

Li, Xueheng; Molleman, Lucas; Van Dolder, Dennie

**Working Paper**

## Conditional punishment: Descriptive social norms drive negative reciprocity

CeDEx Discussion Paper Series, No. 2020-05

**Provided in Cooperation with:**

The University of Nottingham, Centre for Decision Research and Experimental Economics (CeDEx)

*Suggested Citation:* Li, Xueheng; Molleman, Lucas; Van Dolder, Dennie (2020) : Conditional punishment: Descriptive social norms drive negative reciprocity, CeDEx Discussion Paper Series, No. 2020-05, The University of Nottingham, Centre for Decision Research and Experimental Economics (CeDEx), Nottingham

This Version is available at:

<https://hdl.handle.net/10419/228369>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Discussion Paper No. 2020-05

Xueheng Li  
Lucas Molleman  
Dennie van Dolder

April 2020

**Conditional punishment:  
Descriptive social norms drive  
negative reciprocity**



CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit <http://www.nottingham.ac.uk/cedex> for more information about the Centre or contact

Suzanne Robey  
Centre for Decision Research and Experimental Economics  
School of Economics  
University of Nottingham  
University Park  
Nottingham  
NG7 2RD  
Tel: +44 (0)115 95 14763  
[suzanne.robey@nottingham.ac.uk](mailto:suzanne.robey@nottingham.ac.uk)

The full list of CeDEX Discussion Papers is available at

<http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx>

# Conditional punishment: Descriptive social norms drive negative reciprocity

By: Xueheng Li<sup>1,2,†</sup>, Lucas Molleman<sup>2,3,4,\*</sup>, † and Dennie van Dolder<sup>2,5,6,†</sup>

<sup>1</sup> Economics Experimental Lab, Nanjing Audit University, China

<sup>2</sup> Centre for Decision Research and Experimental Economics, University of Nottingham, United Kingdom

<sup>3</sup> Amsterdam Brain and Cognition, University of Amsterdam, The Netherlands

<sup>4</sup> Center for Adaptive Rationality, Max Planck Institute for Human Development Berlin, Germany

<sup>5</sup> School of Business and Economics, Vrije Universiteit Amsterdam, The Netherlands

<sup>6</sup> Tinbergen Institute, The Netherlands

\* Corresponding author, Nieuwe Achtergracht 129-B, 1018 WT Amsterdam, The Netherlands. Email: l.s.molleman@uva.nl

† Equal contributions

This version: 6 April 2020

## Keywords

Cooperation; peer punishment; decision-making experiment; sanctioning; online experiment; conditional strategies

## **Abstract**

Is punishment of free riders driven by descriptive social norms of cooperation and punishment? We conduct experiments in which participants interact in a one-shot social dilemma with punishment. We study how punishment of free riders is influenced by behavior among members of a payoff-irrelevant reference group. Participants can condition punishment on either the level of cooperation or the level of punishment in the reference group, respectively reflecting descriptive norms of cooperation and punishment. We observe considerable heterogeneity in punishment behavior. Among punishers, the most common strategies are to increase punishment with higher levels of cooperation ('norm enforcement'), and to increase punishment with higher levels of punishment in the reference group ('conformist punishment'). By means of a simple dynamic model, we demonstrate that these conditional punishment strategies can substantially promote cooperation: conformist punishment helps cooperation to gain a foothold in a population, and norm enforcement helps to maintain cooperation at high levels. Our study illustrates how punishment is shaped by the social context, and highlights the potential of conditional punishment strategies to promote the emergence and maintenance of cooperation.

## Introduction

For organizations, communities, and society as a whole to function, individuals often have to engage in activities that are costly for themselves, but beneficial for others. Peer punishment is considered to be one of the key mechanisms to explain why humans often cooperate in situations where private and collective incentives do not align: many people are willing to punish those who free ride on the cooperation of others, even if punishment is costly and cannot lead to future benefits (1–8). The threat of punishment makes free riding less attractive and can thereby help maintain cooperation at high levels (4, 5, 9–18).

Given that peer punishment can play a pivotal role in sustaining cooperation, it is critical to understand what factors influence people's willingness to punish. When studying the drivers of peer punishment, laboratory studies typically focus on aspects specific to the interaction at hand, such as peers' cooperation decisions, the cost and impact of punishment, or the potential for future interaction or retaliation (e.g., 4, 10, 15, 19–21). In doing so, these studies generally abstract away from the broader social context in which an interaction takes place. Cross-cultural experiments, however, show that social context matters for the effectiveness of punishment to support cooperation: people from different societies use peer punishment in systematically different ways (3, 22–28). Because societies differ from each other in myriad ways, such cross-cultural comparisons have limited ability to identify exactly which aspects of the social context underlie any observed differences.

In this paper, we investigate an important way in which the social context may influence punishment of free riding: through indicating 'descriptive norms' specifying what behavior is typical in the current interaction setting (29–31). Studies from across the social sciences have shown that people tend to conform to descriptive norms (29, 32–37). In social dilemmas, it has been established that many people are more willing to cooperate if they believe that others will do so as well (8, 20, 28, 31, 38–40). Whether, and if so how, descriptive norms influence peer punishment, however, remains unclear. Here, we first provide experimental evidence that many people condition their punishment of a free-riding partner on descriptive norms of cooperation and punishment. With a simple dynamic model, we then show that such conditional punishment strategies can have pronounced implications for the emergence and maintenance of cooperation in groups.

For the decision to punish a free riding peer, two descriptive norms may be important. First, punishment might be guided by the descriptive norm of cooperation: is free riding the typical action in the population? It has been shown that people often infer injunctive norms (what one ought to do) from descriptive norms (what most people actually do): people tend to judge behaviors that are less common in a population to be less socially appropriate (or 'moral') and

consequently more deserving of punishment (41–48). If people use descriptive norms of cooperation to form moral judgments in this manner, they will judge free riding more harshly when it is atypical, which will increase their willingness to punish. Second, punishment might be guided by a descriptive norm of punishment: is punishment a typical reaction to free riding? Descriptive norms of punishment can signal a ‘principle of social proof’ (49) that free riding is disapproved of, and that punishment is an appropriate and legitimate reaction. Conformity to these norms would lead people to punish free riding if others do so as well. Examining the impact of these two descriptive norms on sanctioning behavior increases our understanding of how the social context can affect individuals' punishment of free riding and thereby influence the emergence and maintenance of cooperation.

To investigate whether descriptive norms of cooperation and punishment impact peer punishment, we conduct a large-scale decision-making experiment. Participants are randomly paired and play a prisoner's dilemma with punishment. Our implementation consists of two stages. In the first stage, participants decide to either ‘cooperate’ or ‘defect’. In the second stage, they decide how severely they want to punish their partner if their partner chose to defect. We add minimal social context by allowing participants to condition their punishment decisions on the levels of cooperation and punishment displayed by participants who previously interacted in the same setting (hereafter, the ‘reference group’). In two between-subject treatments, participants can either condition their punishment decisions on (i) the level of cooperation, or (ii) the level of punishment in the reference group. Importantly, the decisions of members of the reference group do not affect payoffs of the focal participants.

Our setup enables us to classify individual participants according to how their punishment decisions respond to descriptive norms, thereby deepening empirical understanding of individual differences in (conditional) punishment. Individual differences in conditional cooperation have received considerable attention in prior research, indicating that the dynamics of cooperation in groups strongly depend on the interplay of individuals' conditional strategies and their beliefs about others' cooperativeness (40, 50, 51). In sharp contrast, little is known about individual differences in conditional punishment and the way in which these differences may affect the emergence of cooperation. Our experimental design allows us to isolate the possible effects of descriptive norms on punishment from related considerations such as a preference for coordinated punishment or positive reciprocity towards other punishers (7, 52–55). Finally, by creating controlled conditions that systematically differ in terms of descriptive norms of cooperation and punishment, our setup complements cross-cultural experiments on punishment that rely on natural variation in social context (3, 22, 24–28, 56).

Our results demonstrate that on aggregate, people's willingness to punish their free riding partner increases both with the level of cooperation and with the level of punishment in the reference group. Importantly, we observe substantial heterogeneity in how people react to the level of cooperation and the level of punishment. Among punishers, we find that three strategies predominate: 'independent punishment', applying the same punishment intensity irrespective of the descriptive norm, 'norm enforcement', increasing punishment intensity with the fraction of cooperators in the reference group, and 'conformist punishment', increasing punishment intensity with punishment levels in the reference group.

To examine the possible long-term implications of the experimentally observed conditional punishment strategies, we develop a simple dynamic model in which a population of agents recurrently interact in a social dilemma game with punishment similar to our experiment. We use analytical methods and agent-based simulations to evaluate how the experimentally observed punishment strategies can shape cooperation in a population.

The model captures key qualitative features of social norm dynamics, involving prolonged periods of stability and sudden shifts. Moreover, the model shows that, in conjunction with independent punishers, norm enforcement and conformist punishment can effectively support cooperation. Importantly, we find that norm enforcement and conformist punishment play markedly different roles in promoting cooperation: conformist punishment can effectively promote the establishment of cooperation in a population, whereas norm enforcement is particularly effective at maintaining cooperation at high levels. Overall, our model shows that the experimentally observed conditional punishment strategies can have a strong and positive impact on the dynamics of cooperation.

## **Experimental design**

We randomly matched participants in pairs to play a two-stage game in which they could earn points (which were converted into dollars at the end of the game). In the first stage, the two players simultaneously choose to cooperate or defect. Joint payoffs are highest when both partners cooperate, with both earning 18 points. However, each individual can increase their personal payoffs in this stage by choosing to defect: unilateral defection leads to 25 points for self and 9 points for the other. Mutual defection leads to 16 points for each. In the second stage, participants have the opportunity to punish their interaction partner if their partner chose to defect (by design excluding 'antisocial punishment'; see 26), by assigning up to 10 deduction points to them. Each assigned deduction point reduces the participant's payoffs with



1 point, and the partner's payoffs with 3 points. The Nash equilibrium of this one-shot game is to defect in the first stage, and to never assign any deduction points in the second stage.

We report on two separate treatments (total  $N=999$ ), in which participants could condition their punishment on descriptive norms of cooperation (CC treatment;  $N=498$ ) or descriptive norms of punishment (CP treatment;  $N=501$ ). We operationalized these descriptive norms as behavior in a reference group of individuals who previously interacted in the same setting, but who were irrelevant for the payoffs in the current interaction. Participants had to indicate how many deduction points they would assign to their partner (if the partner chose to defect) for a set of situations that vary with respect to the reference group's levels of cooperation or punishment. The actual behavior in the reference group determined which of the situations was implemented and used to calculate payoffs (see Methods for details; the SI shows the experimental materials in full).

