



UvA-DARE (Digital Academic Repository)

Combination of institutional incentives for cooperative governance of risky commons

Sun, W.; Liu, L.; Chen, X.; Szolnoki, A.; Vasconcelos, V.V.

DOI

[10.1016/j.isci.2021.102844](https://doi.org/10.1016/j.isci.2021.102844)

Publication date

2021

Document Version

Final published version

Published in

iScience

License

CC BY-NC-ND

[Link to publication](#)

Citation for published version (APA):

Sun, W., Liu, L., Chen, X., Szolnoki, A., & Vasconcelos, V. V. (2021). Combination of institutional incentives for cooperative governance of risky commons. *iScience*, 24(8), Article 102844. <https://doi.org/10.1016/j.isci.2021.102844>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

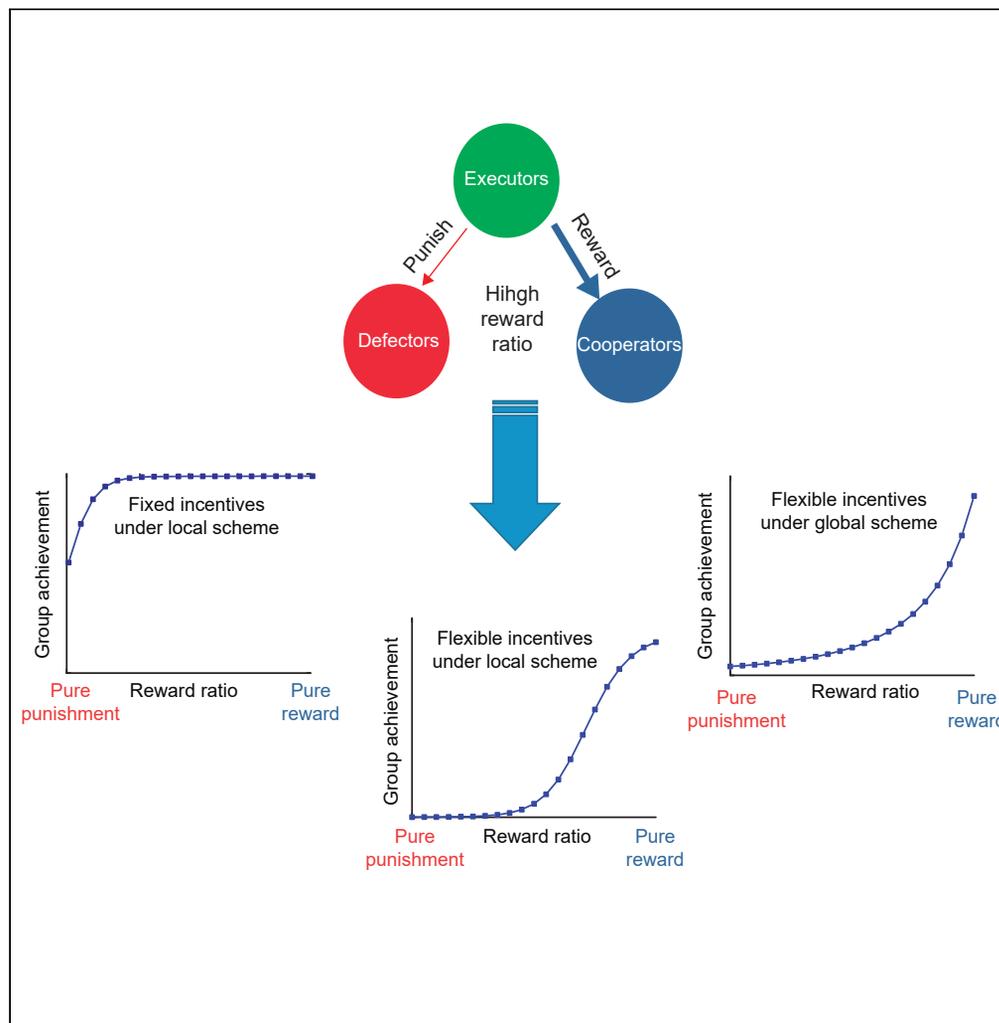
Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Article

Combination of institutional incentives for cooperative governance of risky commons



Weiwei Sun, Linjie Liu, Xiaojie Chen, Attila Szolnoki, Vítor V. Vasconcelos

xiaojiechen@uestc.edu.cn

Highlights

Pure reward in a local scheme is more effective both for fixed and flexible incentives

It can drive the entire population toward a highly cooperative state

Increasing the efficiency of the institution can induce the success of pure reward

A local scheme promotes group success more effectively than a global scheme

Sun et al., iScience 24, 102844
August 20, 2021 © 2021 The Author(s).
<https://doi.org/10.1016/j.isci.2021.102844>



Article

Combination of institutional incentives for cooperative governance of risky commons

Weiwei Sun,^{1,5} Linjie Liu,^{1,5} Xiaojie Chen,^{1,6,*} Attila Szolnoki,² and Vitor V. Vasconcelos^{3,4}

SUMMARY

Finding appropriate incentives to enforce collaborative efforts for governing the commons in risky situations is a long-lasting challenge. Previous works have demonstrated that both punishing free-riders and rewarding cooperators could be potential tools to reach this goal. Despite weak theoretical foundations, policy makers frequently impose a punishment-reward combination. Here, we consider the emergence of positive and negative incentives and analyze their simultaneous impact on sustaining risky commons. Importantly, we consider institutions with fixed and flexible incentives. We find that a local sanctioning scheme with pure reward is the optimal incentive strategy. It can drive the entire population toward a highly cooperative state in a broad range of parameters, independently of the type of institutions. We show that our finding is also valid for flexible incentives in the global sanctioning scheme, although the local arrangement works more effectively.

INTRODUCTION

Our society faces many urgent challenges that can be encapsulated as “the problems of the commons”, such as climate change, desertification of land, environmental pollution, and the list can be easily extended by other examples (Ostrom, 1990). All these pressing problems require us to give up our short-term personal interests for the benefit of long-term collective success (Rand and Nowak, 2013; Tavoni, 2013; Lenton, 2014; Shirado and Christakis, 2020). However, self-interested individuals always prefer their personal interests which can easily lead to a collective failure; hence, the “tragedy of the commons” seems to be inevitable (Hardin, 1968).

As a paradigm, the public goods game has been commonly used to study such dilemmas characterized by the frustration between personal and collective interests (Fehr and Gächter, 2002; Hauert et al., 2002; Szolnoki et al., 2011; Sasaki et al., 2012; Chen et al., 2015; Han et al., 2015; Wang et al., 2019; Ginsberg and Fu, 2019; Domingos et al., 2020; Santos et al., 2021). However, recent research emphasized that the traditional public goods game does not consider the risk of group failure; hence, the so-called collective-risk social dilemma game is a more competent tool to grab the essence of the conflict in several cases (Milinski et al., 2008). An excellent example of this feature is the mitigation effort to prevent climate change (Schroeder et al., 2012; Barrett and Dannenberg, 2012; Marotzke et al., 2020), which is one of the greatest public goods dilemmas that we humans face in the history (Inman, 2009).

Recent experimental and theoretical studies concerning collective-risk social dilemmas have revealed that the risk of collective failure plays an important role in promoting the evolution of cooperation (Milinski et al., 2008, 2011; Wang et al., 2009, Wang et al., 2010; Santos and Pacheco, 2011; Jacquet et al., 2013; Wu et al., 2013; Zhang et al., 2013). More precisely, a high risk can significantly enhance the willingness to cooperate, but cooperation collapses when the risk of failure is low (Santos et al., 2012; Chen et al., 2012, 2014; Pacheco et al., 2014; Vasconcelos et al., 2014, 2015; Diaz and Moore, 2017). Introducing costly punishment into the collective-risk social dilemma game can reverse the disadvantage situation under low risk, even when punishment is voluntary and costly (Vasconcelos et al., 2013). Furthermore, local schemes of sanctioning can do better than global schemes of institutions in improving the general group achievement, even when the risk of collective failure is low (Vasconcelos et al., 2013).

