



UvA-DARE (Digital Academic Repository)

Costly incentives design from an institutional perspective: cooperation, sustainability and affluence

Zhou, X.; Belloum, A.; Lees, M.H.; van Engers, T.; de Laat, C.

DOI

[10.1098/rspa.2022.0393](https://doi.org/10.1098/rspa.2022.0393)

Publication date

2022

Document Version

Final published version

Published in

Proceedings of the Royal Society A-Mathematical, Physical and Engineering Sciences

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Zhou, X., Belloum, A., Lees, M. H., van Engers, T., & de Laat, C. (2022). Costly incentives design from an institutional perspective: cooperation, sustainability and affluence. *Proceedings of the Royal Society A-Mathematical, Physical and Engineering Sciences*, 478(2265), [20220393]. <https://doi.org/10.1098/rspa.2022.0393>

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>)

Research



Cite this article: Zhou X, Belloum A, Lees MH, van Engers T, de Laat C. 2022 Costly incentives design from an institutional perspective: cooperation, sustainability and affluence. *Proc. R. Soc. A* **478**: 20220393. <https://doi.org/10.1098/rspa.2022.0393>

Received: 7 June 2022

Accepted: 9 August 2022

Subject Areas:

computer modelling and simulation, applied mathematics

Keywords:

institutional incentive, sustainability, evolutionary game theory

Author for correspondence:

Xin Zhou

e-mail: x.zhou@uva.nl

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6168135>.

Costly incentives design from an institutional perspective: cooperation, sustainability and affluence

Xin Zhou¹, Adam Belloum¹, Michael H. Lees¹,
Tom van Engers² and Cees de Laat¹

¹Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

²Faculty of Law, University of Amsterdam, Amsterdam, The Netherlands

XZ, 0000-0001-7706-8327

Incentives are usually introduced by the regulator entity (third-party), to promote cooperation in a market. The implementation of incentives is always costly and thus might fail to be enforced sustainably. This work aims at exploring the effects of incentives from an institutional perspective, while coping with the scenario where the third-party is part of the system but not composed by players. The evolutionary game theory (EGT) framework is applied to identify the incentives that lead to pure cooperation. In contrast to traditional EGT, this paper introduces an elimination mechanism that can reduce the market size. The incentives identified in the EGT analysis are further examined in simulation experiments which measure the market size, affluence and sustainability. The findings show: (1) light punishment leads to a reduction of the market size, yet heavier punishment is beneficial to the market size and wealth; (2) mixed incentives will generally lead to different wealth of the third party and of the participants. While under moderate strength, the wealth of both parties is the same and their overall wealth is maximal; (3) for sustainability, pure punishment (resp. reward) is sustainable (resp. unsustainable), the sustainability of mixed incentives depends on both their strength and agents' rationality level.

1. Introduction

In human society, incentives like rewarding norm followers [1,2] and punishing norm violators [3–5] are practical instruments for maintaining the order of a market or a community [6,7]. When participants can gain extra benefits from breaking the norm, self-interests always drive the participants to be a norm violator, and choose to defect or cheat. Incentives, that change the benefits and cost of actions, can largely decrease the rate of non-compliant behaviours and promote collaborations. Since the important role of incentives in real-world management has been recognized, experimental and theoretical studies on designing practical and effective proper incentives are increasing [8,9]. Evolutionary game theory (EGT) is an analytical framework widely applied in analysing and predicting the effect of the incentives [10–12], which relies on the Darwinian process of natural selection that drives participants toward the optimization of reproductive success [13,14]. With EGT, the dynamics of participants' population composition under a specific incentive can thus be observed. Previous game-theoretic studies on the design of incentives mainly focused on the dynamics of cooperation, such as the emergence of cooperation [15,16], the level of cooperation [10], or the sustainability of the cooperation [17], while few of them considering the sustainability of incentives, especially when incentives are enforced by an external thirdparty.

Based on the execution manner of incentives, related studies can be classified into two categories: (1) peer-to-peer incentives executed by participants (also called players) [2,18–20]; (2) institutional incentives executed by the third party [21–23], the third party can be composed by players [24,25], or completely from external [26]. The enforcement of incentives can be costly [27–29], and thus the incentive might be terminated if the resource for implementation cannot cover the cost. Despite the fact that it is reasonable to assume that the resource for implementing incentives is inexhaustible [30,31] under the peer reward or peer punishment scenario, as volunteer punishment and altruistic rewarding can always emerge [32,33]. Nevertheless, it is not the case when incentives are enforced in an institutional manner by the third party [34,35]. The cost of incentives' enforcement can hardly be ignored [25,28,29,36,37], as the high cost can potentially lead to the failure in the enforcement process [24].

A few works have considered minimizing the cost of incentive enforcement from the perspective of optimization, and managed to design cost-effective reward or punishment [38–42] for an external third-party (or external decision-maker [26]). In these studies, the third party is composed of agents that do not belong to the system, and the income of such an external third party was not considered. It is reasonable to ignore the income when incentives are carried out by a third party which is completely external to the system. For instance, when a fraud happens in a market, the income of the judiciary is rarely considered in the adjudication and enforcement process, as such institutions are independent to the market, and supported by the nation. Under this kind of circumstance, focusing on cost and committing to minimizing the cost is of great practical meaning. However, if the third party belongs to the system, but is not composed of players, such as the owner or the maintainer of the market, then not only the cost, but also the income of the third party need to be considered. Because only when the cash flow of the thirdparty is positive can the incentive be executed sustainably. Hence, from an institutional perspective, an effective incentive needs not only to foster collaboration, but also to be sustainable.

In addition to sustainability, incentives' influence on the affluence of participants and the third party is also vital in practice. An incentive that performs well in fostering collaboration may result in undesirable side effects, such as lowering the accumulated wealth of players or shrinking the market size, for example, which is in conflict with the affluence growth and the prolonged development of the market. Consequently, in this study, the evaluating criterion for incentives are extended beyond promoting collaboration to the sustainability and the impact on the market's affluence. Based on these criteria, we try to answer the question as to what kinds of incentives are practical and proper from the institutional perspective.

Pure reward, pure punishment, and mixed incentives are explored, and we first apply the framework of EGT to identify the analytical results of the population dynamics. Then, considering the complexity of the real world, we introduce the income and cost for the third party to evaluate the sustainability of incentives, while relaxing the unlimited population [13] assumption of EGT by assuming that the market capacity is limited and might shrink if participants are bankrupted. In the simulation experiments, the cooperation level, sustainability, and the affluence of both players and the third party are observed. Further, since participants are known as bounded rational [43], we study how the rational level of participants adjusts the process.

Based on the simulation results, the following insights are drawn: (1) participants' rationality can significantly affect of the incentives on promoting cooperation, lower rationality requires stronger incentives to maintain the cooperation level; (2) pure reward incentives promote participants' wealth but can hardly be implemented sustainably; (3) pure punishment is always sustainable. But light punishment can lead to the shrinkage of the market, while strong punishment is helpful for maintaining the market size, and beneficial for improving the affluence of both players and the third party; (4) mixed incentive combined the advantages of pure incentives—usually it can promote the cooperation level sustainably, its sustainability depends on both the strength of the incentive and the rational level of the agents; (5) mixed incentives induce a trade-off between the wealth of the participants and the third party, where moderate strength incentives can maximize the sum of accumulated wealth for both.

The remainder of this paper is organized as follows: §2 introduces the incentive model, including the cost and income of different parties in the market. Section 3 outlines the analytical results derived by the EGT, and with this foundation, we subsequently elaborate the design of the simulation experiments. The experimental results are reported and interpreted in §4. Finally, this article concludes with a discussion of the results and points out some future research directions.

2. Model

In this section, we first introduce the basic pairwise game played by the participants, and the incentives imposed in the market. Then we explain the income and expenditure for the third party when maintaining the market. With these two parts, the dynamic model for all parties is defined.

(a) Pairwise game and incentives

We consider a market with two parties, N participants that have pairwise interactions, and one independent third party that implements incentives for promoting cooperation. As players might lack professional regulator training, rarely can they freely switch roles from participant to maintainer. Hence, the third party here is assumed to be independent, rather than composed of players.

