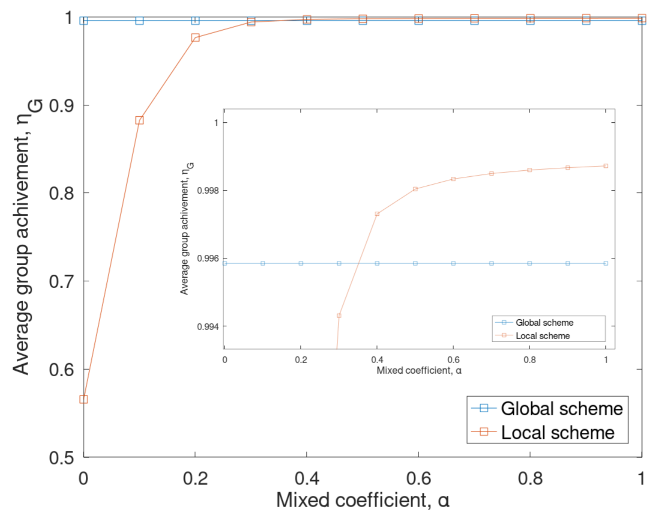

Figure 12: Average group achievement depending on the mixed coefficient for local and global scheme

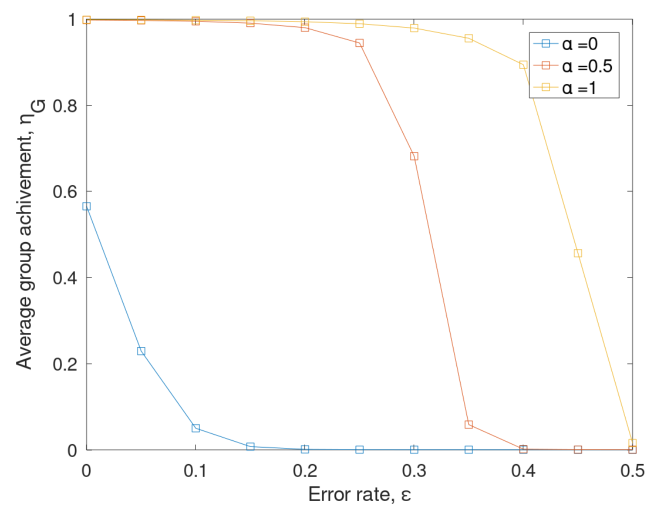

Figure 13: Group achievement depending on error rate ϵ

Discussion

We obtained the same results as the original paper (Sun et al., 2021).

In summary, for both fixed and flexible incentives, the highest group achievement is reached when executors use the pure reward strategy. Defectors are punished and cooperators rewarded with an $\alpha = 1$. The fluctuation of the risk r, the mutation rate μ or the enhancement factor δ do not have an impact on the ranking of the best strategy, which is pure reward. They only influence the time to achieve better results. Furthermore, we show the robustness of the pure reward strategy against the error make by the executors. Finally, we highlight that using a local scheme is more efficient than using a global one.

Humankind dilemmas such as global warming require the cooperation of a sufficient amount of people. That is unlikely to happen without rewarding incentives due to short-term profit and selfishness.

In the case of the coronavirus pandemic, group achievement can be seen as recovering the way people used to live before the pandemic stroke (open cinemas, music festivals…). Punishment can be seen as not allowing non-vaccinated people to take part in activities (flights, eating at a restaurant…).

The mathematical model has proven to be working well. Nonetheless, it is hardly applicable in real life as humans have emotions, unpredictable events can happen, and so on. Furthermore, a pure reward strategy would be hard to implement due to economic reasons. We cannot reward indefinitely. Therefore, a balance between reward and punishment would be needed and would change according to the environment. The environment changing in real time, so would the incentives. The stochastic environment could be compared to human mistakes, corruption, etc.

References

Dupuis, E., Chapelle, A., Dardenne, N., and Renard, S. (2021). Code to compute institutional incentives for cooperative governance of riskycommons and related figures. `https://github.com/emmdupui/LearningDynamicsGroup11`.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.

Sun, W., Liu, L., Chen, X., Szolnoki, A., and Vasconcelos, V. V. (2021). Combination of institutional incentives for cooperative governance of risky commons. *iScience*, 24(8):102844.