Beside punishment (Sigmund et al., 2010), other incentive strategies, like rewarding cooperators or the simultaneous presence of independent incentives, were also considered by previous works (Hilbe and

¹School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

²Institute of Technical Physics and Materials Science, Centre for Energy Research, P.O. Box 49, Budapest 1525, Hungary

³Informatics Institute, University of Amsterdam, Science Park 904, 1098XH Amsterdam, The Netherlands

⁴Institute for Advanced Study, University of Amsterdam, 1012 GC Amsterdam, The Netherlands

⁵These authors contributed equally

⁶Lead contact

*Correspondence: xiaojiechen@uestc.edu.cn
<https://doi.org/10.1016/j.isci.2021.102844>



(Sigmund, 2010; Szolnoki and Perc, 2013). An intriguing question is whether there exists the optimal relation of negative and positive incentives that can be maintained when voluntary and costly. In this paper, we thereby consider a hybrid enforcement strategy which imposes both kinds of incentives on participants simultaneously in the collective-risk social dilemma game. Accordingly, the positive and negative incentives are simultaneously presented within an incentive strategy, called executor strategy, and executors both punish defectors and simultaneously reward cooperators with a certain weight.

Our principal goal is to investigate how the proposed executor strategy influences the group achievement for solving the collective-risk social dilemma and to further find the optimal incentive strategy in the cooperative governance of risky commons when institutions are working at a local or global level. Furthermore, we respectively consider two cases of institutions: when the imposed incentive is fixed, the incentive amount is independent of the actual number of executor players; while when the imposed incentive is flexible, the incentive amount depends on the actual number of executor players (Sigmund et al., 2010; Sasaki and Unemi, 2011; Vasconcelos et al., 2013; Szolnoki and Perc, 2015). We stress that in our framework whether the institutions can be formed depends on the number of executor players in game interactions, and hence, this setting is different from previous works (Chen et al., 2015; Góis et al., 2019) which supposed that the considered centralized institutions are established in advance and can always work stably. In addition, the incentive budget on cooperators and defectors is not constrained in this work, whereas it is fixed in previous works (Chen et al., 2015; Góis et al., 2019). Interestingly, we show that independently of the types of institutions, pure reward strategy can always do better than other combinations of incentives in improving the level of group achievement in local sanctioning schemes. Besides, we find that this observation remains valid for flexible incentives in global sanctioning schemes. Furthermore, we reveal that regardless of the value of risk, a local scheme can promote group success more effectively than a global scheme, no matter whether the imposed incentives are fixed or flexible.

RESULTS

Fixed incentives under a local scheme

We investigate the stationary distribution and the gradient of selection to study the evolutionary dynamics of cooperators (C), defectors (D), and executors (E) in finite well-mixed populations (Vasconcelos et al., 2013). The stationary distribution standing for the prevalence in time of each configuration of the entire population is mapped onto the triangular simplex, in which each dot represents a configuration. The so-called gradient of selection provides the most likely direction of evolution from a given configuration.

In Figure 1, we show the stationary distribution and the gradient of selection when pure reward strategy (i.e., $\alpha = 1$) and pure punishment strategy (i.e., $\alpha = 0$) are considered, respectively. It suggests that the population will spend a significant time near the CE (C means cooperators and E means executors)-edge of the triangle simplex, as shown in Figure 1A. Furthermore, most of the arrows in the simplex flow to the intermediate region of CE -edge, indicating that the entire population evolves toward cooperation dominant states. In Figure 1B, we show that the population will spend most of the time near the configurations in which a lot of individuals are defectors (Vasconcelos et al., 2013). The comparison of these panels suggests that pure reward strategy can lead to better outcomes than pure punishment strategy since the population spends most of the time away from the D vertex in the usage of pure reward strategy. In order to qualify this, we further compute the average group achievement η_G values, which denote the average fraction of groups that succeed in achieving collective targets to measure the level of cooperation. We find that when pure reward strategy is considered, the average group achievement value is 99.87%. This value is much higher than the average group achievement of 56.60% induced by pure punishment strategy.

In order to further explore how the hybrid incentive strategy influences the group achievement and to find the optimal weight of incentives, we show the average group achievement η_G as a function of α for different values of risk r in Figure 2A. We find that the group achievement first grows fast and then increases slowly when increasing the value of α . Thus, we can conclude that pure reward strategy can always do better in improving the level of group achievement than other combinations of incentives. We note that here $c/b=0.1$ was used to obtain Figure 2A, but our findings remain valid for a broad range of c/b ratio.

We further present η_G as a function of risk r for different values of α , as shown in Figure 2B. For the sake of comparison, we also show the results when there is no incentive strategy (see black dash line). We find that the level of group achievement increases as we increase r . Secondly, the introduction of incentive strategy

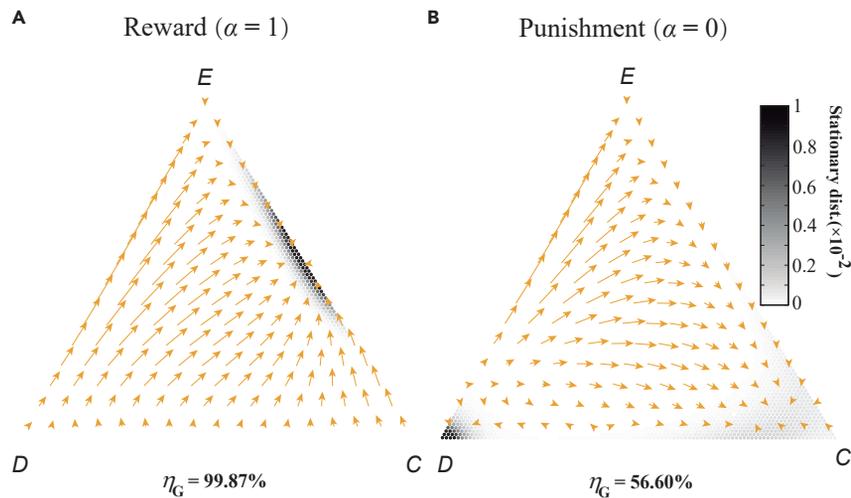


Figure 1. Evolutionary dynamics of cooperators, defectors, and executors under local scheme with fixed incentives

Here, the risk level is $r = 0.2$. In panel A, in the extreme case of $\alpha = 1$, E becomes equivalent to a pure rewarding strategy, while the other extreme case of $\alpha = 0$, where E is a pure punisher, can be seen in panel B. The darker dots in the simplex represent the regions where the population spends more time. Orange arrows represent the most likely direction of evolution when the population leaves the current configuration, obtained by computing the gradient of selection. We set the collective target M to 75% of the group size. Local scheme is established when the number of executors exceeds 25% of the group. Parameter values are $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_r = 0.03$, and $\pi_e = 0.3$.

provides better outcome compared to the traditional two-strategy case. Furthermore, when strategy E is introduced, it is better to adjust the weight factor toward the pure rewarding case which can improve the result further especially in the small risk region.