For homogeneous participants, each of them has the same strategy space $S = \{C, D\}$, and C (resp. D) represents the cooperation (resp. defection) strategy, also known as compliant (resp. non-compliant) behaviour. Choosing C by both of the agents can bring mutual benefits, whereas each of them has the temptation T to betray the other [44]. Such a situation is quite common in our daily life, and is often characterized using the Prisoner's dilemma game (PDG) [31,45]. We select PDG as the basic game model,¹ while the pay-off matrix is given in table 1. The mutual cooperation (resp. defection) profit is R (resp. P), and temptation for defecting is T .

Incentives discussed in this work include pure reward, pure punishment and mixed incentives. For pure reward, mutual cooperators will receive reward R_{CC} . The sucker's payoff is given as R_{CD} . We set $R_{CD} \geq R_{CC}$ to ensure that the reward for a sucker will be no less than the reward of a cooperator in a mutual cooperation. As for pure punishment, mutual defectors will receive the fine F_{DD} . The defector who betrays a cooperator will be fined by F_{CD} . We set $F_{CD} \geq F_{DD}$ which

¹It is worth noting that in our reality some more complicated scenarios can happen, for example, group interactions can replace the mentioned pairwise ones [22,46]. Meanwhile, agents can be heterogeneous in various features like risk preference or spatial position, etc. But in this work, we start with the most basic homogeneous agents playing pairwise PDG game.

Table 1. Pay-off matrix of the Prisoner's dilemma game.

		Player Y	
		C	D
Player X	C	R	-T
	D	T	P

means the punishment for a defector in $[D, C]$ will be no less than a defector in $[D, D]$. Mixed incentives require $R_{CD} + R_{CC} \neq 0$ and $F_{CD} + F_{DD} \neq 0$.

(b) The income and cost for the third party

To evaluate whether the incentive can be carried out in a sustainable way, we assume that only when the accumulated wealth of the third party is non-negative can the incentive be enforced. Hence, the income and cost of the third party are introduced.

In practice, tax, membership fee, or commission fee are common resources imposed by the third party for maintaining the order of the community or market [31,47,48]. This study assumes the income of the third party is composed of two parts: (1) the basic commission fee c_0 paid by the agents in each round [49]; (2) the fine retrieved from the defectors [50,51]. Let M be the amount of participants in the market, and particularly, $M^{(t)}$ denotes the amount at time t ($M^{(0)} = N, M^{(t)} \leq N$). The concrete income depends on M as well as on the population distribution. The population profile is a vector $\mathbf{x} = \{x, y\}$, $x + y = 1$, wherein x (resp. y) is the fraction of the cooperators (resp. defectors). Specifically, $\mathbf{x}^{(t)} = \{x^{(t)}, y^{(t)}\}$ denotes the population profile at time t . Hence, the income of the third party can be expressed as

$$I^{(t)} = M^{(t)}(c_0 + x^{(t)}y^{(t)} F_{CD} + (y^{(t)})^2 F_{DD}). \quad (2.1)$$

Remark 2.1. When implementing pure reward incentives, $M = N$. However, punishment can eliminate agents who fail to pay the fines or commission fee. Thus, $M \leq N$ when implementing punishment.

The cost of incentives is also composed of two parts: (1) the rewards assigned to the participants [52], (2) the cost for imposing the fine [53]. The reason for not considering the cost for imposing the reward is that, in real-world practice, it is usually the agents self-reporting their good behaviour [54], which does not count as the major cost of the third party, hence the cost of rewarding is simplified as the endowed reward. While the cost of enforcing punishment, which relies on detecting and monitoring [7], can be different. Here, we assume the punishment cost is proportional to the probability and the severity of the fine [21], let α ($\alpha \geq 0$) denote the unit cost of punishment for the third party. The value of α depends on the specific cost of enforcement, without loss of generality, we assume $\alpha = 0.3$ as [55]. Accordingly, the expenditure of the third party is defined as:

$$E^{(t)} = (x^{(t)})^2 M^{(t)} R_{CC} + x^{(t)}y^{(t)} M^{(t)} R_{CD} + \alpha M^{(t)} \left(x^{(t)}y^{(t)} F_{CD} + (y^{(t)})^2 F_{DD} \right) \quad (2.2)$$

The wealth of the third party at time t , $W_T^{(t)}$ is given as

$$W_T^{(t)} = I^{(t)} - E^{(t)}, \quad (2.3)$$

and the accumulated wealth of the third party is

$$W_T = \int W_T^{(t)} dt. \quad (2.4)$$

The explicit formula of W_T is given in part 3 of the electronic supplementary material. W_T is required to be non-negative for sustainable incentives.

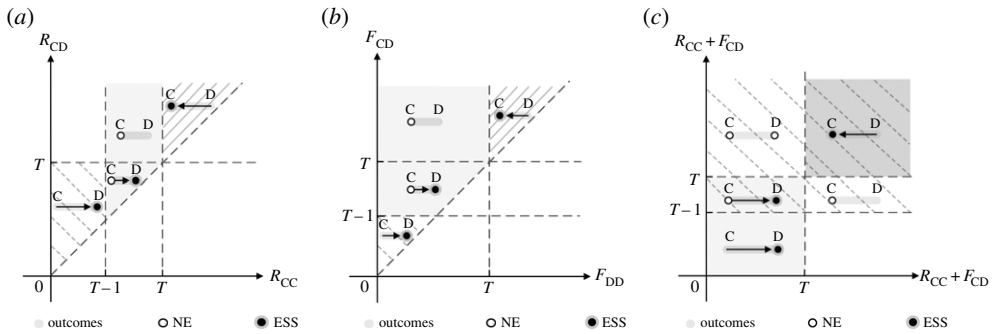


Figure 1. Equilibrium under (a) pure reward, (b) pure punishment and (c) mixed incentives. The requirement for $x^* = 1$ being an NE is that the strength of incentives (R_{CC} , F_{CD} or $R_{CC} + F_{CD}$) being greater than $T - 1$. Further, when the strength (R_{CC} , F_{DD} or $R_{CD} + F_{DD}$) is greater than T , $x^* = 1$ will be an ESS.

3. Analytical results and set-up for simulation experiments

Based on the assumptions described in §2, the population equilibrium with various incentives can be delivered by EGT [56,57]. This section first presents the evolution of the population under different policies, more concretely, the corresponding thresholds of the incentives for achieving the Nash equilibrium (NE) and the evolutionary stable strategy (ESS). It then elaborates the design of the simulation experiments based on the derived analytical results.

(a) Analytical results

The three fixed points are $x^* = 0$, $x^* = 1$, $x^* = (T - R_{CD} - F_{DD}) / (1 + R_{CC} - R_{CD} + F_{CD} - F_{DD})$. In the electronic supplementary material, tables S1–S3 exhibit the requirements for x^* being an NE or ESS under various incentives. Figure 1a–c visualize the results of the population equilibrium by different incentives, concrete derivations are given in part 1 of the electronic supplementary material.

Figure 1a exhibits the results of pure reward incentives ($R_{CD} \geq R_{CC}$ always exists). It can be observed that if $R_{CC} \geq T$, $x^* = 1$ will be the ESS. While given $R_{CC} \leq T - 1$, the ESS will be $x = 0$. When $T \geq R_{CC} \geq T - 1$, $x^* = 1$ is the NE, but the strategy is not robust under this circumstance, which means the cooperation strategy can be invaded by mutants.

Figure 1b shows the results of pure punishment incentives ($F_{CD} \geq F_{DD}$ always exists). When $F_{CD} \geq T - 1$, $x^* = 1$ is always an NE, while $F_{DD} > T$, the $x^* = 1$ becomes an ESS. When the strength of punishment is too light, $F_{CD} \leq T$, then $x = 0$ is the ESS. The practical meaning is that the punishment should be at least greater than T which is the temptation that can be gained from defection, otherwise the agents will have the motivation to become defectors.