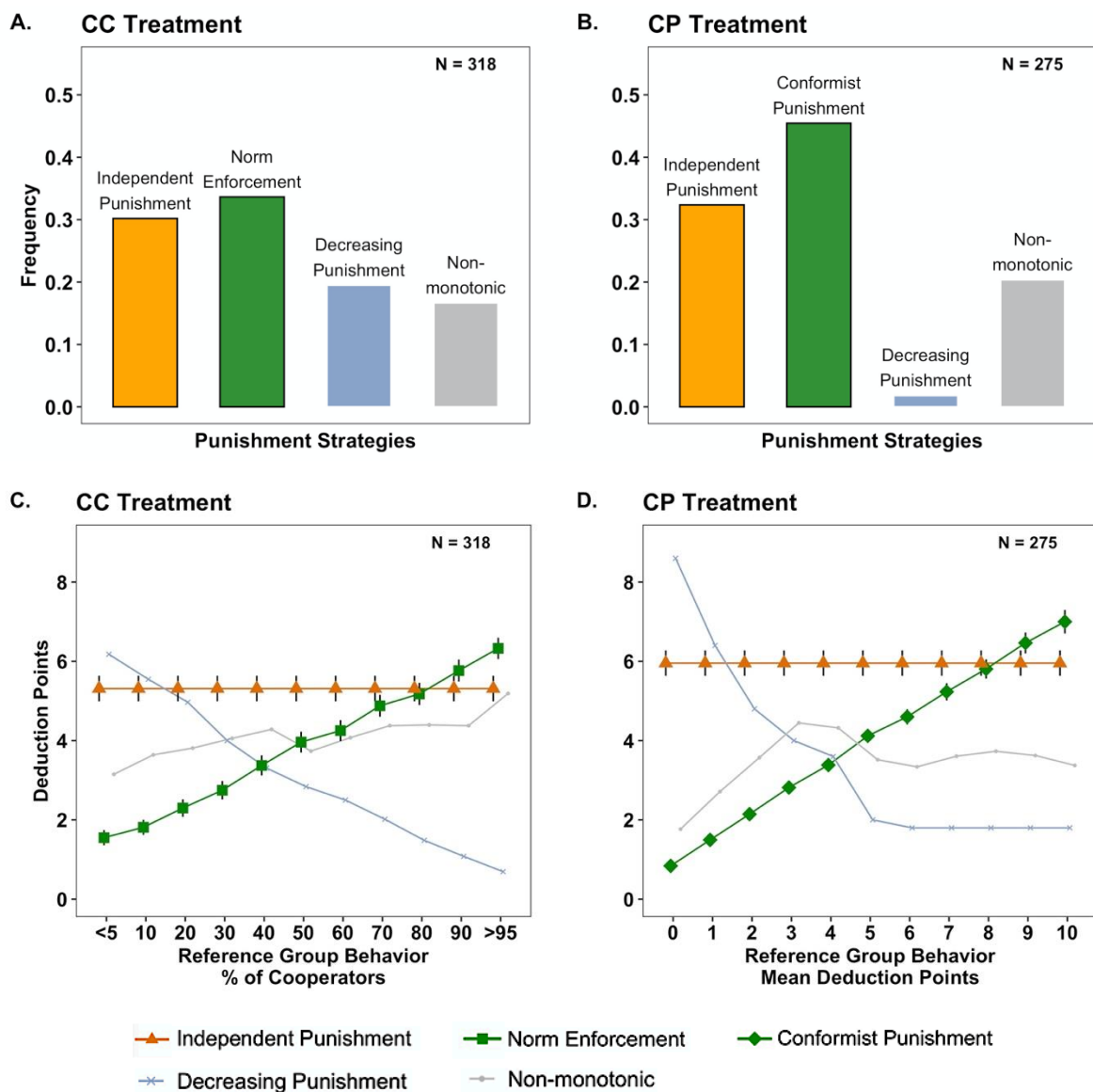
## Results

**Experimental results.** Treatments did not differ in terms of cooperation rates (CC: 68.5%; CP: 65.5%;  $\chi^2(1)=1.02$ ,  $P=0.313$ ). Participants' overall punishment levels, averaged across all situations, were also similar across treatments (CC: 2.68; CP: 2.47 deduction points; two-sample Wilcoxon rank-sum test: d.f.=997,  $z=1.703$ ,  $P=0.089$ ).

On aggregate, behavior in the reference group impacted the participants' punishment decisions: both the fraction of cooperators and the average intensity of punishment had a significantly positive effect on the average number of deduction points that participants assigned to their free riding partners (ordinary least squares regression:  $P<0.01$  for both treatments; Table S1; Fig. S1). We interpret this as evidence that the social context impacts peer punishment, with both descriptive norms of cooperation and descriptive norms of punishment modulating people's overall willingness to punish defectors.

Participants substantially differed in their punishment behavior (Fig. 1). Among participants who punished at least once (64% and 55% for CC and CP, respectively), three distinct punishment strategies predominate (Fig. 1A,B): (i) 'independent punishment', applying the same punishment intensity irrespective of the behavior in the reference group (Fig. 1A,B; orange bars), (ii) 'norm enforcement', monotonically increasing punishment with the level of cooperation in the reference group (Fig. 1A; green bar), and (iii) 'conformist punishment', monotonically increasing punishment with the level of punishment in the reference group (Fig. 1B; green bar). In the CC treatment, a smaller portion of participants decreased their punishment of free riders as cooperation became more common in the reference group (Fig.

1A; blue bar); in the CP treatment, such ‘decreasing punishment’ was virtually absent (Fig. 1B, blue bar). These results indicate that people substantially vary in how they condition punishment of free riders on the levels of cooperation and punishment in the social environment.

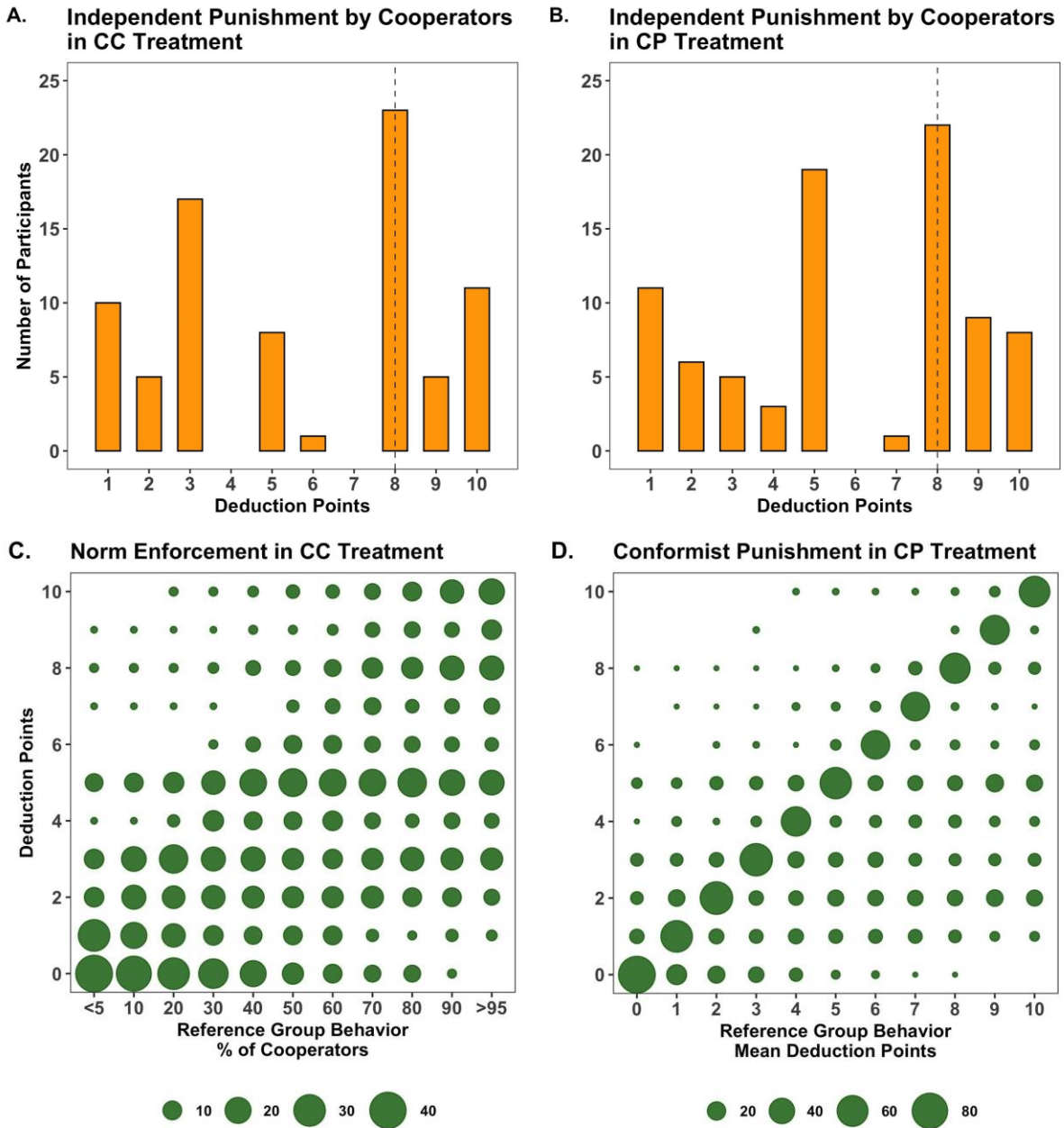


**Fig. 1.** Punishment strategies observed in our experiment. (A and B) Frequency distributions of punishment strategies in the CC and CP treatment, among participants who punish at least once. (C and D) For each strategy, the average number of deduction points (+/- 1 SEM) assigned to free riding partners for the situations in the CC and CP treatment.

Figure 1C and 1D show, for each of the punishment strategies, the average number of deduction points assigned. Participants who engaged in norm enforcement (Fig. 1C; green) and participants who engaged in conformist punishment (Fig. 1D; green) strongly reacted to the level of cooperation and punishment in the reference group. On average, 'norm enforcing' participants assigned 1.6 deduction points when the percentage of cooperators in the reference group was less than 5%. Their punishment increased to 6.3 deduction points when more than 95% of the participants in the (payoff-irrelevant) reference group cooperated (Figure 1C; green line). Similarly, in the CP treatment, participants who punished conformistically assigned about 0.8 deduction points when participants in the reference group assigned 0 deduction points on average. Their punishment increased sharply to 6.5 deduction points when the average number of deduction points assigned by members of the reference group was 10. Taken together, these results show that the punishment behavior of participants who use conditional strategies is strongly affected by the social environment.

For participants who punished independently and cooperated in stage 1, the modal behavior in both treatments was to assign 8 deduction points (Fig. 2A,B). By contrast, assigning 8 deduction points is very rare among independent punishers who defected in stage 1 (see Fig. S3-5 for a full breakdown of punishment decisions by cooperators and defectors in each treatment). This level of punishment equalizes the earnings between a cooperator and their free-riding partner, suggesting that some participants' do not punish to reciprocate the unkind action, but rather to eliminate disadvantageous inequality (Fehr and Schmidt 1999; Raihani & Bshary 2019).

Figure 2C and D show the distributions of assigned deduction points among participants engaging in norm enforcement in the CC treatment (Fig. 2C) and conformist punishment in the CP treatment (Fig 2D). We observe large numbers of data points on the diagonal in the graph for conformist punishment in the CP treatment. This indicates that participants engaging in conformist punishment frequently chose to exactly match the average number of deduction points assigned in the reference group. Norm enforcement in the CC treatment showed a less pronounced pattern.