In addition to the risk of collective failure mentioned above, the mutation rate μ also plays an important role in the evolution of cooperation (Santos and Pacheco, 2011; Vasconcelos et al., 2013). However, it is unclear whether pure reward is still the optimal incentive strategy when individual mutation is considered as it introduces exogenous behavioral changes. To answer this question, in Figure 2C, we show η_G as a function of μ for three different values of α . We find that the group achievement first increases until reaching the maximum, and then, it decreases with the increasing value of μ for different α values. We find that when the mutation value is high, the group achievement level reaches a certain value regardless of the value of α , which corresponds to random play, but it is always the highest for $\alpha = 1$, and this difference is the most conspicuous for small mutation rates (see the inset of Figure 2C). Thus, we can conclude that our main finding is robust against the introduction of mutation.

We further find that the success of pure rewarding strategy for fixed incentives is induced by the efficiency of the institution. As shown in Figure 3A, the institution prevalence increases as we increase α for different risk values of r . We further compute the average fine on defectors and the average reward on cooperators and executors as presented in Figure 3B and observe that as the parameter α increases, the average reward amount on cooperators and executors monotonically increases from zero. On the other hand, the average fine amount on defectors first increases until reaching the maximum and then decreases with increasing the value of α . We can find that for $\alpha = 1$, executors and cooperators have the greatest evolutionary advantage over defectors, which is most conducive to the institution prevalence and the group achievement.

Flexible incentives under a local scheme

Beside fixed incentives, we can also apply flexible values imposed by the executor strategy where the actual level depends on the number of E players. Furthermore, the success of their collective efforts is acknowledged via an enhancement factor δ (Sasaki and Unemi, 2011; Sasaki et al., 2012). Hence, an enhanced fine and reward are distributed among defecting and cooperating players, respectively. In Figure 4, we first show the stationary distribution and the gradient of selection for low and high enhancement factor values

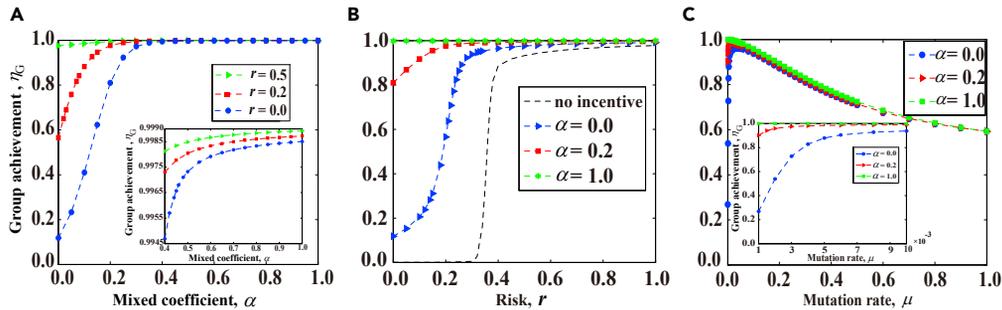


Figure 2. The average group achievement η_G under local scheme with fixed incentives

Panel A shows η_G as a function of α for three different values of r . The inset of panel A shows η_G as a function of α , where the value of α ranges from 0.4 to 1.0.

Panel B shows η_G as a function of r in the cases without incentive strategy E and with three different weight factor values in its presence.

Panel C shows η_G as a function of μ for three different values of α . The inset of panel c shows η_G as a function of μ , where the value of μ ranges from 0.001 to 0.01. Parameter values are $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $\pi_e = 0.3$ in panels A and B; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $r = 0.3$, $\pi_t = 0.03$, and $\pi_e = 0.3$ in panel C.

in the extreme cases of $\alpha = 1$ (pure reward) and $\alpha = 0$ (pure punishment). As shown in Figure 4A, the entire population will spend most time near configurations in which most individuals are cooperators when δ is low. Besides, we observe that most arrows flow to the vertex C and a few arrows point to vertex D. In Figure 4B, the flow of arrows is basically consistent with that observed in Figure 4A. However, the background shadow area near vertex C in Figure 4A is significantly larger than the corresponding region in Figure 4B, which suggests that pure reward strategy has an evolutionary advantage over pure punishment strategy in improving the level of group achievement. For high enhancement factor of the institutional fund, the application of pure reward makes the system more cooperative and the entire population spend most of the time in the vicinity of the CE-edge, as shown in Figure 4C. Besides, most of the arrows in the simplex flow to the regions near the C-corner. The corresponding simplex of pure punishment shown in Figure 4D illustrates very clearly that the dark area is close to vertex D; hence, we can conclude that the usage of pure reward strategy is more beneficial than the application of pure punishment strategy.

We also compute and add the average group achievement η_G values to the simplexes, which also illustrate clearly that pure reward performs better and this impact can be increased for a higher value of enhancement factor of the institution pool. In the following, we present how the group achievement η_G varies with the value of α for different values of δ in Figure 5A. This panel demonstrates clearly that the optimal weight of incentives is $\alpha = 1$, corresponding to the pure reward case. In particular, we find that for low δ (e.g., $\delta = 2$), the level of group achievement approaches zero for any value of α . However, for a larger δ value, the group achievement level grows as α increases and can be significantly improved when the value of α is larger than an intermediate value.

We further present η_G as a function of r for different values of α , as shown in Figure 5B. We also show the results in the case when additional incentive strategy is absent (see black dash line). In general, the group achievement level increases by enhancing the risk level r . Besides, we observe that the introduction of the incentive strategy may elevate the value of η_G , especially in the pure reward extreme case, but the improvement is not as shocking as for the fixed incentives version, shown in Figure 2B. In Figure 5C, we show how the group achievement η_G varies with increasing the enhancement factor δ for different values of α . In general, the group achievement grows as we increase δ , but the most remarkable improvement can be seen in the pure reward limit at $\alpha = 1$. Finally, we also explore the possible impact of mutation rate μ on the group achievement for three different values of α . As Figure 5D shows, the introduction of a small mutation rate can promote the group achievement significantly. However, when the μ becomes high, the group achievement level reaches a certain value regardless of the value of α . But, as previously stated, the best improvement can be obtained again for pure reward strategy.

In order to explore why pure reward strategy is more beneficial to the group achievement than other incentive strategies, we also present the institution prevalence η_I in dependence of the weight factor

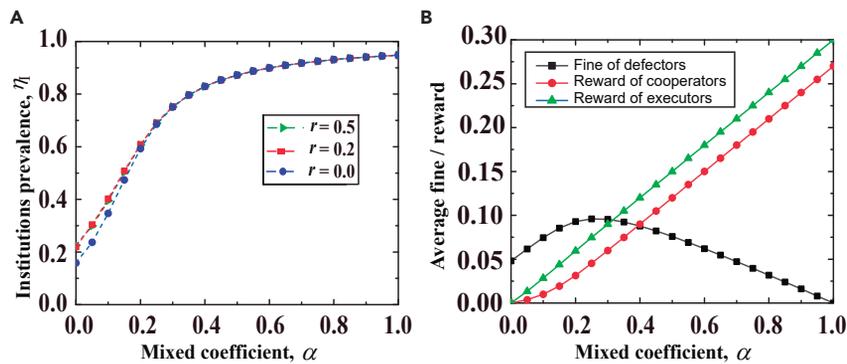


Figure 3. The institution prevalence η_i and the average reward and fine under local scheme with fixed incentives

Panel A shows η_i as a function of α for three different values of r . Panel B shows the average fine on defectors and the average reward on cooperators and executors as a function of α . Parameter values are $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_c = 0.03$, and $\pi_e = 0.3$ in panel A; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $r = 0.2$, $\mu = 1/Z$, $\pi_c = 0.03$, and $\pi_e = 0.3$ in panel B.