Figure 1c shows a more complicated scenario where mixed incentive is applied. What matters to the equilibrium is $R_{CC} + F_{CD}$ (resp. $R_{CD} + F_{DD}$). The sum of the reward and punishment is essentially the difference of the expected pay-off for cooperators and defectors when facing a cooperator (resp. defector). In figure 1c, it is clear that as long as $R_{CD} + F_{DD}$ is less than T , then $x^* = 0$ is always an NE, while only when $R_{CC} + F_{CD}$ is also less than T , $x^* = 0$ can it be robust to mutations. This result indicates that only when the pay-off differences of being a cooperator and being a defector are both less than T , defecting strategy is the ESS. In addition, if $R_{CC} + F_{CD}$ is greater than $T - 1$, $x^* = 1$ becomes the NE, and when both $R_{CC} + F_{CD}$ and $R_{CD} + F_{DD}$ are greater than T , $x^* = 1$ becomes the ESS.

(b) Simulation experiment description

As previously stated, for a specific market or community in practical life, the amount of participants is always finite. Thus, in our simulation experiments, we regulate the initial size of

Table 2. The population evolving algorithm under the pairwise interaction model.

Input:	$N, c_0, \alpha, \beta, R, P, T, S$, incentive mechanism parameters $R_{CC}, R_{CD}, F_{CD}, F_{DD}$, initial population profile $x^{(0)}$, initial accumulated wealth of the third party $W_T^{(0)}$, initial accumulated wealth of all the participants $W_A^{(0)}$, and the total observation time step T^2 .
Output:	Evolution of $x^{(t)}, W_T, W_A$.
Step 1:	If $t < T$, compute the amount of cooperators and defectors based on $x^{(t)}$, then randomly match individuals in pairs, generate the pairwise table . Else go to step 6 .
Step 2:	Based on the pairwise table, calculate the real pay-off of each individual, and generate the pay-off table . If and only if the agent's wealth cannot cover the commission fee c_0 or the fine (F_{CD} or F_{DD}), the agent will be eliminated.
Step 3:	Calculate the income and expense of the third party at time step t , update W_T . Calculate the average pay-off of cooperators, update $\pi(C)$. Calculate the average pay-off of defectors, update $\pi(D)$.
Step 4:	Calculate $x^{(t+1)}$ according to equation (3.1).
Step 5:	Go to step 1.
Step 6:	End.

²We denote the termination time by T , and denote a specific time step by t .

the market. Meanwhile, as participants are of bounded rationality in the real world, which brings more uncertainties to the effect and the sustainability of the incentives, hence the results under different levels of participants' rationality are analysed.

(i) Algorithm: the evolution of population

A Monte Carlo simulation experiment is leveraged to observe the accumulated wealth of the third party and of the participants. This section mainly introduces the algorithm and the design of the simulation experiment.

We assume initially N agents take part in a free market, and each agent has an initial wealth $w_A^{(0)}$, the total initial wealth of participants $W_A^{(0)} = Nw_A^{(0)}$. Players first pay the commission fee c_0 before playing the PDG pairwise. Their wealth then gets updated based on the pay-off matrix and the implemented incentive. Players who cannot afford the commission fee or fine will be eliminated from the market. Their population profile $x = \{x, y\}$ will evolve with the following dynamic mechanism: let $\pi(C)$ (resp. $\pi(D)$) denote the average pay-off of cooperating (resp. defecting) strategy. Based on the EGT framework, the dominant strategy can have more next generations. With probability $p_1 = 1/(1 + \exp^{-\beta(\pi(C) - \pi(D))})$, the defector will imitate the cooperating strategy, and the cooperator with $p_2 = 1/(1 + \exp^{-\beta(\pi(D) - \pi(C))})$ will adopt a defecting strategy [58], β ($\beta \in [0, \infty)$) denotes selection intensity, which represents the rational level of participants [59,60]. A larger β indicates a more rational participant, and if $\beta = 0$ the participant will choose to be a collaborator or defector randomly. Consequently, x will be updated as

$$x^{(t+1)} = (1 - x^{(t)}) p_1 + x^{(t)} (1 - p_2). \quad (3.1)$$

Then at time step $t + 1$, the new generation of participants will be matched in pairs again to have another round of interaction. In the evolutionary process, we observe $x^{(t)}$ to evaluate cooperation level, the accumulated wealth of the third party W_T to measure the sustainability, and the accumulated wealth of all the participants W_A to represent the affluence of the market. The pseudo code of the simulation algorithm is shown in table 2.

Table 3. Simulation experiment set-up.

model parameters	symbol	value
number of initial participants	N	100
commission fee	c_0	0.5
cost-related coefficient	α	0.3
rational level of participants	β	[1, 2, 4]
mutual cooperation payoff	R	1
mutual defection payoff	P	0
temptation payoff	T	2
sucker's payoff	S	-2
population profile	$x^{(0)}$	{0.25, 0.75}
initial wealth of participants	$W_A^{(0)}$	1000
initial wealth of the third party	$W_T^{(0)}$	1000

The settings of the parameters related to the market feature are presented in table 3. With these settings, the population evolving algorithm can be implemented to observe the effect of various incentives.

(ii) Parameter settings for incentives

For observing the performance of the different incentive policies, three series of simulation experiments are designed: pure reward, pure punishment and mixed incentives.

- **Pure reward:** Based on the analytical results in §3a, only when $R_{CC} \geq T - 1$ and $R_{CD} \geq T$, can $x^* = 1$ be an NE or ESS. In simulation experiments $T = 2$, and we set $R_{CC} = 1 + 0.25i$, $R_{CD} = 2 + 0.25i$ where $i \in \mathbb{N}, i \leq 8$.
- **Pure punishment:** Similarly, we set $F_{CD} = 2 + 0.25i$, and $F_{DD} = 0 + 0.25i$, $i \in \mathbb{N}, i \leq 8$.
- **Mixed:** For mixed incentives, we set both $R_{CC} + F_{CD}$ and $R_{CD} + F_{DD}$ varying from 1 to 3 and 2 to 4, respectively, with the increment of 0.25. There is no unique standard to set the rate of R_{CC} in $R_{CC} + F_{CD}$ (resp. F_{DD} in $R_{CD} + F_{DD}$). The rate depends on to what extent the system focuses on positive or negative incentive. In these experiments, to guarantee that $R_{CD} \geq R_{CC}$ and $F_{CD} \geq F_{DD}$, while minimizing the difference between rewarding and punishment, we set the rate to 0.2 (the proof of the rate, and the concrete settings for these four parameters are given in part 2 of the electronic supplementary material).

4. Experimental results and interpretation

This section first shows the effect of incentives on promoting cooperation, then elaborates the influence on the wealth of different parties, and finally discusses the sustainability of the market under these incentives.

(a) Effect of incentives on cooperators' population and market size

In a market with finite size, with the same population profile x , the expected pay-off of different strategies is not entirely fixed, due to the different possible pairs. Thereby, unlike the analytical results shown in figure 1, the dynamics of x exhibits a chaotic behaviour, fluctuating over time, but does not reach a fixed value. Figure 2 represents the evolution of x in the market contains 100 participants. Notwithstanding the fact that there is no strict 'stable state' that can be reached, the determinism of x can be assessed by computational means [61]. With $T = 500$, $x^{(t)}$ exhibits a

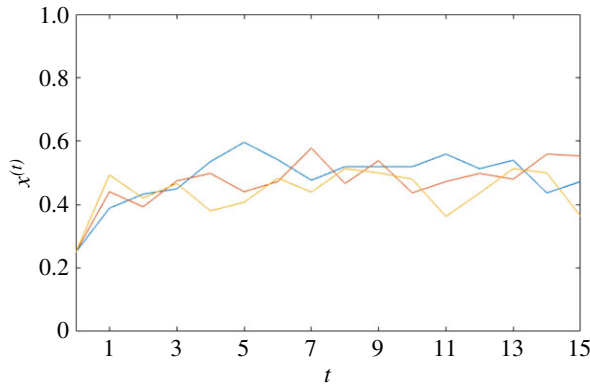


Figure 2. The dynamics of $x^{(t)}$ in 15 time steps of three repeating experiments. Parameters, $x^{(0)} = 0.25$, $\beta = 1$, $R_{CC} = 1$, $R_{CD} = 2$. (Online version in colour.)