**Fig. 2.** Punishment behavior for the most common punishment strategies. (A and B) Distributions of deduction points assigned by participants who punished independently and cooperated in the first stage of the game. The mode behavior for both treatments (assigning 8 deduction points; vertical dotted line) equalizes the earnings of a cooperator and a free rider. (C and D) Deduction points assigned by participants who engaged in 'norm enforcement' in the CC treatment, and 'conformist punishment' in the CP treatment. Dot sizes reflect the numbers of observations.

**Dynamic model.** Our experimental results reveal that people’s punishment of free riders is shaped by descriptive norms of cooperation and punishment, and that various punishment strategies (conditional and unconditional) co-exist. This raises the question of how the observed punishment strategies interact to drive dynamics of cooperation over time. To address this question, we study a simple model using analytical methods and agent-based simulations. The model allows us to examine how the dynamics of cooperation depends on the composition of the population regarding the agents’ punishment strategies.

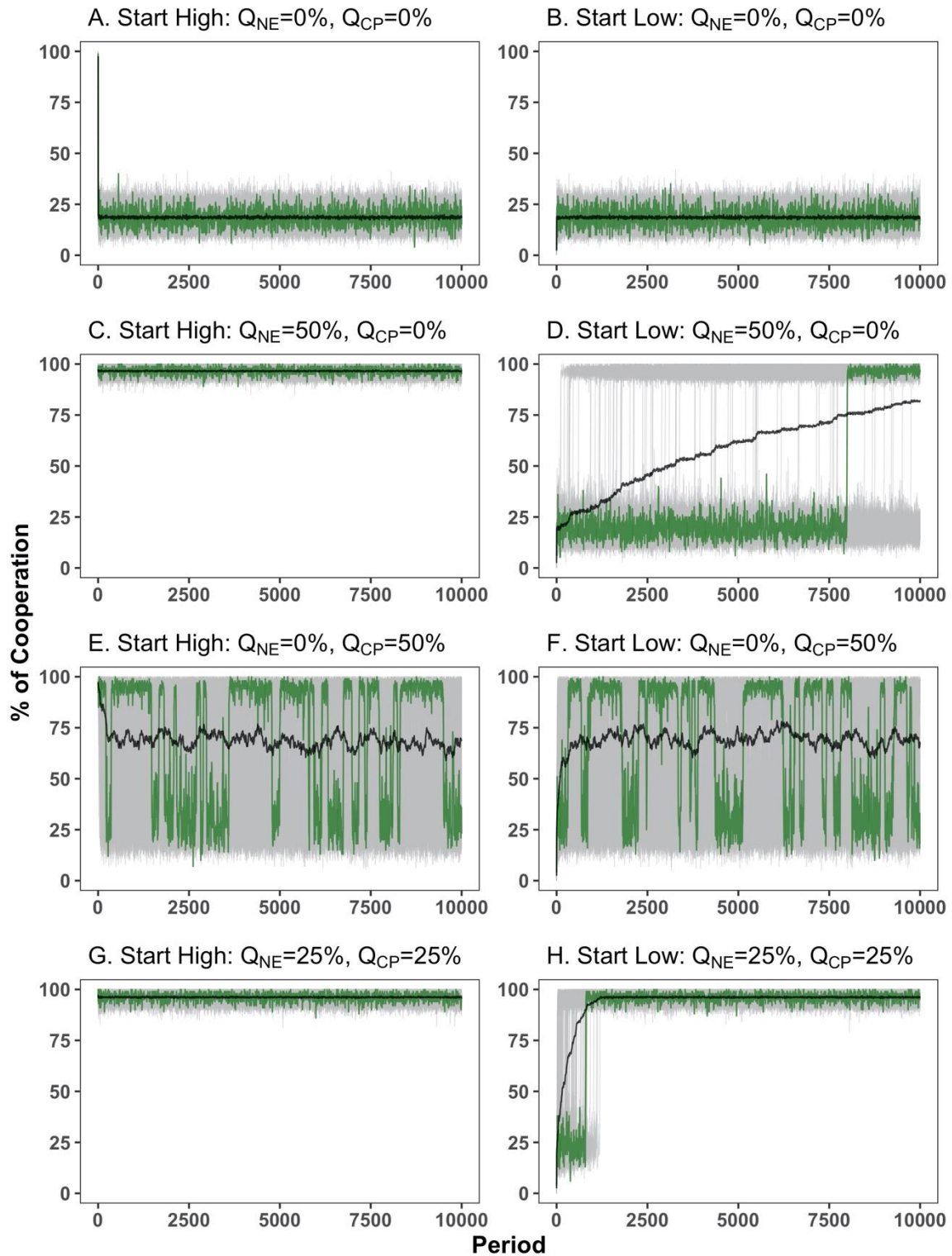
We consider a population of  $n$  agents who interact repeatedly for  $T$  periods in a setting similar to our experiment. In each period, agents ( $i$ ) are randomly matched into pairs, ( $ii$ ) choose whether to cooperate or defect, and ( $iii$ ) choose whether to punish their partner if their partner defects. For ease of exposition and to facilitate tractability, we model both cooperation and punishment as binary decisions. In each period, each agent samples  $m$  agents from the population and counts how many of them cooperated and how many of them were willing to punish defectors in the previous period. The counts divided by  $m$  become their beliefs about the rates of cooperation and punishment in the current period (respectively denoted by  $b_c$  and  $b_p$ ). An agent cooperates in the current period if they believe that the proportion of punishers exceeds a threshold ( $b_p > \theta_c$ ), and defects otherwise.

An agent’s punishment strategy determines whether they punish a defecting partner. Based on our experimental results, we consider four punishment strategies: ( $i$ ) *independent punishment*: punish irrespective of beliefs; ( $ii$ ) *norm enforcement*: punish if the perceived cooperation rate exceeds a threshold ( $b_c > \theta_{NE}$ ); ( $iii$ ) *conformist punishment*: punish if the perceived punishment rate exceeds a threshold ( $b_p > \theta_{CP}$ ); and ( $iv$ ) *never punish*. For the sake of exposition, we will here focus on the case where  $\theta_c = \theta_{NE} = \theta_{CP} = 0.5$ . In the Supplementary Analysis, we present results for arbitrary threshold values. We further assume that agents’ punishment strategies are fixed throughout all periods and mutually exclusive. The dynamics are stochastic: with probability  $\varepsilon > 0$ , an agent makes a mistake and behaves randomly; with complementary probability  $1 - \varepsilon$  the agent behaves according to its strategy (Young 1993, Kandori et al. 1993; see Methods for more details).

Our goal is to assess how relative frequencies of independent punishment ( $Q_{IP}$ ), norm enforcement ( $Q_{NE}$ ), and conformist punishment ( $Q_{CP}$ ) affect the dynamics of cooperation. First, we derive analytical results about the stationary distribution of the dynamic when the observation sample is large ( $m=n$ ) and the mistake probability is vanishingly small ( $\varepsilon \rightarrow 0$ ). The stationary distribution reflects the relative frequencies of different population states in the long run ( $T \rightarrow \infty$ ). We show that if  $Q_{IP} + \frac{1}{2}(Q_{NE} + Q_{CP}) > \frac{1}{2}$ , only the cooperation equilibrium

occurs with a positive frequency in the stationary distribution; conversely, if  $Q_{IP} + \frac{1}{2}(Q_{NE} + Q_{CP}) < \frac{1}{2}$ , only the defection equilibrium occurs with a positive probability in the stationary distribution (see SI, Supplementary Analysis for formal proof). This analysis shows that, perhaps unsurprisingly, independent punishment is the most potent strategy for promoting cooperation. Importantly, however, when independent punishment is not sufficiently frequent, conditional strategies of norm enforcement and conditional punishment can be key for sustaining cooperation in the long run.

Next, we use simulations to examine the short-run dynamics of our model: how conditional punishment strategies drive the emergence and breakdown of cooperation, and how their relative frequencies affect the time it takes for a population between states of high and low cooperation. Simulations also allow us to consider small observation samples and non-negligible mistake probabilities. To account for path-dependence, the simulations consider different starting conditions by varying agents' initial beliefs about the rates of cooperation and punishment. To evaluate how norm enforcement and conformist punishment affect cooperation, we fix the frequency of independent punishers at thirty percent and vary the frequencies of the conditional punishment strategies. Further robustness checks are detailed at the end of this section.



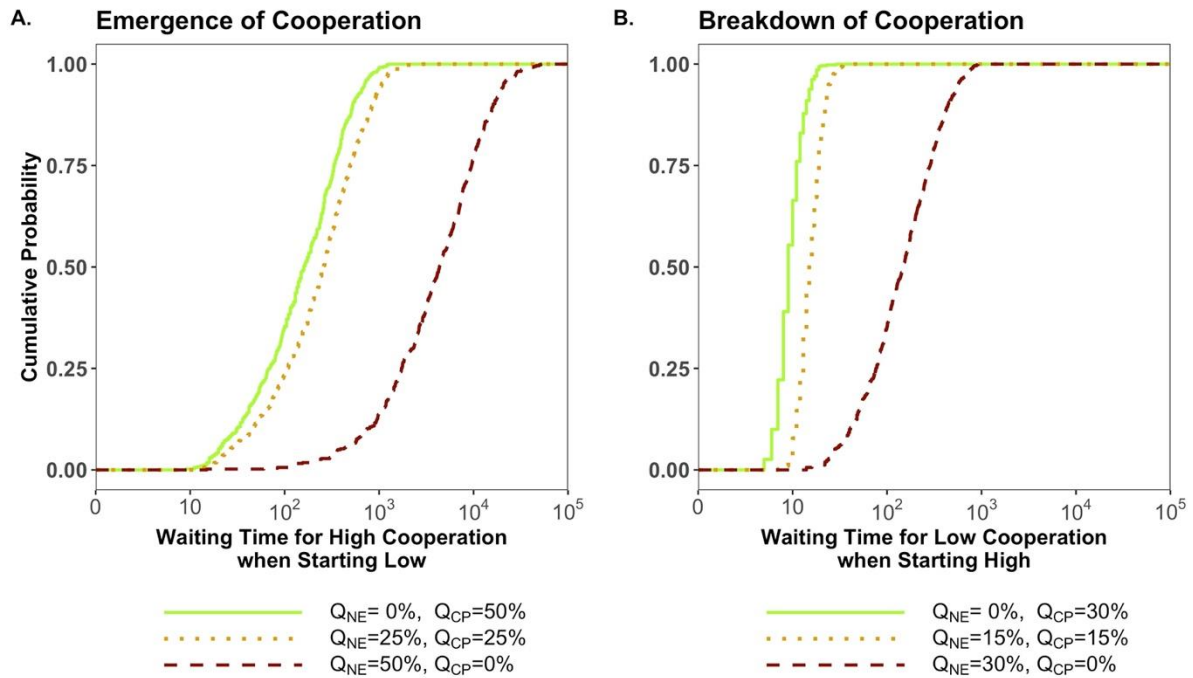
**Fig. 3.** Effects of conditional punishment strategies on cooperation dynamics. Across all panels, we hold fixed the frequency of independent punishers at 30%.  $Q_{NE}$  is the frequency of norm enforcement, and  $Q_{CP}$  is the frequency of conformist punishment. Columns of panels vary agents' initial beliefs regarding the frequencies of punishment and cooperation in the population, either starting high ( $b_c = b_p = 0.75$ ; left column) or starting low ( $b_c = b_p = 0.25$ ; right column; see Methods for details). In each panel, black lines show mean cooperation rates over time across 100 simulation runs; grey lines show individual runs, with a representative run highlighted in green. Further simulation settings:  $n = 100$ ,  $m = 10$ ,  $\varepsilon = 0.05$ .