α , as shown in Figure 6A. We find that for flexible incentives, the institution prevalence increases as the weight factor increases for different values of δ , which means that the η_i value reaches the maximum for $\alpha = 1$. This indicates that the efficiency and prevalence of such institution can be best enhanced by the pure reward strategy. In addition, we see that the average reward amount on cooperators and executors increases from zero as we increase the α value (Figure 6B). On the contrary, the average fine on defectors slowly decreases and reaches to zero when α is one (Figure 6B). We then find that the evolutionary advantage of cooperators and executors over defectors can be best promoted by the pure reward strategy since their incentive difference can reach the maximum for $\alpha = 1$, which can facilitate the success of group achievement.

DISCUSSION

In this study, we have introduced a combined incentive strategy to address the collective-risk social dilemma game, where the corresponding players punish free-riders and reward cooperators simultaneously. The fundamental question is how to share available resources to execute these tasks. We have found that pure reward strategy can lead to the highest level of group achievement, no matter whether we apply fixed or flexible incentives for governing the commons. It was demonstrated that when the risk value of collective failure, the mutation rate, or the enhancement factor for institutional sanctioning is changed to some extent, our conclusion about the superiority of pure reward remains intact. Furthermore, we stress that this finding is not only valid under a local scheme of sanctioning but also valid for flexible incentives under global level of institutions (see supplemental information where the latter option is discussed in details). The comparison of the mentioned cases also reveals that the application of local incentives works more efficiently.

Previous works have demonstrated the advantage of a polycentric governance approach in solving the collective-risk social dilemma and found that the local scheme of sanctioning based on pool punishment is more effective than that associated with a single, global one for providing better conditions both for cooperation to thrive and for ensuring the maintenance of such sanctioning (Vasconcelos et al., 2013; Green et al., 2014; Cole, 2015). In our work, we also confirm this conclusion, and we further find that it is valid not only for fixed but also for the case when incentives are flexible. Thus, our work has greatly extended the validity of previous conclusions.

In our work, the incentives on cooperators and defectors are imposed by executors, a third combined strategy introduced into the collective-risk social dilemma. This consideration of the enforcement strategy as a voluntary enforcement involves a more complex three-strategy system. Indeed, this is a kind of a bottom-up self-regulation for influencing the evolution of cooperation (Sugiarto et al., 2017). We stress that this approach is different from the top-down-like incentive mechanism used in previous works (Chen et al., 2015; Góis et al., 2019), in which cooperators can be rewarded and defectors can be punished directly by external centralized institutions. These works have concluded that the adaptive hybridization of incentives can best promote the evolution of cooperation both in the traditional public goods game

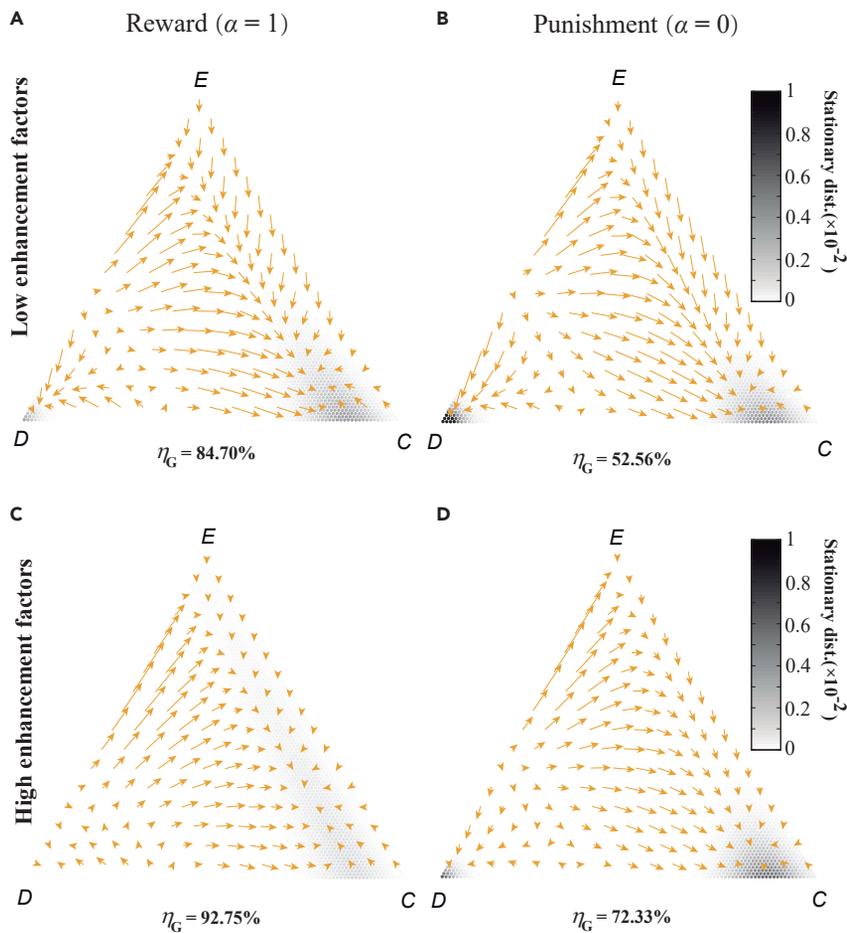


Figure 4. Evolutionary dynamics of cooperator, defector, and executor strategy under local scheme with flexible incentives

The risk level is $r = 0.35$. Panels A and C show the case of pure reward ($\alpha = 1$). Panels B and D show the case of pure punishment ($\alpha = 0$). Parameter values are $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $\delta = 1.4$ in panels A and B; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $\delta = 3$ in panels C and D.

(Chen et al., 2015) and in the collective-risk social dilemma game (Góis et al., 2019). When we consider the exogenous stability of these institutions, we find that pure reward strategy can better promote the evolution of cooperation than hybrid incentive strategies in the collective-risk social dilemma. This is because in our framework the incentive budget on individuals is not constrained, and the evolutionary advantage of cooperators and executors over defectors can be significantly promoted due to the maximal incentive difference induced by the pure reward strategy. Hence, the efficiency of institution can be best enhanced by the pure reward strategy, and from an evolutionary perspective, we show how an incentive strategy can be maintained as an endogenously chosen institution for cooperative governance of risky commons. Thus, our work can enrich the knowledge of the emergence of institutions for governing the commons (Sigmund et al., 2010; Vasconcelos et al., 2013; Dannenberg and Gallier, 2020).

Many research studies have demonstrated that reward or punishment is a viable approach in promoting the evolution of public cooperation (Szolnoki et al., 2011; Sasaki and Unemi, 2011; Sasaki et al., 2012). However, identifying the best way to distribute incentives remains an open question, given the difficulty in assessing the advantages and disadvantages of each possibility in different scenarios (Hilbe and Sigmund, 2010; Szolnoki and Perc, 2013). The majority of previous research studies addressing the “stick vs. carrot” management strategy concluded that punishment is more effective than reward in sustaining common cooperation in the conventional public goods game (Sigmund, 2007; Sasaki et al., 2012). However, the conventional