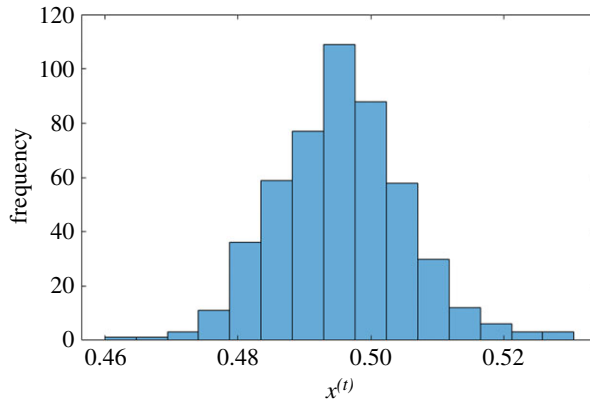


Figure 3. The frequency distribution of $x^{(t)}$ from time step 5 to 500. Parameters, $x^{(0)} = 0.25$, $\beta = 1$, $R_{CC} = 1$, $R_{CD} = 2$. (Online version in colour.)

normal distribution as shown in figure 3. Since the expectation of $x^{(t)}$, $E(x^{(t)})$ always exists, and depends on the incentive, the mean value of x^t is chosen to reflect the effect of incentives on the population profile. Further, to determine the termination time \mathbf{T} , T -test was applied to trials with $\mathbf{T} = [350, 300, 250, 200, 150, 100, 50, 30, 20]$, $E(x^{(t)})$ ($t \in [5, \mathbf{T}]$) is required to have no statistical significant difference with $E(x^{(t)})$ ($t \in [5 : 500]$), p -value is greater than 0.05. After trials, \mathbf{T} is finally ensured to be 30.

The heat map in figure 4 shows the expectation of $x^{(t)}$ under various incentives, and the heat indicates the expected rate of cooperators in the market. As can be observed, the stronger incentives lead to higher expected $x^{(t)}$. In addition, compare plots horizontally, as β increases, the expected $x^{(t)}$ becomes higher with respect to the same strength of the incentive. That is due to the fact that when participants are more rational, the selection will depend more on the payoff.

However, not all the subplots are in line with these rules: in subplot (f) of figure 4, the outcome of pure punishment with $\beta = 4$ is different. This fact is associated with the size of the market M . To better elaborate this abnormal phenomenon, we draw a bubble plot to outline $M^{(30)}$ over the heat map, and the area of the bubble indicates the value of $M^{(30)}$. From (a–c) and (g–i), it can be observed that under pure reward incentives or mixed incentives, the market size is stable, and there are not many agents eliminated. While under pure punishment incentives, the size of the market changed dramatically, varying from 0 to 100. The abnormal cells shown in (f)

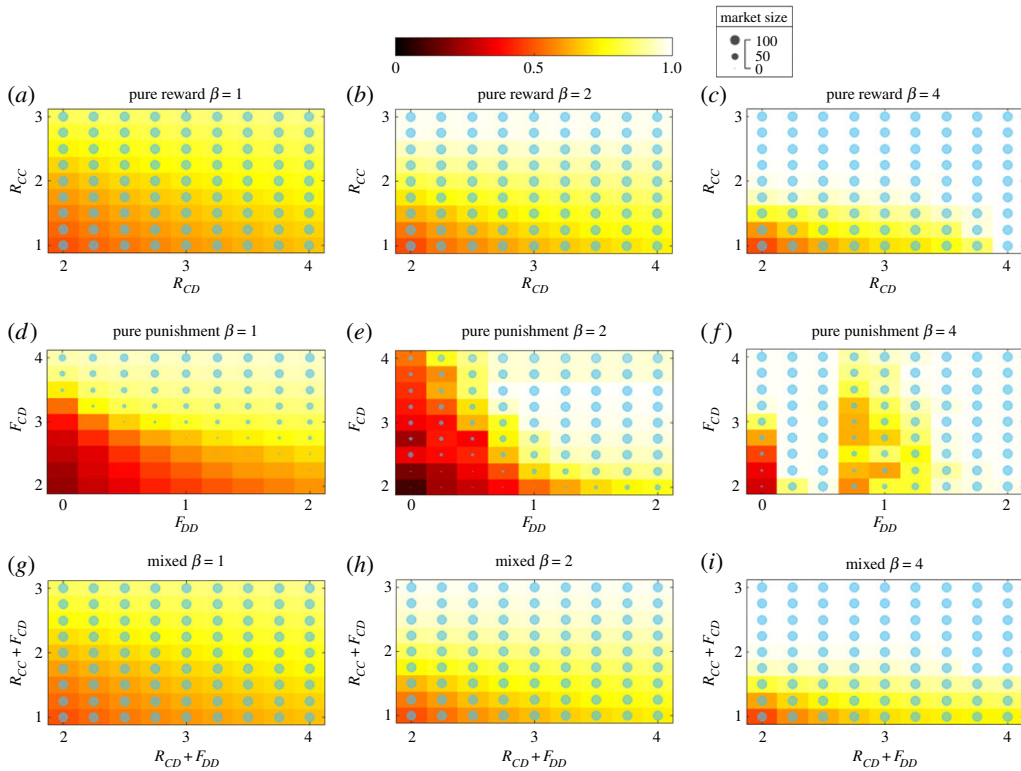


Figure 4. The expectation of $x^{(t)}$ and $M^{(30)}$ under different incentives. The heat represents the value of $x^{(t)}$, the size of the bubble represents the market size. Subplots (a–c), (d–f) and (g–i) represent the results under pure reward, pure punishment and mixed incentives, respectively. Incentives with higher strength are generally corresponding to higher $x^{(t)}$. Exceptions exhibit when the size of the market shrinks, which is caused by too light punishments. A higher rational level of participants also contributes to higher cooperation level. (Online version in colour.)

correspond to the situations where the size of the market shrinks a lot, $x^{(t)}$ no longer strongly depends on the incentives, but on the choice of few bounded rational participants. An *ad hoc* example would be that a market with $M^{(t)} = 4$, if only one agent chooses defection, $x^{(t)}$ is 75%, then after this agent is eliminated in this round, two of the remaining ones are selected to interact at time step $t + 1$, if only one of them choose to defect, then $x^{(t+1)}$ will drop down to 50%, but if this agent also chooses to cooperate, then $x^{(t+1)}$ will increase to 1. Thus, when M is small, the effect of randomness will counteract the effect of incentives on the results, which leads to the abnormal cells in subplot (f).

Remark 4.1. Under mixed incentives, the size of the market can also shrink a little bit when the incentive is not strong enough, but this result is not completely visible in figure 4. We adjusted the scale in figure 7 to visualize this pattern in a clearer way.

The size of the market under various incentives presents some interesting patterns. Since M changes a lot under pure punishment incentives, we focus on the result shown on subplots (d–f). Firstly, M gradually becomes larger from the lower triangular part to the upper triangular part in each subplot. It reveals a counterintuitive phenomenon, that M will be larger under more severe punishment. The explanation is, the heavy punishment increases the pay-off difference between cooperators and defectors, hence those bounded rational participants tend to choose cooperation. While under weak punishment, participants tend to choose to defect repeatedly, especially when β is low, hence, they are easier to be eliminated from the market. Secondly, the effect of punishment has two aspects. On the one hand, it decreases the pay-off of defectors;