Figure 3 shows the dynamics of cooperation in situations where independent punishment is not sufficiently frequent to sustain cooperation by itself. We first confirm that, if independent punishers alone are too rare to support cooperation on their own, and neither of the conditional punishment strategies is present in the population, cooperation never emerges in our simulations (Fig. 3A,B). Next, we consider cases where independent punishment is complemented with conditional punishment strategies, raising the overall frequency of punishers. The presence of norm enforcement has a strong stabilizing effect once high levels of cooperation have been achieved (Fig. 3C). However, it might take considerable time for cooperation to emerge (Fig. 3D). These dynamics are driven by a positive feedback loop between norm enforcement and cooperation, locking a population into a state of either high or low cooperation, making it hard to transition from one state to the other.

By contrast, in the presence of conformist punishers cooperation readily emerges, but is not stable (Fig. 3E,F). The population alternates between states with low and high levels of cooperation, with rapid shifts between these states. These dynamics are driven by another positive feedback loop: when levels of cooperation and punishment are low, some agents may punish their free riding partner due to mistakes or—in the case of conformist punishers—due to sampling bias. In turn, these stochastic events may prompt other conformist punishers to punish too in the next period, thereby increasing the levels of cooperation and punishment even more, and possibly tipping the population to high levels of cooperation and punishment. However, similar stochastic processes may also cause cooperation to suddenly break down when conformist punishers stop punishing when they happen to underestimate the level of punishment in the population.

When both conformist punishment and norm enforcement are present in the population—but keeping the overall frequency of conditional punishment the same—cooperation rapidly emerges and remains stable at high levels (Fig. 3G,H). Conformist punishers still amplify the impact of stochasticity when cooperation is low, facilitating the emergence of cooperation. Subsequently, norm enforcement locks the population into a state of high cooperation. This result highlights that the concerted action of conformist punishment and norm enforcement can efficiently support cooperation.





**Fig. 4.** Effects of conditional punishment strategies on the emergence and breakdown of cooperation. Lines show the cumulative probability of cooperation to rise above 75 percent (A) or fall below 25 percent (B), as a function of time. Time is shown on a logarithmic scale, and each line represents 500 simulation runs. Across both panels, we hold fixed the frequency of independent punishers at 30%.  $Q_{NE}$  is the frequency of norm enforcement, and  $Q_{CP}$  is the frequency of conformist punishment. Frequencies of these conditional strategies were chosen such that—according to our analytical results—cooperation would emerge (Panel A) or break down (panel B) in the long run. Initial beliefs regarding cooperation and punishment levels start low in Panel A ( $b_c = b_p = 0.25$ ), and high in Panel B ( $b_c = b_p = 0.75$ ; see Methods for details). Each simulation runs for 100,000 ( $10^5$ ) periods. Further simulation settings:  $n = 100$ ,  $m = 10$ ,  $\varepsilon = 0.05$ . Results for additional population compositions with regard to punishment strategies confirm the general pattern shown here (Fig. S8).

These results indicate that different conditional punishment strategies can promote cooperation in different ways: conformist punishment facilitates the emergence of cooperation; norm enforcement helps to maintain it after its emergence. Figure 4 confirms these insights. When a population starts from a state of low cooperation, the presence of conformist punishment, rather than norm enforcement, can strongly increase the rate at which it shifts to a state of high cooperation (Fig. 4A). Conversely, the presence of norm enforcement can substantially extend the time that a population remains in a state of high cooperation (Fig. 4B).

In the Supplementary Information we examine the generalizability and robustness of our model results. We confirm that our main model results hold across different ranges of relative frequencies of the various (conditional) punishment strategies and different initial beliefs about cooperation and punishment in the population (Fig. S9-10). Furthermore, we show that the presence of agents who decrease their punishment of free riding as cooperation becomes more common—as observed in the CC treatment (‘decreasing punishment’ in Fig. 1A)—

destabilizes the non-cooperative equilibrium. By itself, decreasing punishment cannot support high levels of cooperation. However, in conjunction with other conditional punishment strategies, norm enforcement in particular, decreasing punishment can boost the likelihood that a population reaches high and stable levels of cooperation (Fig. S11).

## Discussion

Our experiment provides large-scale behavioral evidence that punishment of free riding in social dilemmas is shaped both by descriptive norms of cooperation (“is free riding a typical action in the population?”) and by descriptive norms of punishment (“what is the typical punishment reaction to free riding?”). On aggregate, punishment increases both with the level of cooperation and the level of punishment in a payoff-irrelevant reference group. At the individual level, we observe substantial heterogeneity in how people react to these descriptive norms. Whereas a sizable fraction of participants punishes independently of what others are doing ('independent punishment'), at least as many participants display conditional punishment strategies, increasing their punishment either with higher levels of cooperation ('norm enforcement) or with higher level of punishment ('conformist punishment') in the reference group. Overall, our experimental results support the emerging view that conditional strategies are not limited to the domain of positive reciprocity (i.e., cooperation; 40, 49, 50, 56), but are also important in the domain of negative reciprocity (i.e., punishment; 47, 53, 54, 57, 58).

Our finding that people punish free riding more when cooperation is more common provides novel behavioral evidence for the idea that people infer injunctive norms (what is 'moral') from descriptive norms (what is 'common'; 41–46, 48, 60). In doing so, we complement existing research that largely relied on (non-incentivized) moral judgments (43, 45, 46, 48; see 47 for a rare exception). Our behavioral approach, however, does not allow us to pin down the psychological mechanisms underlying the different punishment strategies. Previous evidence suggests that norm enforcement may be driven by increased disapproval of free riding when cooperation is common (47). Similarly, conformist punishers' observing others punishing defectors may increase their own disapproval of defection. Alternatively, it could be also that conformist punishers follow a simple heuristic of copying what others are doing (47, 61, 62). The finding that conformist punishers frequently chose to exactly match the average punishment of others (Fig. 2D) suggests that the latter may be more likely. Future work should combine behavioral data with survey data to investigate to what extent (conditional)

punishment reflects changes in people's moral judgments after observing others' actions, and to what extent it reflects people's conformist inclinations.

An important open question for understanding conditional punishment strategies is whether people who condition their punishment behavior on that of others do so consistently across different decision settings. Although it seems plausible that some individuals are generally more responsive than others to their social environment, it remains an open question whether individuals who engage in norm enforcement when informed of cooperation rates in their environment, would also tend to punish conformistically when informed of punishment rates. Similarly, conditional punishment strategies might correlate with well-studied strategies of conditional cooperation. Experiments addressing these associations would provide deeper insights into the behavioral architecture of cooperation and punishment, contributing to ongoing debates around the generality of strategies across settings involving positive and negative reciprocity (51, 55, 63).

Our model demonstrates how the experimentally identified conditional punishment strategies can have important implications for cooperation dynamics. Analytical results reveal that conformist punishment strategies can considerably broaden the set of conditions under which cooperation can emerge and persist in the long run. Agent-based simulations yield deeper insights into the different roles that norm enforcement and conformist punishment play in this dynamic. Norm enforcers punish free riders when the cooperation rate in the population is relatively high, which makes them effective in maintaining cooperation. However, they do not punish when free riding predominates and are therefore of little help for cooperation to emerge from scratch. In contrast, conformist punishers sanction free riders as long as sufficiently many others do—irrespective of the cooperation rate—and can, therefore, play a valuable role in helping cooperation gain a foothold in a population.

Whereas our experiment shows that the behavior of an individual can be influenced by what the collective is doing, our model illustrates how these individual strategies can subsequently impact collective dynamics. We deliberately employ a simple stylized model to illustrate the basic effects of conditional punishment strategies on the dynamics of cooperation. Despite its simplifying assumptions (e.g., mutually exclusive punishment strategies, binary punishment and cooperation choices, random re-matching after every interaction), our model produces intuitive and robust results. Moreover, the model is able to capture key qualitative features of the dynamics of social norms: prolonged periods of stability which are punctuated by tipping points, where one norm is rapidly replaced by another (Fig. 3; 63). In line with the results obtained by (47), we find that especially the positive social feedback provided by norm enforcers is critical to capture these patterns in norm dynamics.

Our simulations illustrate how conformist punishment can amplify stochastic events, leading to both rapid alternation between the emergence and breakdown of cooperation in a population (Fig. 3, 4). In contrast, norm enforcement can engender a process of positive feedback with cooperation, locking a population into a state of either high or low levels of cooperation, making it hard to transition to the other state (Fig. 3). These results give pointers for efficiently promoting desirable behaviors, such as voting, tax compliance, or energy conservation. In particular, facilitating the observability of (or accessibility to) information about other people's behavior may be effective when the majority of the population displays the desired behavior: this information can boost norm enforcement, ensuring that adherence to the present norm remains high. Conversely, when a majority of the population shows the undesired behavior, it may be more effective to provide people with information that informs them that many people disapprove of the undesirable behavior. Such information may trigger conformist punishment and shift the system to the more desirable outcome.

## Methods

**Experimental Procedures.** We recruited participants from Amazon Mechanical Turk (MTurk; average age 35.5 (s.d.=10.3), range 18-71; 43% male) during September 2017 and September 2019. Our main experimental results do not differ between these two waves of data collection (Fisher's exact test does not reject the hypothesis that the distributions of punishment strategies (Fig. 1A,B) are the same; CC treatment:  $X^2(4)=3.472$ ,  $P=0.482$ ; CP treatment:  $\chi^2(4) = 2.320$ ,  $P=0.677$ ). Hence we pool the data in our analysis. We restricted our sample to the United States for reasons of comprehension of English instructions. The only other participation criterion was to have at least 95% of previous HITs approved (HITs are jobs performed on MTurk; see Supplementary Information, Experimental Procedures). The experiment was programmed in LIONESS Lab (65), code is available in the online GitHub repository associated with this paper; experimental instructions are documented in full in the Supplementary Information. Ethical approval was given by the Research Ethics Committee at the School of Economics, University of Nottingham, UK.

After reading the instructions and passing compulsory control questions, participants entered stage 1 and made their binary cooperation decisions. In stage 2, participants completed another set of compulsory control questions (see Figs. S6 and S7 for details), before we asked them to provide their punishment responses to descriptive norms of cooperation and punishment. We used the strategy method (66) to obtain a full punishment profile for each individual (55, 67–69). In the CC treatment, we operationalized the descriptive norm of

cooperation as the fraction of cooperative choices in a payoff-irrelevant reference group (sampled from a pre-recorded pool; details below). We presented participants with eleven situations regarding the proportion of cooperators in this reference group, spanning the full range of possible outcomes. For each of these situations, participants had to indicate how many deduction points they would assign to their current interaction partner. In the CP treatment, we operationalized the descriptive norm of punishment as the average intensity of punishment in the reference group, and participants indicated for each possible situation how many deduction points they would assign to their current interaction partner.