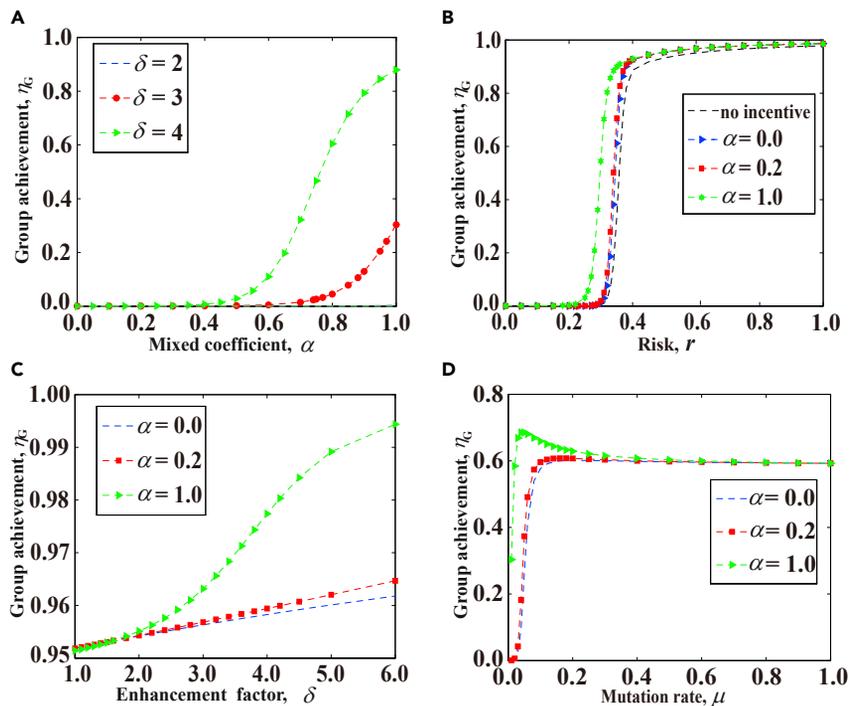


Figure 5. The average group achievement η_G under local scheme with flexible incentives

Panel A shows η_G as a function of α for three different values of δ . Panel B shows η_G as a function of r in the cases without incentive strategy E and with three different values of α . Panel C shows η_G as a function of δ for three different values of α . Panel D shows η_G as a function of μ for three different values of α . Parameter values are $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $r = 0.2$ in panel A; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $\delta = 2$ in panel B; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $r = 0.5$ in panel C; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, $r = 0.2$, and $\delta = 3$ in panel D.

public goods game fails to capture the significant feature of many real social scenarios, such as climate change (Green et al., 2014). In this work, instead of the conventional public goods game, we apply the collective-risk social dilemma game which has been proved as a framework to investigate the inherent problems of collective cooperation regarding avoiding dangerous climate change and other problems of this type (Santos and Pacheco, 2011), and we show that the optimal sanctioning policy is pure reward, no matter whether fixed or flexible incentives are considered. Thus, our work may unveil the effects of incentives on cooperative governance of risky commons in a real scenario.

Stimulating some nations or regions to reduce their greenhouse gas emissions is a fundamental task to mitigate climate change. How to design effective incentive measures is particularly important for this purpose. Since the Paris Agreement, we are witnessing the transformation of some countries' environmental governance to high-quality management (Rogelj et al., 2016). The establishment of the bonus-penalty mechanism, including giving appropriate compensation fees to emission reduction and imposing some fines on areas where pollution emission exceeds the prescribed quantity, has effectively promoted coordinated emission reduction (Xue et al., 2020). However, it is still unclear how to weigh rewards and punishments in realistic yet complicated situations. Our research may provide some advices for policy makers from a theoretical perspective: pure reward is more effective than pure punishment and other hybrid incentive measures in achieving emission reduction targets, especially if the stability of the institution is of concern.

Limitations of the study

The current study focuses on exploring the optimal incentive strategy in the scenario where the combination of punishment and reward is involved. However, it still has some limitations, which could be further extended. In this work, we consider the institution-type incentives with which punishment and reward are used. Indeed, there are some other types of incentives in the real society, and the implementation of incentive strategies is not necessarily institutional (Van Lange et al., 2014; Perc et al., 2017). In the future

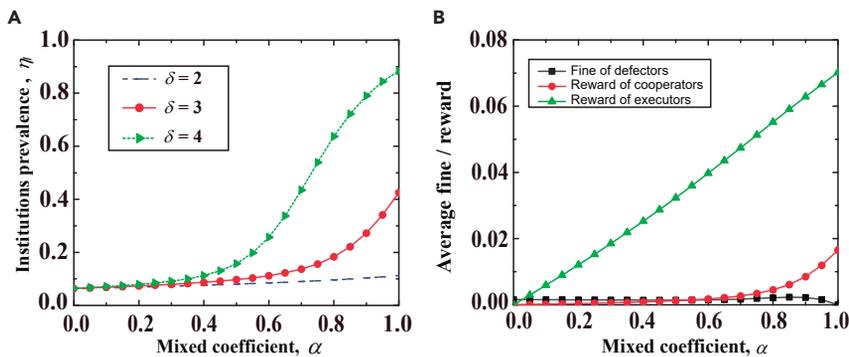


Figure 6. The institution prevalence η and the average reward and fine under local scheme with flexible incentives

Parameter values are $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\pi_t = 0.03$, and $r = 0.2$ in panel A; $Z = 100$, $N = 4$, $c = 0.1$, $b = 1$, $\mu = 1/Z$, $\delta = 3$, $\pi_t = 0.03$, and $r = 0.2$ in panel B.

work, we can thus consider different types of incentive strategies, e.g., social exclusion (Sasaki and Uchida, 2013) or different implementation ways of incentive, e.g., peer-type incentives (Hauert et al., 2007; Sigmund et al., 2010). In a recent work (Han et al., 2021), it is shown that peer punishment is more efficient than peer reward in promoting safety compliance and in mitigating the risk from artificial intelligence unsafe development. However, it is unknown whether pure reward strategy is still the best incentive strategy to promote the evolution of cooperation when peer-type incentives are considered in the collective-risk social dilemmas. In addition, in this work, we consider that whether incentives can be provided or not depends on the number of executors. Indeed, the number of cooperators can be used as a criterion to decide whether or not to provide incentives (Han and Tran-Thanh, 2018). In this scenario, it is interesting to explore the optimal incentive strategy for cooperative governance of risky commons. Moreover, implementation of local incentive requires additional information, and hence, it is also necessary to consider the cost of information collection for executors. Therefore, it is worth further exploring how the introduction of observation cost or information collection cost affects cooperative governance of risky commons (Szolnoki and Chen, 2015). Our work reveals that pure reward strategy is a more efficient incentive strategy for cooperative governance of risky commons. This may result from that executors bear the cost of punishing defectors (Szolnoki and Perc, 2017), while the beneficiary of rewarding mechanism actually includes themselves, which needs to be verified in the future work. Indeed, considering second-order sanctions is important in the context of pool incentives (Fowler, 2005; Perc, 2012; García and Traulsen, 2019), so it may also open an interesting research avenue to explore the efficiency of incentives in the collective-risk social dilemma game. Previous studies have shown that the existence of second-order free-riders plays an important role in the emergence of cooperation and institutions (Sigmund et al., 2010; Sasaki and Unemi, 2011; Szolnoki and Perc, 2017). Therefore, it is worth investigating whether the introduction of second-order punishment and second-order reward can improve cooperative governance of risk commons (Van Lange et al., 2014). Furthermore, our research explores the optimal incentive strategy by adjusting the fixed relative weight parameter for the combination of punishment and reward. But indeed, this parameter can be time varying, depending on the population states. On the other hand, on the premise of ensuring a sufficiently good outcome, how to optimize the cost of providing incentives is another important issue of incentive design (Han and Tran-Thanh, 2018). Hence, it is meaningful to investigate the optimal incentive strategy in this scenario by means of optimal control theory or reinforcement learning approach (Wang et al., 2019; Ratliff et al., 2019). In this work, we have considered a well-mixed interaction where individuals perform random interactions. However, the interactions among individuals are typically not random but rather that they are limited to a set of neighbors in a structured population, which could be described by a complex interaction network (Santos and Pacheco, 2011; Szolnoki et al., 2011; Pinheiro et al., 2012; Szolnoki and Perc, 2013; Perc et al., 2017). Thus, it could be interesting to explore the optimal incentive strategy in structured populations. We believe that future work considering these extensions will be valuable and improve our understanding of the role of incentive strategies in cooperative governance of risk commons.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- METHOD DETAILS
 - Collective-risk social dilemma game
 - Settings of fixed incentives
 - Settings of flexible incentives
 - Evolutionary dynamics

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2021.102844>.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61976048 and 62036002) and the Fundamental Research Funds for the Central Universities of China.