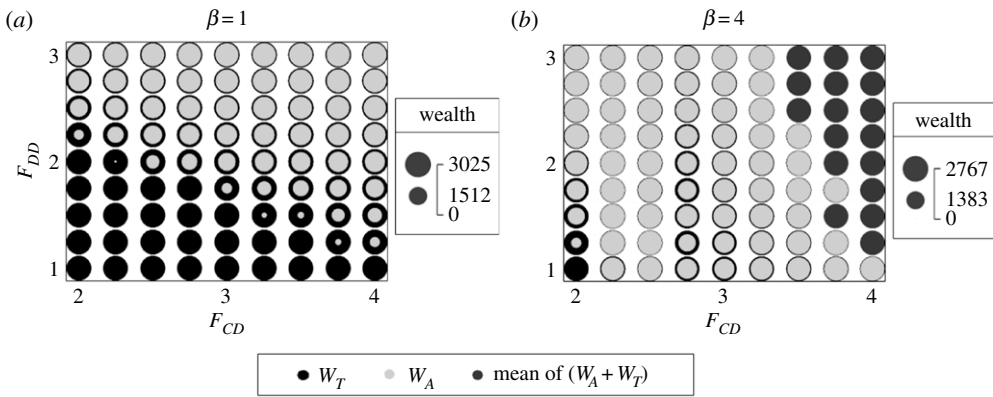


Figure 5. The accumulated wealth of the third party (W_T , represented by black bubbles) and of the agents (W_A , represented by light grey bubbles) under pure punishment incentives. W_T is always greater than W_A . The difference between W_A and W_T diminishes as the punishment becomes heavier, especially when participants enjoy a higher rational level ($\beta = 4$).

when it becomes heavier, the pay-off difference between cooperators and defectors gets larger, and thus can promote cooperation and maintain M . On the other hand, if the punishment is moderate, it can easily eliminate the agents who repeatedly choose to defect, and thus make M drop abnormally. In subplot(f), when $F_{DD} \in [0, 0.5]$, the influence of promoting cooperation is dominant, M climbs as F_{DD} increases. While when $F_{DD} \in [0.75, 1.5]$, the influence of elimination becomes dominant, and causes a large decrease in M . But when F_{DD} becomes heavier, the effect on promoting cooperation becomes dominant again and M increases again.

The experimental results in figure 4 indicate: (1) as a general rule, the expected $x^{(t)}$ increases as the incentive becomes stronger, and this effect is more obvious with higher β . However, under pure punishment incentives, the shrinking market might involve low $x^{(t)}$; (2) in terms of the market size M , under pure reward as well as mixed incentives, M is stable, while under pure punishment incentives, counterintuitively, M increases as the incentives go heavier.

(b) Effect of incentives on the accumulated wealth of different parties

The parties in the market include the third party who implements the incentives, as well as the agents who join the market. Both of these two parties can gain or lose utilities in the market. This section aims at analysing the accumulated wealth of these two parties. Since under pure reward incentives it is always the third party subsidizing the cooperators, it can be expected that the accumulated wealth of the third party, W_T , decreases until the third party goes bankrupt, and the accumulated wealth of the agents, W_A , keeps increasing. By contrast, for pure punishment and mixed incentives, the relative size of W_T and W_A can exhibit interesting behaviours, hence, we focus on the results of pure punishment and mixed incentive.

The accumulated wealth of the agents W_A and the third party W_T is represented by bubble plots (figures 5 and 6), and bubble areas represent the amount. Note that the smaller bubbles are drawn over the larger ones, so that the relative difference of the income of the two parties can be observed easily. Different parties are distinguished by colours, black links to W_T and light grey links to W_A . Particularly, when the difference between W_A and W_T is small ($|\log(W_A) - \log(W_T)| \leq 0.35$), the difference is hardly visible, and we thus use the grey colour to represent $(W_A + W_T)/2$.

Figures 5 and 6 reveal that W_A increases as the incentives become stronger. As for W_T , under pure punishment incentives, W_T is always greater than W_A , as the income of the third party contains the retrieved fine. Nevertheless, as the incentives become stronger, the population of the defector decreases, and the income of the third party will thus drop. Accordingly, the difference between W_A and W_T declines as shown in figure 5. This phenomenon becomes more visible as β

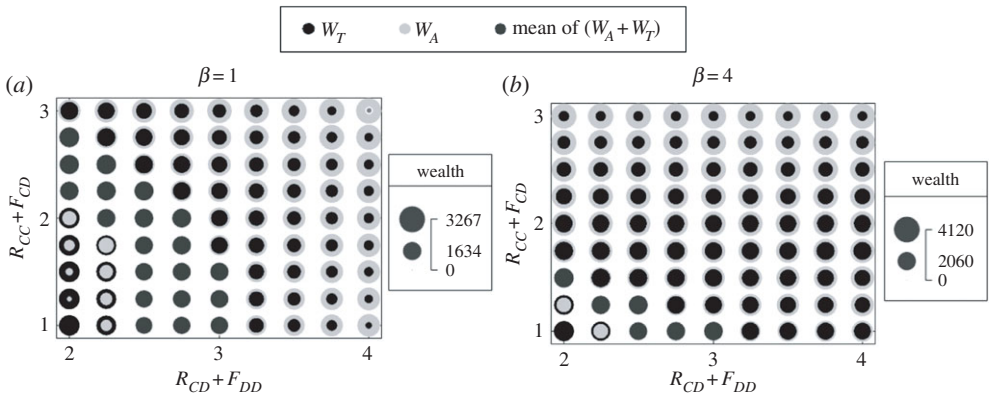


Figure 6. The accumulated wealth of the third party (W_T) and of the agents (W_A) under mixed incentives. There is a trade-off between W_A and W_T , W_A decreases while W_T increases as the incentives become stronger. Meanwhile, Pareto optimization space exists for increasing $W_A + W_T$. In addition, when participants are more rational ($\beta = 4$), both the upper limitations of W_A and W_T increase.

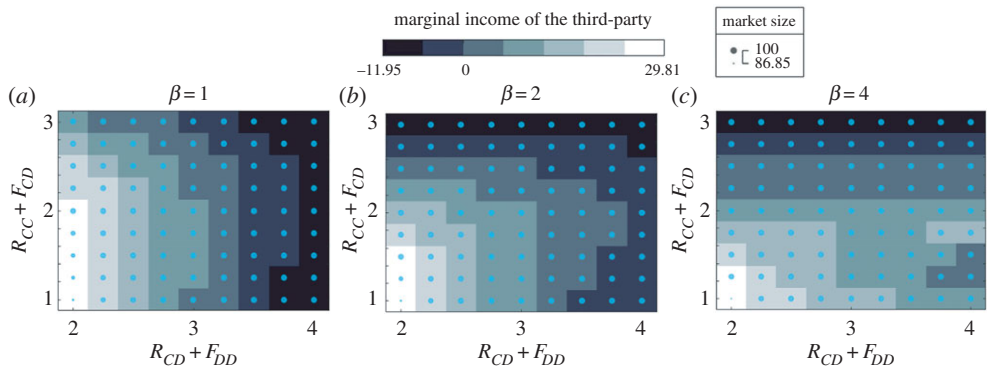


Figure 7. Sustainability of mixed incentives, $x^{(0)} = 0.25$. The colours of the cells represent the marginal income of the third party. The result shows that the marginal income decreases as the incentive become stronger, while with different β , the sustainable incentives can be different. Under mixed incentives, the market size is always greater than 86. (Online version in colour.)

is greater, which is because having more rational participants means being easier to achieve $x^* = 1$ under the same incentives, which reduces the fine-based income of the third party. Also, for this reason, the upper bound of W_T drops as β increases.

Under mixed incentives, the pattern becomes more complicated as shown in figure 6, W_T declines as the incentives become stronger, while W_A exhibits an opposite trend. It can be observed that W_A catches up with W_T and surpasses it as mixed incentive becomes stronger. In addition, this figure exhibits an obvious trade-off between W_A and W_T . Yet, we can also note the Pareto optimization exists when the incentives are of moderate strength under mixed incentives, when the sum of the accumulated wealth of these two parties is maximized.

When comparing subplots (a) and (b) in figure 6, the relative size of W_T with W_A depends on both the strength of the incentives and β . Under low β as shown in subplot (a), when $R_{CD} + F_{DD}$ is small, W_T is much greater than W_A , while as $R_{CD} + F_{DD}$ increases, W_T decreases and W_A increases. When $R_{CD} + F_{DD}$ is greater than 3, W_A becomes obviously greater than W_T . However, under high β as shown in subplot (b), it is mainly $R_{CC} + F_{CD}$ that depends on the trend of W_T and W_A . When $R_{CC} + F_{CD}$ becomes stronger, W_T decreases. Especially, when $R_{CC} + F_{CD} \geq 1.75$,

W_T become lower than W_A , as $R_{CC} + F_{CD}$ increases further, the difference of these two parties enlarges.