The pre-recorded pool consisted of a total of 273 MTurkers who played a prisoner's dilemma with punishment mirroring our experiment (cooperation rate: 69%; average punishment of free riding partners: 2.7 deduction points). For each dyad in the main experiment, we independently sampled 50 participants from the pre-recorded pool to form the reference group. The behavior of the reference group defined the situation that was used to calculate participants' earnings. Since participants did not know which situation was the actual one beforehand, they were incentivized to consider each situation as if it was real.

Once participants had completed the two decision making stages of the experiment, they were placed in a lobby, in which they would be matched with another participant as soon as they completed their decisions as well. Excluding the time spent in the lobby, our experiment on average lasted 9.9 minutes. In our experiment, participants could earn points which were converted to US dollars at the end of the experiment (20 points were worth \$1.00). Average earnings were \$1.96 (range \$0.41 - \$2.51), which translates to an hourly wage of \$12.00.

We define independent punishment as using the same (non-zero) level of punishment across all situations. We defined conditional punishment strategies of norm enforcement and conformist punishment as showing a weakly monotonic increase in punishment in responses to increasing levels of cooperation (CC treatment) and punishment (CP treatment) in the reference group. This approach based on monotonicity is a conservative way to identify conditional punishment strategies: an alternative classification method based on linear regression models would lead all individuals with non-monotonic response patterns (cf. Fig. 1) to be identified as using either independent, increasing, and decreasing punishment strategies.

**Dynamic model.** In the first period of the simulations, agents are endowed with initial beliefs about the norms of cooperation and punishment, and respond to the beliefs according to their specified strategies. For Starting High in Fig. 3 and Fig. 4B, agents initially believe that 75% of the agents in the population will cooperate ( $b_c = 0.75$ ) and 75% of the agents would punish free riding ( $b_p = 0.75$ ). For all agents, the payoff maximizing response to these beliefs is to

cooperate; independent punishers, norm enforcers, and conformist punishers punish their defecting partner when holding these beliefs. For Starting Low in Fig. 3 and Fig. 4A, agents have initial beliefs  $b_c = b_p = 0.25$ . For all agents, the payoff maximizing response to these beliefs is to defect; only independent punishers punish defectors when holding these beliefs. In each subsequent period, each agent updates their beliefs by sampling  $m$  agents from the population with probability  $u$ . We set  $u = 0.5$  in our simulations; our analytical results apply to any  $u$  with  $0 < u < 1$ . Assuming  $u < 1$  prevents that all agents simultaneously update in a period with probability one (69–72; see Remarks in Supplementary Analysis, Long-run Equilibrium). Agents make mistakes with probability  $\varepsilon$ . Mistakes are independent across agents, periods, and cooperation and punishment decisions. Full simulation code is available from the public repository associated with this paper ([https://github.com/LucasMolleman/LMD\\_Conditional\\_punishment](https://github.com/LucasMolleman/LMD_Conditional_punishment)).

## Acknowledgments

We thank Simon Gächter, Björn Lindström, Jesse Niebaum, Wouter van den Bos and Manwei Liu for useful comments and discussions. We thank Benjamin Beranek for help with data collection. L.M. acknowledges support from the European Research Council (Grant: ERC-AdG 295707 COOPERATION), an Amsterdam Brain and Cognition Project grant 2018, and the Jacobs Foundation. D.v.D. acknowledges support from the Economic and Social Research Council through the Network for Integrated Behavioural Sciences (Grant: ES/K002201/1), and the Netherlands Organisation for Scientific Research (Grant: 452-17-004).

## References

1. E. Ostrom, J. Walker, R. Gardner, Covenants with and without a sword: self-governance is possible. *Am. Polit. Sci. Rev.* **86**, 404–417 (1992).
2. K. Sigmund, Punish or perish? Retaliation and collaboration among humans. *Trends Ecol. Evol.* **22**, 593–600 (2007).
3. S. Gächter, B. Herrmann, Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 791–806 (2009).
4. E. Fehr, S. Gächter, Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* **90**, 980–994 (2000).
5. E. Fehr, U. Fischbacher, S. Gächter, Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nat.* **13**, 1–25 (2002).

6. E. Fehr, S. Gächter, Altruistic punishment in humans. *Nature* **415**, 137–140 (2002).
7. F. Guala, Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *Behav. Brain Sci.* **35**, 1–15 (2012).
8. E. Fehr, I. Schurtenberger, Normative foundations of human cooperation. *Nat. Hum. Behav.* **2**, 458–468 (2018).
9. B. Rockenbach, M. Milinski, The efficient interaction of indirect reciprocity and costly punishment. *Nature* **444**, 718–723 (2006).
10. M. Egas, A. Riedl, The economics of altruistic punishment and the maintenance of cooperation. *Proc. R. Soc. B Biol. Sci.* **275**, 871–878 (2008).
11. S. Gächter, E. Renner, M. Sefton, The long-run benefits of punishment. *Science* **322**, 1510 (2008).
12. N. Nikiforakis, H.-T. Normann, A comparative statics analysis of punishment in public-good experiments. *Exp. Econ.* **11**, 358–369 (2008).
13. D. G. Rand, A. Dreber, T. Ellingsen, D. Fudenberg, M. A. Nowak, Positive interactions promote public cooperation. *Science* **325**, 1272–5 (2009).
14. M. J. Crockett, L. Clark, M. D. Lieberman, G. Tabibnia, T. W. Robbins, Impulsive choice and altruistic punishment are correlated and increase in tandem with serotonin depletion. *Emotion* **10**, 855–862 (2010).
15. N. Nikiforakis, Feedback, punishment and cooperation in public good experiments. *Games Econ. Behav.* **68**, 689–702 (2010).
16. R. P. Cubitt, M. Drouvelis, S. Gächter, Framing and free riding: emotional responses and punishment in social dilemma games. *Exp. Econ.* **14**, 254–272 (2011).
17. N. J. Raihani, A. Thornton, R. Bshary, Punishment and cooperation in nature. *Trends Ecol. Evol.* **27**, 288–295 (2012).
18. A. A. Arechar, S. Gaechter, L. Molleman, Conducting interactive experiments online. *Exp. Econ.* **21**, 99–131 (2018).
19. G. Camera, M. Casari, Cooperation among strangers under the shadow of the future. *Am. Econ. Rev.* **99**, 979–1005 (2009).
20. S. Gächter, F. Kölle, S. Quercia, Reciprocity and the tragedies of maintaining and providing the commons. *Nat. Hum. Behav.* **1**, 650–656 (2017).
21. N. J. Raihani, R. Bshary, Punishment: one tool, many uses. *Evol. Hum. Sci.* **1**, e12 (2019).
22. A. E. Roth, V. Prasnikar, M. Okuno-Fujiwara, S. Zamir, Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study. *Am. Econ. Rev.*, 1068–1095 (1991).
23. J. Henrich, Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *Am. Econ. Rev.* **90**, 973–979 (2000).

24. H. Oosterbeek, R. Sloof, G. van de Kuilen, Cultural differences in ultimatum game experiments: evidence from a meta-analysis. *Exp. Econ.* **7**, 171–188 (2004).
25. J. Henrich, *et al.*, Costly punishment across human societies. *Science* **312**, 1767–1770 (2006).
26. B. Herrmann, C. Thöni, S. Gächter, Antisocial punishment across societies. *Science* **319**, 1362–1367 (2008).
27. J. Henrich, *et al.*, Markets, religion, community size, and the evolution of fairness and punishment. *Science* **327**, 1480–1484 (2010).
28. J. P. Henrich, *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter* (Princeton University Press, 2016).
29. R. B. Cialdini, C. A. Kallgren, R. R. Reno, A focus theory of normative conduct: a theoretical refinement and reevaluation of the role of norms in human behavior. *Adv. Exp. Soc. Psychol.* **24**, 201–234 (1991).
30. A. Berkowitz, “An overview of the social norms approach” in *Challenging the Culture of College Drinking: A Socially Situated Health Communication Campaign*, 2005th Ed., (Hampton Press, 2005).
31. C. Bicchieri, *The grammar of society: the nature and origins of social norms* (Cambridge University Press, 2006).
32. R. Bond, P. B. Smith, Culture and conformity: a meta-analysis of studies using Asch’s (1952b, 1956) line judgment task. *Psychol. Bull.* **119**, 111–137 (1996).
33. R. B. Cialdini, R. R. Reno, C. A. Kallgren, A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.* **58**, 1015–1026 (1990).
34. B. S. Frey, S. Meier, Social comparisons and pro-social behavior: testing “conditional cooperation” in a field experiment. *Am. Econ. Rev.* **94**, 1717–1722 (2004).
35. J. M. Nolan, P. W. Schultz, R. B. Cialdini, N. J. Goldstein, V. Griskevicius, Normative social influence is underdetected. *Pers. Soc. Psychol. Bull.* **34**, 913–923 (2008).
36. J. M. Burger, *et al.*, Nutritious or delicious? The effect of descriptive norm information on food choice. *J. Soc. Clin. Psychol.* **29**, 228–242 (2010).
37. M. Hallsworth, J. A. List, R. D. Metcalfe, I. Vlaev, The behavioralist as tax collector: using natural field experiments to enhance tax compliance. *J. Public Econ.* **148**, 14–31 (2017).
38. J. Elster, Social norms and economic theory. *J. Econ. Perspect.* **3**, 99–117 (1989).
39. J. Elster, *The cement of society: a survey of social order* (Cambridge University Press, 1989).
40. U. Fischbacher, S. Gächter, E. Fehr, Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404 (2001).
41. D. Hume, *A treatise of human nature* (Courier Corporation, 2003).