AUTHOR CONTRIBUTIONS

All authors designed and performed the research as well as wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 10, 2021

Revised: June 23, 2021

Accepted: July 8, 2021

Published: August 20, 2021

REFERENCES

- Barrett, S., and Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proc. Natl. Acad. Sci. USA* *109*, 17372–17378.
- Cole, D.H. (2015). Advantages of a polycentric approach to climate change policy. *Nat. Clim. Change* *5*, 114–118.
- Chen, X., Szolnoki, A., and Perc, M. (2012). Risk-driven migration and the collective-risk social dilemma. *Phys. Rev. E* *86*, 036101.
- Chen, X., Zhang, Y., Huang, T.-Z., and Perc, M. (2014). Solving the collective-risk social dilemma with risky assets in well-mixed and structured populations. *Phys. Rev. E* *90*, 052823.
- Chen, X., Sasaki, T., Brännström, Å., and Dieckmann, U. (2015). First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* *12*, 20140935.
- Dannenberg, A., and Gallier, C. (2020). The choice of institutions to solve cooperation problems: a survey of experimental research. *Exp. Econ.* *23*, 716–749.
- Diaz, D., and Moore, F. (2017). Quantifying the economic risks of climate change. *Nat. Clim. Change* *7*, 774.
- Domingos, E.F., Grujić, J., Burguillos, J.C., Kirchsteiger, G., Santos, F.C., and Lenaerts, T. (2020). Timing uncertainty in collective risk dilemmas encourages group reciprocity and polarization. *iScience* *23*, 101752.
- Fehr, E., and Gächter, S. (2002). Altruistic punishment in humans. *Nature* *415*, 137–140.
- Fowler, J.H. (2005). Second-order free-riding problem solved? *Nature* *437*, E8.
- García, J., and Traulsen, A. (2019). Evolution of coordinated punishment to enforce cooperation from an unbiased strategy space. *J. R. Soc. Interface* *16*, 20190127.
- Ginsberg, A., and Fu, F. (2019). Evolution of cooperation in public goods games with stochastic opting-out. *Games* *10*, 1.
- Góis, A.R., Santos, F.P., Pacheco, J.M., and Santos, F.C. (2019). Reward and punishment in climate change dilemmas. *Sci. Rep.* *9*, 16193.
- Green, J.F., Sterner, T., and Wagner, G. (2014). A balance of bottom-up and top-down in linking climate policies. *Nat. Clim. Change* *4*, 1064–1067.
- Han, T.A., Pereira, L.M., and Lenaerts, T. (2015). Avoiding or restricting defectors in public goods games? *J. R. Soc. Interface* *12*, 20141203.
- Han, T.A., Pereira, L.M., Lenaerts, T., and Santos, F.C. (2021). Mediating artificial intelligence developments through negative and positive incentives. *PLOS ONE* *16*, e0244592.
- Han, T.A., and Tran-Thanh, L. (2018). Cost-effective external interference for promoting the evolution of cooperation. *Sci. Rep.* *8*, 15997.
- Hardin, G. (1968). The tragedy of the commons. *Science* *162*, 1243–1248.
- Hauert, C., De Monte, S., Hofbauer, J., and Sigmund, K. (2002). Replicator dynamics for optional public goods games. *J. Theor. Biol.* *218*, 187–194.
- Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., and Sigmund, K. (2007). Via freedom to coercion: the emergence of costly punishment. *Science* *316*, 1905–1907.
- Helbing, D. (1993). Boltzmann-like and Boltzmann-Fokker-Planck equations as a foundation of behavioral models. *Phys. A* *196*, 546–573.
- Hilbe, C., and Sigmund, K. (2010). Incentives and opportunism: from the carrot to the stick. *Proc. R. Soc. B* *277*, 2427–2433.
- Imhof, L.A., Fudenberg, D., and Nowak, M.A. (2005). Evolutionary cycles of cooperation and defection. *Proc. Natl. Acad. Sci. USA* *102*, 10797–10800.
- Inman, M. (2009). The climate change game. *Nat. Clim. Change* *1*, 130–133.