This pattern is connected to β : higher β indicates more rational participants, under the same incentives, and x is easier to evolve to 1. Thus, with a high β , as the mixed incentive becomes stronger, more cooperators will be in the market, which rapidly increases the cost of $R_{CC} + F_{CD}$. That is why W_T decreases dramatically as $R_{CC} + F_{CD}$ increases. However, when β is low, the defectors always exist in the market, as $R_{CD} + F_{DD}$ becomes stronger, not only the management cost R_{CD} increases, the income from F_{DD} decreases (lower y induced by the stronger incentive). Consequently, with low β , $R_{CD} + F_{DD}$ dominates the change of the wealth.

Summarizing the effect of incentives on W_A and W_T , the findings show that under pure punishment incentives, $W_T > W_A$, but $|W_T - W_A|$ will reduce as the punishment become heavier; under mixed incentives, W_T decreases as the mixed incentive becomes stronger, while W_A presents an opposite trend. It is possible to improve $W_T + W_A$ by choosing the moderate strength incentives.

(c) The sustainability of the mixed incentives

In this section, the slope of W_T is applied to evaluate the sustainability, as it can represent the marginal income of the third party and predict the trend of W_T . An incentive is regarded as sustainable if this value is positive. For pure reward incentives, the third party always gives rewards to the agents, W_T is always decreasing, and by their nature they cannot be sustainable in the long term. Pure punishment incentives are on the contrary always sustainable, as the third party can collect a fine from the agents in addition to the commission fee. Thus, in §4c, we only discuss the result for mixed incentives.

Subplots (a–c) of figure 7 show that as the mixed incentives become stronger, the slope of W_T turns from positive to negative, which indicates that those lighter incentives enjoy lower implementation cost, and can therefore provide the third party with more marginal income. The strength of the reward or of the punishment is negatively related to the marginal income of the third-party. Light mixed incentives may also result in M shrinking, but M is always greater than 86.

When comparing subplots (a–c) of figure 7 horizontally, the value of β has a strong influence on the sustainability of mixed incentives. When $\beta = 1$, participants are less rational, the market is always mixed with defectors, as indicated by the subplot (g) of figure 4. If the incentives become stronger, the cost of R_{CD} will increase. It explains why when $R_{CC} + F_{CD}$ is greater than 3, the slope becomes negative, and incentives become unstable. When $\beta = 4$, more rational participants will choose to be cooperators, which increases the cost of R_{CC} . Thus, when $R_{CC} + F_{CD}$ is greater (≥ 2.75), incentives become unsustainable as shown in subplot (c) of figure 7. For subplot (b), when $\beta = 2$, it combines the feature of subplots (a) and (c).

In summary, the results of the three series of simulation experiments reveal that, after setting up the initial market size and introducing the elimination mechanism, various effects on the market's collaboration level, on the affluence of different parties, and on their own sustainability emerge. Generally, x increases under stronger incentives. Compared with pure reward, pure punishment incentives potentially eliminate participants. It is also worth mentioning that the heavier punishment is beneficial for preventing M from shrinking and promoting W_A , which can be considered counterintuitive. Whereas under mixed incentives, the wealth of these two parties shows the opposite trend, and we can always choose moderate strength incentives to maximize $W_A + W_T$. In terms of sustainability, pure reward cannot be sustainable, but pure punishment is always sustainable, mixed incentives' sustainability depends on both β and the strength of the incentive.

5. Conclusion and discussion

During the last decades, there have been tremendous efforts to design proper incentives. Although in the field of political economics, the sustainability of incentives is considered to be

an important criterion in evaluating the real-world incentives or policies [62], in the theoretical exploration, sustainability is not always considered, especially when the incentives are carried out by the third party (or external decision-maker [26]). Our work, from the institutional perspective, considers: (1) the sustainability of incentives by evaluating the enforcement cost and the income of the third party; and (2) the effect of incentives on the market's affluence by introducing the elimination mechanism. Hence, the evaluation criteria of incentives include three aspects: cooperation level, sustainability, and the influence on promoting market affluence. The motivation for such consideration is twofold. First, we believe that when designing incentives, it is critical to consider sustainability at the institutional level to ensure the practicality of incentives. Secondly, the market can be considered as an ecological system; a flourishing market not only depends on the population of cooperators, but also on the size of the market, and the accumulated wealth of all parties.

Our simulation experiments show that a pure reward is unsustainable, driven by the fact that the third party is required to continuously provide high subsidies. However, this conclusion might be different under different assumptions for the income of the third party, for example, Sasaki and Uchida proposed a sustainable pure reward incentive based on volunteer reward pool, which assumes that the fund in rewarding pool enjoys an interest rate, and will be shared by rewarders and cooperators [63]. In our work, we focus on institutional enforced incentives; the fund for rewarding incentives is purely gathered from the commission fee, which has no interest rate, and thus, the rewarding pool will not increase spontaneously. Further, to reach a high cooperating level, we assume the reward has to be greater than T , which indicates a high cost of rewarding. These two different assumptions lead to the opposite conclusions on the sustainability of pure reward. The seemingly contradictory results illustrate the subtlety of pure reward incentives, which depends on the nature of the specific system.

For the case of pure punishment, there are a few published papers which explored the requirements for sustainable pool punishment, for example, Matjaž *et al.* pointed out that the punishment for the second-order free-riders can lead to sustainable pool-punishment in a population structured public goods game [22,24]; Sarah *et al.* relaxed this requirement by introducing the signaling effect of participants knowing whether a punishment institution was established [23]. Different from the traditional pool punishment, Lee, Colin and Szolnoki proposed hiring the mercenary punishers from players with the collected tax to counteract the second-order free-riders [25]. The main difference to these pool or mercenary punishment models is that we assume the third party cannot be composed of players, considering the role of manager or regulator of the market is usually fixed by a group of professional agents. In our work, the punishing pool is composed of the collected commission fee and the retrieved fine, and the enforcement cost is a rate α of fine ($\alpha = 0.3$ [55]), these settings make the costly punishment sustainable. Besides, by introducing the elimination mechanism, we are able to study the impact of the incentive on the size of the market. We find that with the introduction of elimination mechanism, the size of the market is shrinkable. Light punishments can even lead to the collapse of the market, yet more severe punishment can maintain the market size better, and thus promote the affluence of the participants as well as of the third party in the long term.

A few previous studies have explored how mixed incentives should be combined to improve cooperation at an institutional level. Chen *et al.* found that the switch of incentive from reward to punishment based on the population of cooperators can effectively promote cooperation [28]. Fang *et al.* pointed out that mixed incentives which enjoy synergistic effects perform better on promoting cooperation [64]. In our study, we observe that mixed incentives can not only lead to a high cooperation level, but also have some advantages in terms of sustainability and affluence compared with pure incentives. More specifically, mixed incentives perform better than pure reward in terms of sustainability, whereas stronger mixed incentives usually correspond to higher implementation costs. Whether the incentive is sustainable depends on the level of rationality of the participants, as well as the strength of the incentive. Compared with pure punishment, mixed incentives can always maintain the size of the market ($M > 86$) and prevent the market from collapsing, which guarantees W_A . Furthermore, the experimental results reveal the trade-off

between the wealth of the third party (W_T) with the agents (W_A). As the incentives become stronger, W_T decreases and W_A increases, while in moderate strength, we can improve $W_T + W_A$ and balance the affluence level of the two parties. Thus, to a certain extent, mixed incentives combine the advantages of pure incentives.

For future work, our model could be extended by applying the evaluation criteria of incentives into more complicated scenarios. For example, in our model the punishment or reward always happens, while in real life probabilistic sanctioning can be an alternative way to reduce the cost of incentive implementation [65,66]. Correspondingly, facing such uncertain punishment or reward, the players might not necessarily be risk neutral. They can be risk-averse facing punishments [30] or risk-seeking facing rewards [67]. Such extension of assumptions can adjust the third party's cost in enforcing incentives, and tune the expected pay-off of individuals, hence influence the effect of incentives. Considering the constantly evolving environment, flexible incentives are potentially cost-efficient [28,39,68], evaluating their effect on affluence, and their sustainability is of vital practical importance. By such an extension, we expect more guidance that, with realistic management significance, can be drawn on to deal with these complicated circumstances.