42. H. H. Kelley, Moral evaluation. *Am. Psychol.* **26**, 293–300 (1971).
43. K. M. McGraw, Subjective probabilities and moral judgments. *J. Exp. Soc. Psychol.* **21**, 501–518 (1985).
44. M. Chudek, J. Henrich, Culture-gene coevolution, norm-psychology, and the emergence of human prosociality. *Trends Cogn. Sci.* **15**, 218–226 (2010).
45. K. Eriksson, D. Cownden, M. Ehn, P. Strimling, “Altruistic” and “antisocial” punishers are one and the same. *Rev. Behav. Econ.* **1**, 209–221 (2014).
46. C. M. Tworek, A. Cimpian, Why do people tend to infer “ought” from “is”? The role of biases in explanation: *Psychol. Sci.* (2016) (April 4, 2020).
47. B. Lindström, S. Jangard, I. Selbing, A. Olsson, The role of a “common is moral” heuristic in the stability and change of moral norms. *J. Exp. Psychol. Gen.* **147**, 228–242 (2018).
48. M. R. Welch, *et al.*, “But everybody does it...”: the effects of perceptions, moral pressures, and informal sanctions on tax cheating. *Sociol. Spectr.* **25**, 21–52 (2005).
49. R. B. Cialdini, M. R. Trost, “Social influence: social norms, conformity and compliance” in *The Handbook of Social Psychology, Vols. 1 and 2 (4th Ed.)*, D. T. Gilbert, S. T. Fiske, G. Lindzey, Eds. (McGraw-Hill, 1998), pp. 151–192.
50. U. Fischbacher, S. Gächter, Social preferences, beliefs, and the dynamics of free riding in public good experiments. *Am. Econ. Rev.* **100**:1, 541–556 (2010).
51. T. O. Weber, O. Weisel, S. Gächter, Dispositional free riders do not free ride on punishment. *Nat. Commun.* **9**, 2390 (2018).
52. M. Casari, L. Luini, Cooperation under alternative punishment institutions: an experiment. *J. Econ. Behav. Organ.* **71**, 273–282 (2009).
53. M. Casari, L. Luini, Peer punishment in teams: expressive or instrumental choice? *Exp. Econ.* **15**, 241–259 (2012).
54. K. Kamei, Conditional punishment. *Econ. Lett.* **124**, 199–202 (2014).
55. L. Molleman, F. Kölle, C. Starmer, S. Gächter, People prefer coordinated punishment in cooperative interactions. *Nat. Hum. Behav.* (2019).
56. J. Henrich, *et al.*, In search of homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* **91**, 73–78 (2001).
57. C. Keser, F. Van Winden, Conditional cooperation and voluntary contributions to public goods. *Scand. J. Econ.* **102**, 23–39 (2000).
58. O. FeldmanHall, A. R. Otto, E. A. Phelps, Learning moral values: another’s desire to punish enhances one’s own punitive behavior. *J. Exp. Psychol. Gen.* **147**, 1211 (2018).
59. J.-Y. Son, A. Bhandari, O. FeldmanHall, Crowdsourcing punishment: Individuals reference group preferences to inform their own punitive decisions. *Sci. Rep.* **9**, 1–15 (2019).

60. R. Sugden, "The motivating power of expectations" in *Rationality, Rules, and Structure*, Theory and Decision Library., J. Nida-Rümelin, W. Spohn, Eds. (Springer Netherlands, 2000), pp. 103–129.
61. G. Gigerenzer, Why heuristics work. *Perspect. Psychol. Sci.* **3**, 20–29 (2008).
62. G. Gigerenzer, Moral satisficing: rethinking moral behavior as bounded rationality. *Top. Cogn. Sci.* **2**, 528–554 (2010).
63. A. Peysakhovich, M. A. Nowak, D. G. Rand, Humans display a "cooperative phenotype" that is domain general and temporally stable. *Nat. Commun.* **5** (2014).
64. H. P. Young, The evolution of social norms. *Annu. Rev. Econ.* **7**, 359–387 (2015).
65. M. Giamattei, L. Molleman, K. Seyed Yahosseini, S. Gachter, Lioness Lab -- a free web-based platform for conducting interactive experiments online. *SSRN Electron. J.* (2019).
66. R. Selten, "Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments" in *Beiträge Zur Experimentellen Wirtschaftsforschung*, H. Sauermann, Ed. (Mohr, 1967), pp. 136–168.
67. A. Falk, E. Fehr, U. Fischbacher, Driving forces behind informal sanctions. *Econometrica* **73**, 2017–2030 (2005).
68. J. Brandts, G. Charness, The strategy versus the direct-response method: A first survey of experimental comparisons. *Exp. Econ.* **14**, 375–398 (2011).
69. S. L. Cheung, New insights into conditional cooperation and punishment from a strategy method experiment. *Exp. Econ.* **17**, 129–153 (2014).
70. H. P. Young, The evolution of conventions. *Econometrica* **61**, 57–84 (1993).
71. H. P. Young, *Individual strategy and social structure: an evolutionary theory of institutions* (Princeton University Press, 2001).
72. I. Eshel, L. Samuelson, A. Shaked, Altruists, egoists, and hooligans in a local interaction model. *Am. Econ. Rev.* **88**, 157–179 (1998).
73. W. H. Sandholm, *Population games and evolutionary dynamics* (MIT Press, 2010).

# Supplementary Materials

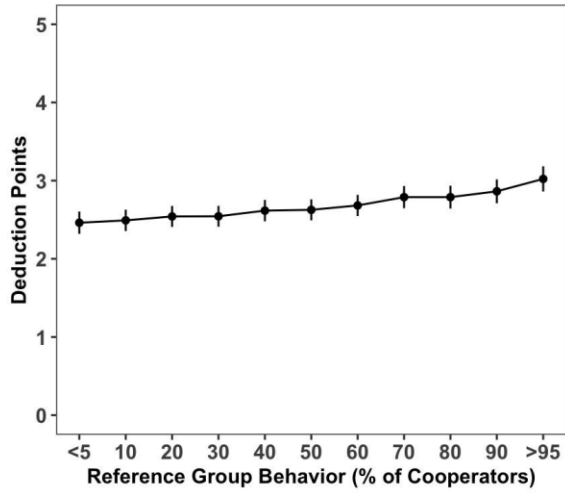
For Li, Molleman, Van Dolder: 'Conditional punishment: Descriptive social norms drive negative reciprocity'

## Contents

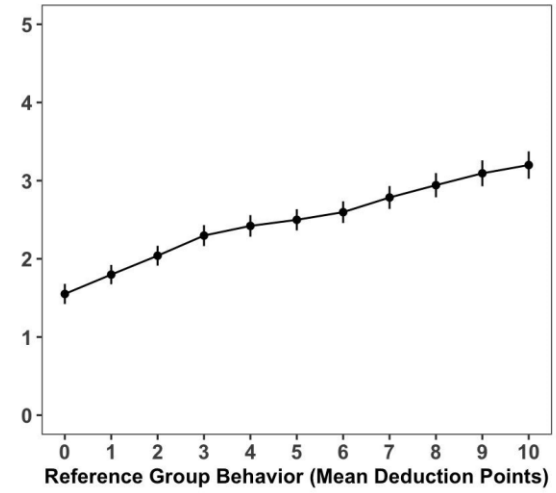
1. Supplementary Figures
2. Supplementary Tables
3. Supplementary Analysis
4. Experimental Procedures and Materials

## Supplementary Figures

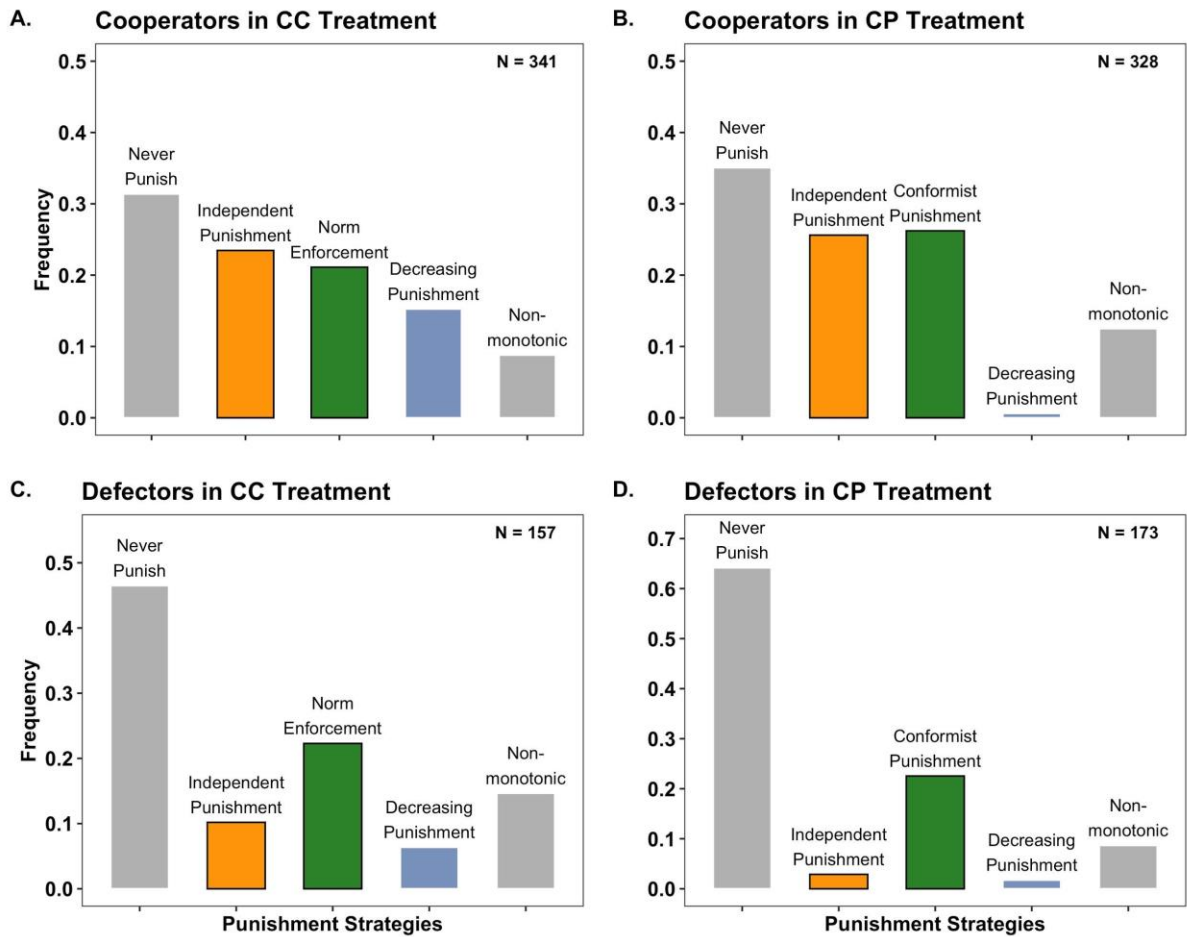
A. CC Treatment (N = 498)



B. CP Treatment (N = 501)

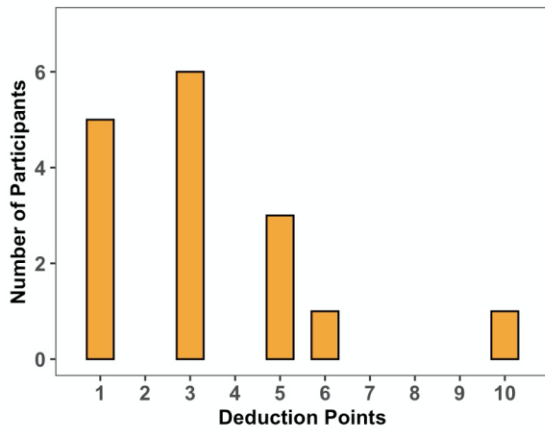


**Fig. S1.** Average intensity of punishment as a function of cooperation and punishment among participants in the payoff-irrelevant reference group. Panel A (B) summarizes decisions in the CC (CP) treatment, showing the average deduction points for each of the situations presented to participants (see screenshots of punishment stage in the experiment). Error bars indicate standard errors of the means (SEM). For statistical analysis, see Table S1.