- Jacquet, J., Hagel, K., Hauert, C., Marotzke, J., Röhl, T., and Milinski, M. (2013). Intra- and intergenerational discounting in the climate game. *Nat. Clim. Change* 3, 1025–1028.
- Kampen, N.V. (2007). *Stochastic Processes in Physics and Chemistry* (North-Holland).
- Lenton, T.M. (2014). Tipping climate cooperation. *Nat. Clim. Change* 4, 14–15.
- Marotzke, J., Semmann, D., and Milinski, M. (2020). The economic interaction between climate change mitigation, climate migration and poverty. *Nat. Clim. Change* 10, 518–525.
- Milinski, M., Sommerfeld, R.D., Krambeck, H.-J., Reed, F.A., and Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proc. Natl. Acad. Sci. USA* 105, 2291–2295.
- Milinski, M., Röhl, T., and Marotzke, J. (2011). Cooperative interaction of rich and poor can be catalyzed by intermediate climate targets. *Clim. Change* 109, 807–814.
- Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press).
- Pacheco, J.M., Vasconcelos, V.V., and Santos, F.C. (2014). Climate change governance, cooperation and self-organization. *Phys. Life Rev.* 11, 2212–2217.
- Perc, M. (2012). Sustainable institutionalized punishment requires elimination of second-order free-riders. *Sci. Rep.* 2, 344.
- Perc, M., Jordan, J.J., Rand, D.G., Wang, Z., Boccaletti, S., and Szolnoki, A. (2017). Statistical physics of human cooperation. *Phys. Rep.* 687, 1–51.
- Pinheiro, F.L., Santos, F.C., and Pacheco, J.M. (2012). How selection pressure changes the nature of social dilemmas in structured populations. *New J. Phys.* 14, 073035.
- Rand, D.G., and Nowak, M.A. (2013). Human cooperation. *Trends Cogn. Sci.* 17, 413–425.
- Ratliff, L.J., Dong, R., Sekar, S., and Fiez, T. (2019). A perspective on incentive design: challenges and opportunities. *Annu. Rev. Control Robot. Auton. Syst.* 2, 305–338.
- Rogelj, J., Den Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K., and Meinshausen, M. (2016). Paris Agreement climate proposals need a boost to keep warming well below 2°C. *Nature* 534, 631–639.
- Santos, F.C., and Pacheco, J.M. (2011). Risk of collective failure provides an escape from the tragedy of the commons. *Proc. Natl. Acad. Sci. USA* 108, 10421–10425.
- Santos, F.C., Vasconcelos, V.V., Santos, M.D., Neves, P., and Pacheco, J.M. (2012). Evolutionary dynamics of climate change under collective-risk dilemmas. *Math. Models Methods Appl. Sci.* 22, 1140004.
- Santos, F.P., Levin, S.A., and Vasconcelos, V.V. (2021). Biased perceptions explain collective action deadlocks and suggest new mechanisms to prompt cooperation. *iScience* 24, 102375.
- Sasaki, T., and Unemi, T. (2011). Replicator dynamics in public goods games with reward funds. *J. Theor. Biol.* 287, 109–114.
- Sasaki, T., Brännström, Å., Dieckmann, U., and Sigmund, K. (2012). The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proc. Natl. Acad. Sci. USA* 109, 1165–1170.
- Sasaki, T., and Uchida, S. (2013). The evolution of cooperation by social exclusion. *Proc. R. Soc. B* 280, 20122498.
- Schroeder, H., Boykoff, M.T., and Spiers, L. (2012). Equity and state representations in climate negotiations. *Nat. Clim. Change* 2, 834–836.
- Shirado, H., and Christakis, N.A. (2020). Network engineering using autonomous agents increases cooperation in human groups. *iScience* 23, 101438.
- Sigmund, K. (2007). Punish or perish? retaliation and collaboration among humans. *Trends Ecol. Evol.* 22, 593–600.
- Sigmund, K., Silva, H.D., Traulsen, A., and Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature* 466, 861–863.
- Sugiarto, H.S., Chuang, N.N., Lai, C.H., and Chew, L.Y. (2017). Emergence of cooperation in a coupled socio-ecological system through a direct or an indirect social control mechanism. *J. Phys. Commun.* 1, 055019.
- Szabó, G., and Tóke, C. (1998). Evolutionary prisoner's dilemma game on a square lattice. *Phys. Rev. E* 58, 69.
- Szolnoki, A., Szabó, G., and Czakó, L. (2011). Competition of individual and institutional punishments in spatial public goods games. *Phys. Rev. E* 84, 046106.
- Szolnoki, A., and Perc, M. (2013). Correlation of positive and negative reciprocity fails to confer an evolutionary advantage: phase transitions to elementary strategies. *Phys. Rev. X* 3, 041021.
- Szolnoki, A., and Chen, X. (2015). Benefits of tolerance in public goods games. *Phys. Rev. E* 92, 042813.
- Szolnoki, A., and Perc, M. (2015). Antisocial pool rewarding does not deter public cooperation. *Proc. R. Soc. B* 282, 20151975.
- Szolnoki, A., and Perc, M. (2017). Second-order free-riding on antisocial punishment restores the effectiveness of prosocial punishment. *Phys. Rev. X* 7, 041027.
- Tavoni, A. (2013). Building up cooperation. *Nat. Clim. Change* 3, 782–783.
- Van Lange, P.A.M., Rockenbach, B., and Yamagishi, T. (2014). *Reward and Punishment in Social Dilemmas* (Oxford University Press).
- Van Segbroeck, S., Pacheco, J.M., Lenaerts, T., and Santos, F.C. (2012). Emergence of fairness in repeated group interactions. *Phys. Rev. Lett.* 108, 158104.
- Vasconcelos, V.V., Santos, F.C., and Pacheco, J.M. (2013). A bottom-up institutional approach to cooperative governance of risky commons. *Nat. Clim. Change* 3, 797–801.
- Vasconcelos, V.V., Santos, F.C., and Pacheco, J.M. (2015). Cooperation dynamics of polycentric climate governance. *Math. Models Methods Appl. Sci.* 25, 2503–2517.
- Wang, J., Fu, F., and Wang, L. (2010). Effects of heterogeneous wealth distribution on public cooperation with collective risk. *Phys. Rev. E* 82, 016102.
- Wang, J., Fu, F., Wu, T., and Wang, L. (2009). Emergence of social cooperation in threshold public goods games with collective risk. *Phys. Rev. E* 80, 016101.
- Vasconcelos, V.V., Santos, F.C., Pacheco, J.M., and Levin, S.A. (2014). Climate policies under wealth inequality. *Proc. Natl. Acad. Sci. USA* 111, 2212–2216.
- Wang, S., Chen, X., and Szolnoki, A. (2019). Exploring optimal institutional incentives for public cooperation. *Commun. Nonlinear Sci. Numer. Simul.* 79, 104914.
- Wu, T., Fu, F., Zhang, Y., and Wang, L. (2013). The increased risk of joint venture promotes social cooperation. *PLOS ONE* 8, e63801.
- Xue, L., Wang, W., and Zhang, M. (2020). Research on bonus-penalty mechanism of pollution abatement: a case study of the northeastern region of China. *J. Clean. Prod.* 267, 122069.
- Zhang, Y., Fu, F., Wu, T., Xie, G., and Wang, L. (2013). A tale of two contribution mechanisms for nonlinear public goods. *Sci. Rep.* 3, 2021.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
r together with b , c , Z , N , M , β , and μ		
for collective-risk social dilemma in a finite population	Santos and Pacheco	https://doi.org/10.1073/pnas.1015648108
π_t together with π_e and n_E for fixed incentives	Vasconcelos et al.	https://doi.org/10.1038/nclimate1927
δ together with π_t for flexible incentives	Sasaki and Unemi	https://doi.org/10.1016/j.jtbi.2011.07.026
Software and algorithms		
MATLAB R2014A	MathWorks	https://www.mathworks.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Xiaojie Chen (xiaojiechen@uestc.edu.cn).

Materials availability

No materials were newly generated for this paper.

Data and code availability

The data that support the results of this study are available from the corresponding authors upon request. The code used to generate the figures is available from the corresponding authors upon request.

METHOD DETAILS

Collective-risk social dilemma game

We consider a finite well-mixed population of Z individuals. From the entire population, N individuals are selected randomly to form a group for playing a collective-risk social dilemma game. In the group, each individual has an endowment b at the beginning of game. Then, everyone decides whether or not to contribute an amount c to the common pool. Furthermore, we require a minimum collective contribution Mc ($0 < M \leq N$) or a minimum number M of cooperators to ensure the benefit of everyone within the group. According to the traditional setup, cooperators contribute to reach the collective target while defectors not. If the collective target is not reached, all participants within the group will lose their remaining endowments with probability r ($0 \leq r \leq 1$). Otherwise, individuals in the group retain their endowments. Notably, the parameter r represents the risk level of collective failure (Santos and Pacheco, 2011).

Based on the above description, the payoffs of defectors and cooperators in a group having j_C cooperators and $N - j_C$ defectors obtained from the game can be respectively written as

$$\begin{aligned}\Pi_D(j_C) &= b\Theta(j_C - M) + (1 - r)b[1 - \Theta(j_C - M)], \\ \Pi_C(j_C) &= \Pi_D(j_C) - c,\end{aligned}$$

where $\Theta(x)$ is the Heaviside function, that is, $\Theta(x) = 0$ if $x < 0$, being one otherwise.

In the following, we introduce a hybrid incentive strategy into the collective-risk social dilemma game and define two different kinds of enforcing institutions. Namely, in the first case the imposed incentives are fixed, while in the other case they are flexible.

Settings of fixed incentives

We then introduce a third strategy into the collective-risk social dilemma game, that is, an executor (E) who not only contributes an amount c to the common pool, who but also contributes an investment to the

sanctioning pool at a cost π_t . This makes it possible to punish free-riders and to reward cooperative players simultaneously. Accordingly, when the number of executors exceeds the given threshold n_E , each executor imposes an α portion of π_e as positive incentive on every cooperators, while $1 - \alpha$ portion of executive pool serves to punish free-riders. Here, α represents the relative weight between positive and negative incentives. In an extreme case, when $\alpha = 0$, executors only choose to punish defectors and they become simple punishers. This simplified situation was discussed in a work of Vasconcelos et al. (Vasconcelos et al., 2013). In the other extreme situation, when $\alpha = 1$, executors only choose to reward cooperators and E becomes a simple rewarding strategy. To sum up our model, the payoffs of the competing strategies where j_C cooperators and j_E executors are present in the group can be written as

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

where Δ corresponds to the sanction function, which relies on whether the sanction is local or global. When the local scheme is considered, Δ is given by $\Theta(j_E - n_E)$, which means that the enforcement is executed at the group level if the number of executors in the group is not less than the group threshold. However, when the global scheme is considered, $\Delta = \Theta(i_E - n_E)$, which means that incentives are applied in the whole population if the total number of executors i_E is not less than the corresponding threshold level.