Data accessibility. The code for the simulations in the paper is available at: <https://bitbucket.org/uva-sne/simulation-experiment-code-of-rspa-2022-0393/>. The data are provided in the electronic supplementary material [69].

Authors' contributions. X.Z.: conceptualization, formal analysis, methodology, visualization, writing—original draft, writing—review and editing; A.B.: conceptualization, methodology, project administration, supervision, writing—review and editing; M.H.L.: conceptualization, methodology, validation, writing—review and editing; T.v.E.: conceptualization, funding acquisition, investigation, project administration, supervision, writing—review and editing; C.d.L.: funding acquisition, investigation, project administration, supervision, validation, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This paper builds upon the work done within the Dutch NWO Research project 'Data Logistics for Logistics Data' (DL4LD, www.dl4ld.net), supported by the Dutch Top consortia for Knowledge and Innovation 'Institute for Advanced Logistics' (TKI Dinalog, www.dinalog.nl) of the Ministry of Economy and Environment in the Netherlands and the Dutch Commit-to-Data initiative (<https://commit2data.nl/>), grant no.: 628.009.001.

References

1. Szolnoki A, Perc M. 2010 Reward and cooperation in the spatial public goods game. *EPL* **92**, 38003. (doi:10.1209/0295-5075/92/38003)
2. Wu Y, Chang S, Zhang Z, Deng Z. 2017 Impact of social reward on the evolution of the cooperation behavior in complex networks. *Sci. Rep.* **7**, 41076. (doi:10.1038/srep41076)
3. Hintze A, Adami C. 2015 Punishment in public goods games leads to meta-stable phase transitions and hysteresis. *Phys. Biol.* **12**, 046005. (doi:10.1088/1478-3975/12/4/046005)
4. Gürerk Ö, Irlenbusch B, Rockenbach B. 2006 The competitive advantage of sanctioning institutions. *Science* **312**, 108–111. (doi:10.1126/science.1123633)
5. Przepiorka W, Diekmann A. 2013 Individual heterogeneity and costly punishment: a volunteer's dilemma. *Proc. R. Soc. B* **280**, 20130247. (doi:10.1098/rspb.2013.0247)
6. Tan F, Xiao E. 2018 Third-party punishment: retribution or deterrence? *J. Econ. Psychol.* **67**, 34–46. (doi:10.1016/j.joep.2018.03.003)
7. Balliet D, Mulder LB, Van Lange PA. 2011 Reward, punishment, and cooperation: a meta-analysis. *Psychol. Bull.* **137**, 594. (doi:10.1037/a0023489)
8. Traulsen A, Röhl T, Milinski M. 2012 An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proc. R. Soc. B* **279**, 3716–3721. (doi:10.1098/rspb.2012.0937)
9. Okada I. 2020 A review of theoretical studies on indirect reciprocity. *Games* **11**, 27. (doi:10.3390/g11030027)
10. Wu Y, Zhang B, Zhang S. 2017 Probabilistic reward or punishment promotes cooperation in evolutionary games. *Chaos Solitons Fractals* **103**, 289–293. (doi:10.1016/j.chaos.2017.06.015)

11. Yang Y, Yang W. 2019 Does whistleblowing work for air pollution control in China? A study based on three-party evolutionary game model under incomplete information. *Sustainability* **11**, 324. (doi:10.3390/su11020324)
12. Gu C, Wang X, Zhao J, Ding R, He Q. 2020 Evolutionary game dynamics of Moran process with fuzzy payoffs and its application. *Appl. Math. Comput.* **378**, 125227. (doi:10.1016/j.amc.2020.125227)
13. Roca CP, Cuesta JA, Sánchez A. 2009 Evolutionary game theory: temporal and spatial effects beyond replicator dynamics. *Phys. Life Rev.* **6**, 208–249. (doi:10.1016/j.plrev.2009.08.001)
14. Hofbauer J, Sigmund K. 1998 *Evolutionary games and population dynamics*. Cambridge, UK: Cambridge University Press.
15. Pei H, Yan G, Wang H. 2021 Reciprocal rewards promote the evolution of cooperation in spatial Prisoner's dilemma game. *Phys. Lett. A* **390**, 127108. (doi:10.1016/j.physleta.2020.127108)
16. Jiménez R, Lugo H, Cuesta JA, Sánchez A. 2008 Emergence and resilience of cooperation in the spatial Prisoner's dilemma via a reward mechanism. *J. Theor. Biol.* **250**, 475–483. (doi:10.1016/j.jtbi.2007.10.010)
17. Zhu G, Liu H, Feng M. 2018 Sustainability of information security investment in online social networks: an evolutionary game-theoretic approach. *Mathematics* **6**, 177. (doi:10.3390/math6100177)
18. Sigmund K, Hauert C, Nowak MA. 2001 Reward and punishment. *Proc. Natl Acad. Sci. USA* **98**, 10757–10762. (doi:10.1073/pnas.161155698)
19. Takesue H. 2018 Evolutionary Prisoner's dilemma games on the network with punishment and opportunistic partner switching. *EPL* **121**, 48005. (doi:10.1209/0295-5075/121/48005)
20. Wang Z, Xia CY, Meloni S, Zhou CS, Moreno Y. 2013 Impact of social punishment on cooperative behavior in complex networks. *Sci. Rep.* **3**, 3055. (doi:10.1038/srep03055)
21. Couto MC, Pacheco JM, Santos FC. 2020 Governance of risky public goods under graduated punishment. *J. Theor. Biol.* **505**, 110423. (doi:10.1016/j.jtbi.2020.110423)
22. Perc M, Jordan JJ, Rand DG, Wang Z, Boccaletti S, Szolnoki A. 2017 Statistical physics of human cooperation. *Phys. Rep.* **687**, 1–51. (doi:10.1016/j.physrep.2017.05.004)
23. Schoenmakers S, Hilbe C, Blasius B, Traulsen A. 2014 Sanctions as honest signals—the evolution of pool punishment by public sanctioning institutions. *J. Theor. Biol.* **356**, 36–46. (doi:10.1016/j.jtbi.2014.04.019)
24. Perc M. 2012 Sustainable institutionalized punishment requires elimination of second-order free-riders. *Sci. Rep.* **2**, 1–6. (doi:10.1038/srep00344)
25. Lee HW, Cleveland C, Szolnoki A. 2022 Mercenary punishment in structured populations. *Appl. Math. Comput.* **417**, 126797. (doi:10.1016/j.amc.2021.126797)
26. Cimpeanu T, Perret C, Han TA. 2021 Cost-efficient interventions for promoting fairness in the ultimatum game. *Knowl.-Based Syst.* **233**, 107545. (doi:10.1016/j.knosys.2021.107545)
27. Ohtsubo Y, Masuda F, Watanabe E, Masuchi A. 2010 Dishonesty invites costly third-party punishment. *Evol. Hum. Behav.* **31**, 259–264. (doi:10.1016/j.evolhumbehav.2009.12.007)
28. Chen X, Sasaki T, Brännström Å, Dieckmann U. 2015 First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *J. R. Soc. Interface* **12**, 20140935. (doi:10.1098/rsif.2014.0935)
29. Hintze A, Staudacher J, Gelhar K, Pothmann A, Rasch J, Wildegger D. 2020 Inclusive groups can avoid the tragedy of the commons. *Sci. Rep.* **10**, 1–8. (doi:10.1038/s41598-020-79731-y)
30. Nakamura M. 2019 Rare third-party punishment promotes cooperation in risk-averse social learning dynamics. *Front. Phys.* **6**, 156. (doi:10.3389/fphy.2018.00156)
31. Gao S, Liang J. 2020 Cooperation under institutional incentives with perfect and imperfect observation. *Phys. Lett. A* **384**, 126723. (doi:10.1016/j.physleta.2020.126723)
32. Henrich J *et al.* 2006 Costly punishment across human societies. *Science* **312**, 1767–1770. (doi:10.1126/science.1127333)
33. Rabellino D, Morese R, Ciaramidaro A, Bara BG, Bosco FM. 2016 Third-party punishment: altruistic and anti-social behaviours in in-group and out-group settings. *J. Cogn. Psychol.* **28**, 486–495. (doi:10.1080/20445911.2016.1138961)
34. Gächter S. 2012 Carrot or stick? *Nature* **483**, 39–40.
35. Villatoro D, Andrighetto G, Sabater-Mir J, Conte R. 2011 Dynamic sanctioning for robust and cost-efficient norm compliance. In *Proc. of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, 16-22 July*, pp. 414–419. California: AAAI Press.