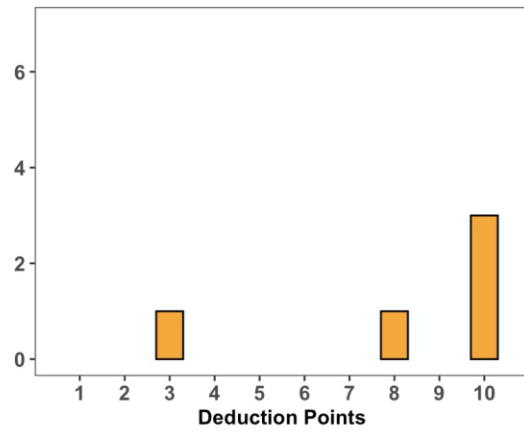


**Fig. S2.** Distributions of punishment strategies in each treatment, broken down by cooperation decision in Stage 1.

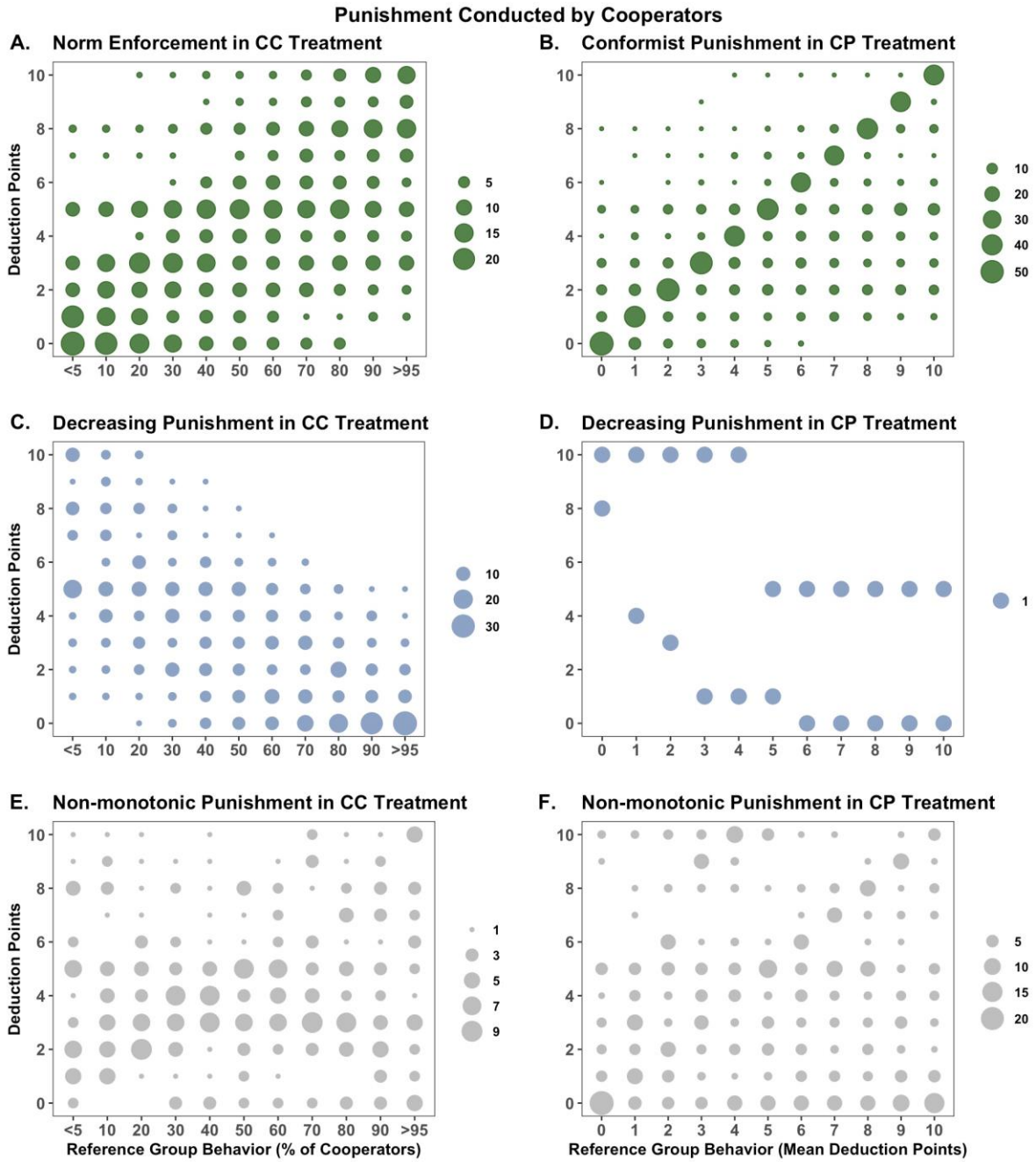
**A. Independent Punishment by Defectors in CC Treatment**



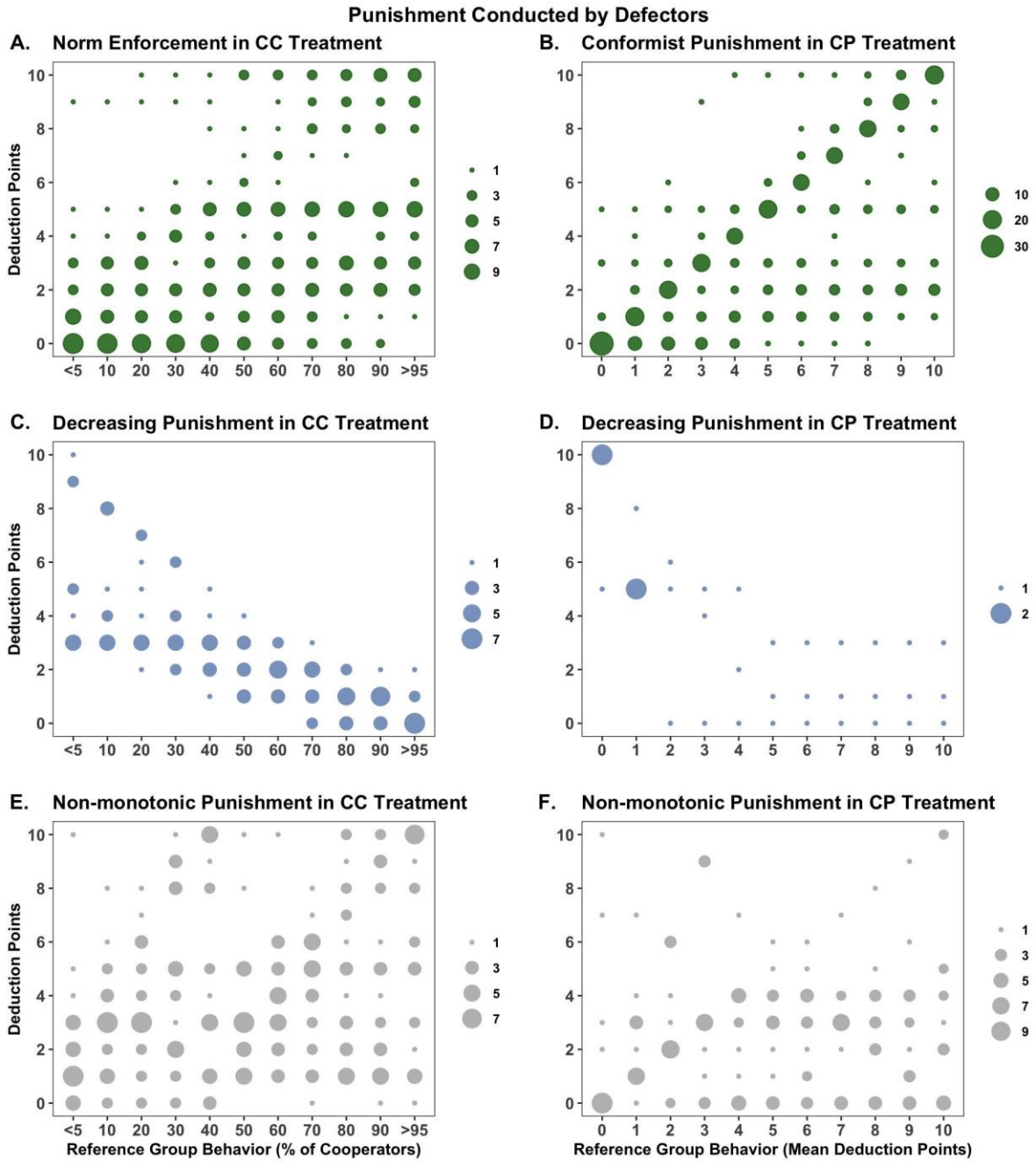
**B. Independent Punishment by Defectors in CP Treatment**



**Fig. S3.** Distribution of deduction points for independent punishers who defected in Stage 1. As in the equivalent figure showing behavior of independent punishers who cooperated in Stage 1 (main text Fig. 2A,B), vertical axes show counts. Note that, in contrast to cooperators, defectors do not equalize payoffs between themselves and their partners by assigning 8 deduction points (potentially explaining why this response was much more frequent among cooperators than among defectors).

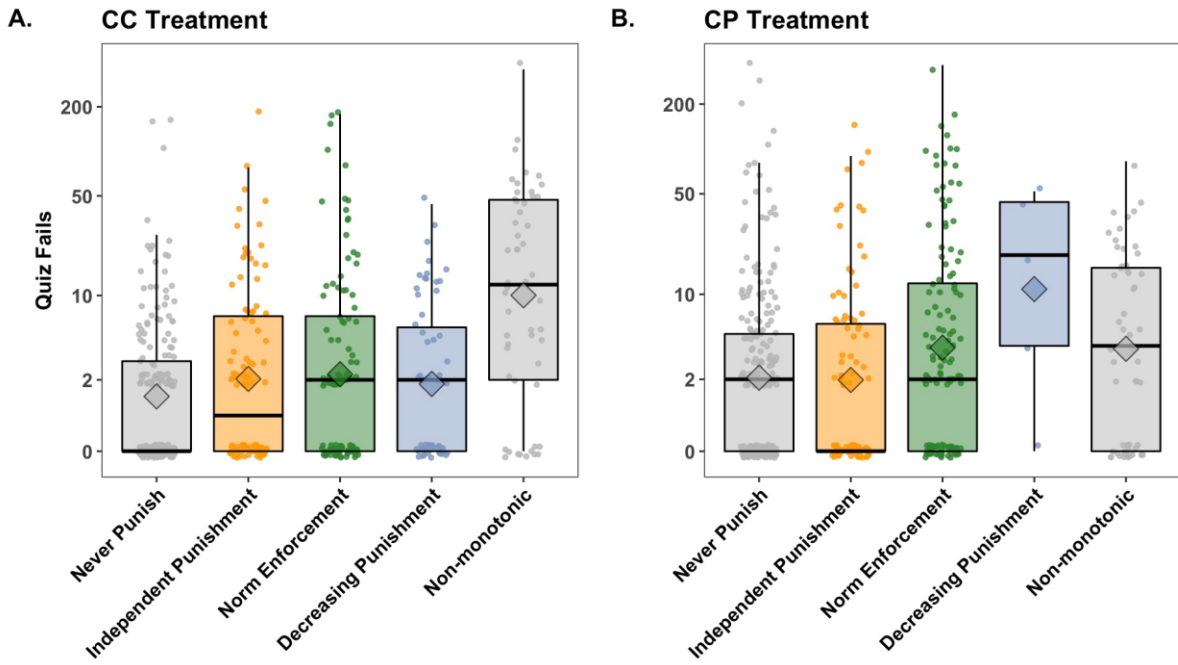


**Fig. S4.** Deduction points assigned by cooperators, broken down by treatments and punishment strategy. Sizes of dots indicate numbers of observations.