Settings of flexible incentives

Alternatively, the level of incentive π_e may not be fixed but may depend on the number of executors. For instance, the institutional pool may behave as a typical public goods pool and its impact can be proportional to the simple sum of individual contributions π_e of executor players with an enhancement factor $\delta > 1$. Notably, in the full reward extreme case this idea was used in one previous work (Sasaki and Unemi, 2011). More importantly, in our extended model we do not just apply a combined incentives strategy, but we still hold the collective-risk feature of the institutional pool. Therefore, as previously, the application of incentives depends sensitively on whether the n_E threshold number of executors are present or not. Hence, the modified payoff values for flexible incentives are

$$\begin{aligned}\Pi_D(j_C, j_E) &= \Pi_D(j_C + j_E) - (1 - \alpha)\frac{\pi_t\delta j_E}{N - j_C - j_E}\Delta, \\ \Pi_C(j_C, j_E) &= \Pi_D(j_C + j_E) + \alpha\frac{\pi_t\delta j_E}{j_C + j_E}\Delta - c, \\ \Pi_E(j_C, j_E) &= \Pi_D(j_C + j_E) + \alpha\frac{\pi_t\delta j_E}{j_C + j_E}\Delta - c - \pi_t.\end{aligned}$$

Evolutionary dynamics

The average payoffs of the aforementioned strategies in a configuration $\mathbf{i} = (i_C, i_E, i_D)$, characterized by the number of cooperators i_C , the number of executors i_E , and the number of defectors $i_D = Z - i_C - i_E$, can be computed by using a multivariate hypergeometric sampling (Hauert et al., 2007) as follows

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

where we note that $j_C \leq i_C$, $j_E \leq i_E$, and $j_D \leq i_D$ and thus impose that the binomial coefficients satisfy $\binom{i}{j} = 0$ if $i < j$.

To describe how individuals adopt new strategies over time, we consider a stochastic process of pairwise comparison (Van Segbroeck et al., 2012). In each evolutionary time step, a randomly chosen individual with strategy L adopts the strategy R of another randomly selected individual from the population with probability $1/[1 + e^{-\beta(f_R - f_L)}]$, where the parameter β corresponds to the intensity of selection (Szabó and Tóke, 1998). For $\beta \rightarrow 0$, we have random drift, while for $\beta \rightarrow \infty$ the imitation process is more inclined to prefer the strategies that yield higher payoffs (Sigmund et al., 2010). Without loss of generality, we set up $\beta = 5.0$, meaning that it is very likely that the better performing players will pass their strategy to others, yet it is also possible that players will occasionally learn from a less successful individual. This setting is also consistent with the previous work (Vasconcelos et al., 2013). We would like to stress that our main results remain valid when the value of β is changed appropriately (see Figure S5 in supplemental information). In addition to the imitation, we also allow strategy mutation, and we assume that an individual adopts a different randomly selected strategy with probability μ . Accordingly, the transition probability between strategies R and L under the above mutation-selection process can be written as

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

where i_L (i_R) represents the number of individuals with strategy L (R) in the population. Therefore, for a given configuration $\mathbf{i} = (i_C, i_E, i_D)$, the probability that the number of cooperators increases or decreases by one in the entire population is written as

$$T_{\mathbf{i}}^{C\pm} = T_{(i_C \pm 1, i_E \mp 1, i_D)} + T_{(i_C \pm 1, i_E, i_D \mp 1)}.$$

Similarly, the probability that the number of executors increases or decreases by one is given by

$$T_{\mathbf{i}}^{E\pm} = T_{(i_C \mp 1, i_E \pm 1, i_D)} + T_{(i_C, i_E \pm 1, i_D \mp 1)}.$$

Since the strategy update process depends only on the current configuration of the system, the associated evolutionary dynamics can be described as a Markov process over a two-dimensional space (Kampen, 2007) where the probability density function we designate by $p_{\mathbf{i}}(t)$, satisfies the following Master Equation

$$p_{\mathbf{i}}(t + \tau) - p_{\mathbf{i}}(t) = \sum_{\mathbf{i}'} \{ T_{\mathbf{i}\mathbf{i}'} p_{\mathbf{i}'}(t) - T_{\mathbf{i}'\mathbf{i}} p_{\mathbf{i}}(t) \}.$$

Here $T_{\mathbf{i}\mathbf{i}'}$ denotes the transition probability from the configuration \mathbf{i}' to \mathbf{i} per unit time τ (Imhof et al., 2005). Accordingly, we can obtain the so-called stationary distribution $\bar{p}_{\mathbf{i}}(t)$, by searching the eigenvector associated with the eigenvalue 1 of the transition matrix $\mathbf{T} = [T_{\mathbf{i}\mathbf{j}}]$ (Kampen, 2007). As an important quantity for describing the evolutionary dynamics of strategies in the system, the stationary distribution can characterize the fraction of time that the population spends in each possible configuration of the finite population.

In addition to the stationary distribution $\bar{p}_{\mathbf{i}}(t)$, another central quantity for describing the evolutionary dynamics of strategies is the gradient of selection $\nabla_{\mathbf{i}}$, which describes the most likely evolutionary path of change of population configuration with time. Employing the Kramers-Moyal expansion of the Master Equation (Helbing, 1993), we can obtain the gradient of selection as

$$\nabla_{\mathbf{i}} = (T_{\mathbf{i}}^{C+} - T_{\mathbf{i}}^{C-}) \mathbf{u}_C + (T_{\mathbf{i}}^{E+} - T_{\mathbf{i}}^{E-}) \mathbf{u}_E,$$

where \mathbf{u}_C and \mathbf{u}_E are unit vectors defining as a basis of the two dimensional simplex.

Furthermore, in order to investigate the role of different sanctioning policies, we provide a key quantity $a_G(\mathbf{i})$ to compute the fraction of groups that reach a given threshold, that is, the collective target (Vasconcelos et al., 2013). We should compute the average group achievement from the perspective of the entire population and not obtain the value for groups centered on a given C , E , or D . For each possible configuration $\mathbf{i} = (i_C, i_E, i_D)$, accordingly we adopt the multivariate hypergeometric distribution to calculate $a_G(\mathbf{i})$ as

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

Subsequently, by using $a_G(i)$ combined with the stationary distribution $\bar{p}_i(t)$, we can calculate the average group achievement η_G as an important quantity which refers to the average proportion of groups that achieve the collective target in all possible group configurations, that is, $\eta_G = \sum_i \bar{p}_i a_G(i)$.

We further compute the fraction of groups that reach n_E executors for local institution or whether a global institution can be formed for the configuration i . Accordingly, we should compute the institution prevalence from the perspective of the entire population and not obtain the value for groups centered on a given C , E , and D . In both cases, we designate this quantity by $a_I(i)$, which can be calculated as

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

We subsequently use $a_I(i)$ combined with the stationary distribution $\bar{p}_i(t)$ to compute the institution prevalence η_I averaging over all possible configuration i , which is given as $\eta_I = \sum_i \bar{p}_i a_I(i)$.