36. Helbing D, Szolnoki A, Perc M, Szabó G. 2010 Punish, but not too hard: how costly punishment spreads in the spatial public goods game. *New J. Phys.* **12**, 083005. (doi:10.1088/1367-2630/12/8/083005)
37. Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K. 2007 Via freedom to coercion: the emergence of costly punishment. *Science* **316**, 1905–1907. (doi:10.1126/science.1141588)
38. Han TA, Tran-Thanh L. 2018 Cost-effective external interference for promoting the evolution of cooperation. *Sci. Rep.* **8**, 1–9. (doi:10.1038/s41598-018-34435-2)
39. Wang S, Chen X, Szolnoki A. 2019 Exploring optimal institutional incentives for public cooperation. *Commun. Nonlinear Sci. Numer. Simul.* **79**, 104914. (doi:10.1016/j.cnsns.2019.104914)
40. Duong MH, Han TA. 2021 Cost efficiency of institutional incentives for promoting cooperation in finite populations. *Proc. R. Soc. A* **477**, 20210568. (doi:10.1098/rspa.2021.0568)
41. Ohdaira T. 2021 Cooperation evolves by the payoff-difference-based probabilistic reward. *Eur. Phys. J. B* **94**, 1–8. (doi:10.1140/epjb/s10051-021-00239-z)
42. Wang S, Liu L, Chen X. 2022 Incentive strategies for the evolution of cooperation: analysis and optimization. *EPL* **136**, 68002. (doi:10.1209/0295-5075/ac3c8a)
43. Conlisk J. 1996 Why bounded rationality? *J. Econ. Lit.* **34**, 669–700.
44. Nash J. 1951 Non-cooperative games. *Ann. Math.* **54**, 286–295. (doi:10.2307/1969529)
45. Mieth L, Buchner A, Bell R. 2021 Moral labels increase cooperation and costly punishment in a Prisoner's dilemma game with punishment option. *Sci. Rep.* **11**, 1–13. (doi:10.1038/s41598-021-89675-6)
46. Szolnoki A, Chen X. 2021 Cooperation and competition between pair and multi-player social games in spatial populations. *Sci. Rep.* **11**, 1–9. (doi:10.1038/s41598-021-91532-5)
47. Kanazawa T, Fukumoto Y, Ushio T, Misaka T. 2009 Replicator dynamics with Pigovian subsidy and capitation tax. *Nonlinear Anal.: Theory Methods Appl.* **71**, e818–e826. (doi:10.1016/j.na.2008.11.072)
48. Wang S, Liu L, Chen X. 2021 Tax-based pure punishment and reward in the public goods game. *Phys. Lett. A* **386**, 126965. (doi:10.1016/j.physleta.2020.126965)
49. Zhou X, Chen K, Wen H, Lin J, Zhang K, Tian X, Wang S, Lev B. 2021 Integration of third-party platforms: does it really hurt them? *Int. J. Prod. Econ.* **234**, 108003. (doi:10.1016/j.ijpe.2020.108003)
50. Estache A. 2020 Institutions for infrastructure in developing countries: what we know. . . and the lot we still need to know. In *The handbook of economic development and institutions*, pp. 634–688. Princeton, NJ: Princeton University Press.
51. Li J, Liu Y, Wang Z, Xia H. 2021 Egoistic punishment outcompetes altruistic punishment in the spatial public goods game. *Sci. Rep.* **11**, 1–13. (doi:10.1038/s41598-021-85814-1)
52. Alders P, Schut FT. 2019 The 2015 long-term care reform in the Netherlands: getting the financial incentives right? *Health Policy* **123**, 312–316. (doi:10.1016/j.healthpol.2018.10.010)
53. Polinsky AM, Shavell S. 2000 The economic theory of public enforcement of law. *J. Econ. Lit.* **38**, 45–76. (doi:10.1257/jel.38.1.45)
54. Friesen L, Gangadharan L. 2013 Designing self-reporting regimes to encourage truth telling: an experimental study. *J. Econ. Behav. Org.* **94**, 90–102. (doi:10.1016/j.jebo.2013.08.007)
55. Quan J, Liu W, Chu Y, Wang X. 2017 Stochastic evolutionary voluntary public goods game with punishment in a quasi-birth-and-death process. *Sci. Rep.* **7**, 1–14. (doi:10.1038/s41598-017-16140-8)
56. Crawford VP. 1989 Learning and mixed-strategy equilibria in evolutionary games. *J. Theor. Biol.* **140**, 537–550. (doi:10.1016/S0022-5193(89)80113-4)
57. Gintis H *et al.* 2000 *Game theory evolving: a problem-centered introduction to modeling strategic behavior*. Princeton, NJ: Princeton University Press.
58. Wardil L. 2013 The evolution of cooperation in mixed games. *Chaos Solitons Fractals* **56**, 160–165. (doi:10.1016/j.chaos.2013.07.018)
59. Szabó G, Tóke C. 1998 Evolutionary Prisoner's dilemma game on a square lattice. *Phys. Rev. E* **58**, 69. (doi:10.1103/physreve.58.69)
60. Szolnoki A, Vukov J, Szabó G. 2009 Selection of noise level in strategy adoption for spatial social dilemmas. *Phys. Rev. E* **80**, 056112. (doi:10.1103/PhysRevE.80.056112)
61. Leon F. 2014 Design and evaluation of a multiagent interaction protocol generating behaviours with different levels of complexity. *Neurocomputing* **146**, 173–186. (doi:10.1016/j.neucom.2014.04.058)

62. Andrews J, Borgerhoff Mulder M. 2018 Cultural group selection and the design of REDD+: insights from Pemba. *Sustainab. Sci.* **13**, 93–107. (doi:10.1007/s11625-017-0489-2)
63. Sasaki T, Uchida S. 2014 Rewards and the evolution of cooperation in public good games. *Biol. Lett.* **10**, 20130903. (doi:10.1098/rsbl.2013.0903)
64. Fang Y, Benko TP, Perc M, Xu H, Tan Q. 2019 Synergistic third-party rewarding and punishment in the public goods game. *Proc. R. Soc. A* **475**, 20190349. (doi:10.1098/rspa.2019.0349)
65. Liu L, Chen X, Szolnoki A. 2019 Evolutionary dynamics of cooperation in a population with probabilistic corrupt enforcers and violators. *Math. Models Methods Appl. Sci.* **29**, 2127–2149. (doi:10.1142/S0218202519500428)
66. Chen X, Szolnoki A, Perc M. 2014 Probabilistic sharing solves the problem of costly punishment. *New J. Phys.* **16**, 083016. (doi:10.1088/1367-2630/16/8/083016)
67. Zeng W, Li M, Chen F. 2016 Cooperation in the evolutionary iterated Prisoner's dilemma game with risk attitude adaptation. *Appl. Soft Comput.* **44**, 238–254. (doi:10.1016/j.asoc.2016.03.025)
68. Sun W, Liu L, Chen X, Szolnoki A, Vasconcelos VV. 2021 Combination of institutional incentives for cooperative governance of risky commons. *Iscience* **24**, 102844. (doi:10.1016/j.isci.2021.102844)
69. Zhou X, Belloum A, Lees MH, van Engers T, de Laat C. 2022 Costly incentives design from an institutional perspective: cooperation, sustainability and affluence. Figshare. (doi:10.6084/m9.figshare.c.6168135)