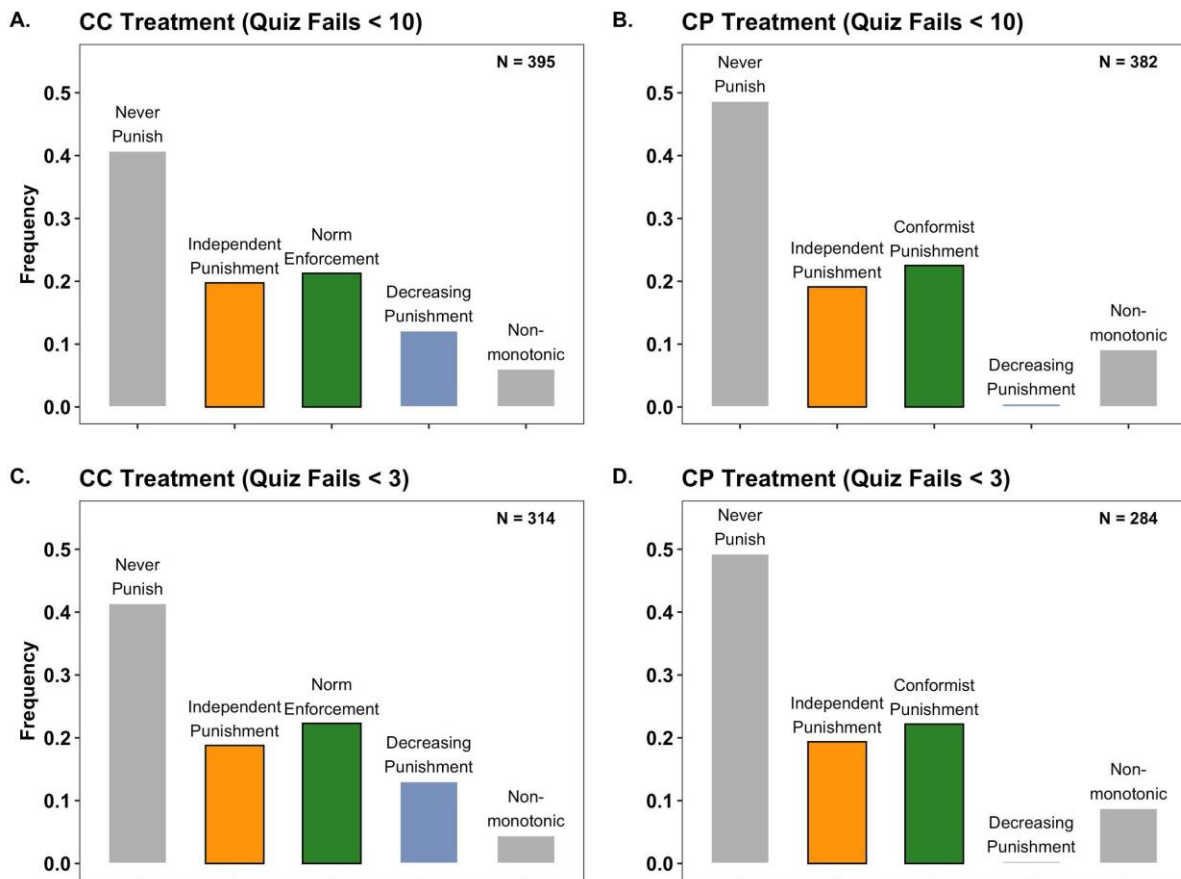


**Fig. S5.** Deduction points assigned by defectors, broken down by treatment and punishment strategy. Sizes of dots indicate numbers of observations.

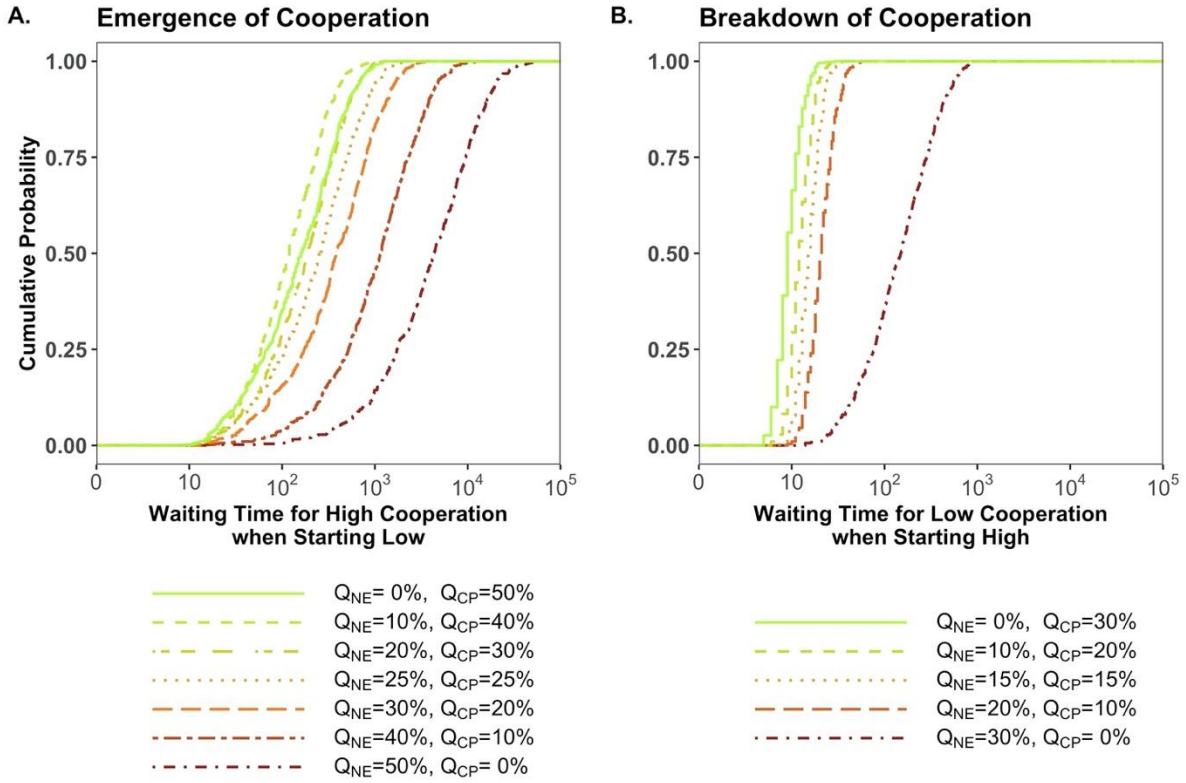




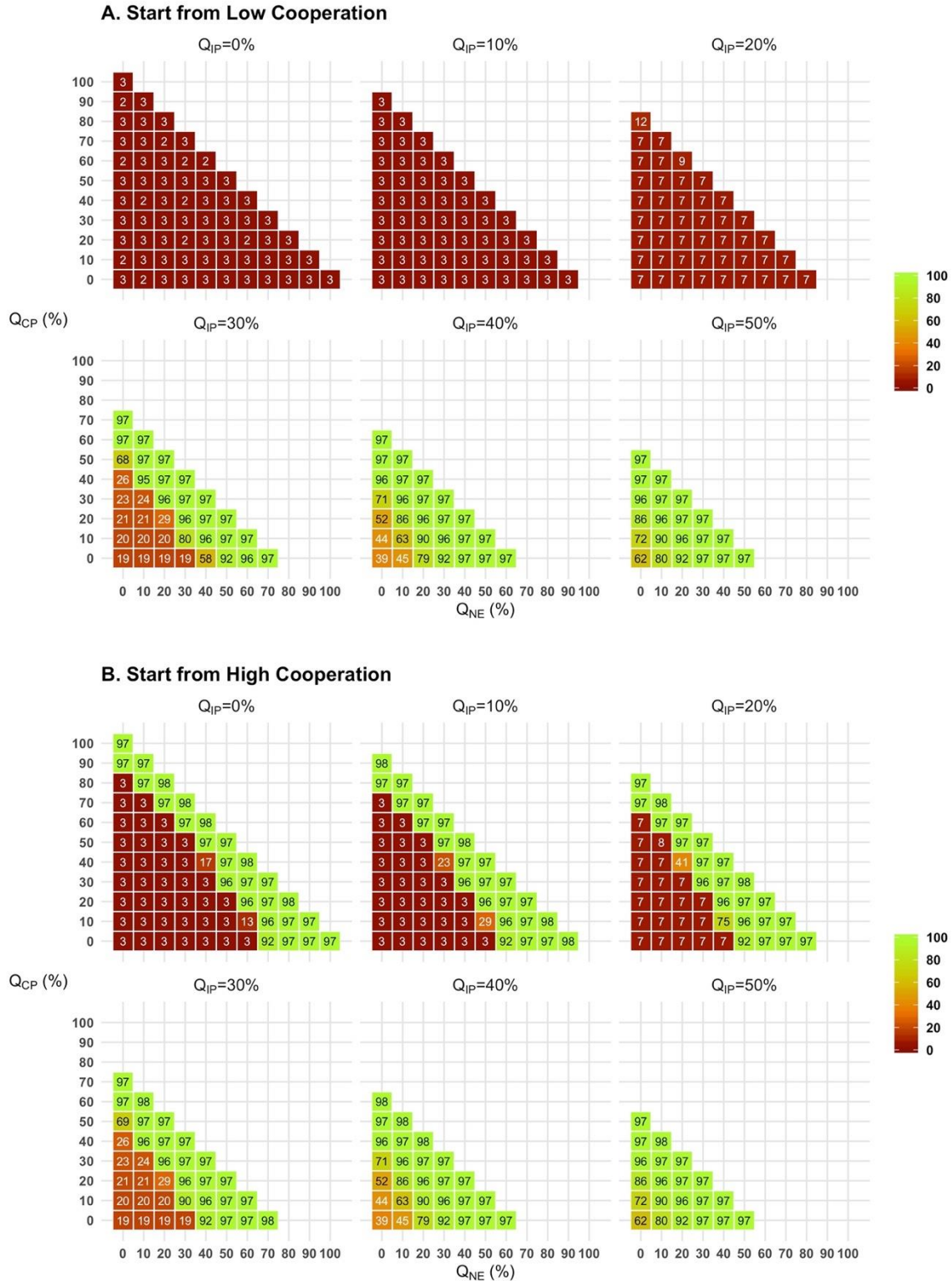
**Fig. S6.** Distributions of failed attempts at the compulsory control questions, broken down by punishment strategy in each treatment. Each jittered data point shows a participant. The y-axis is displayed in log scale to account for outliers with many attempts before passing the 9 control questions (7 of which are open questions about game payoffs). Diamond symbols show the means.



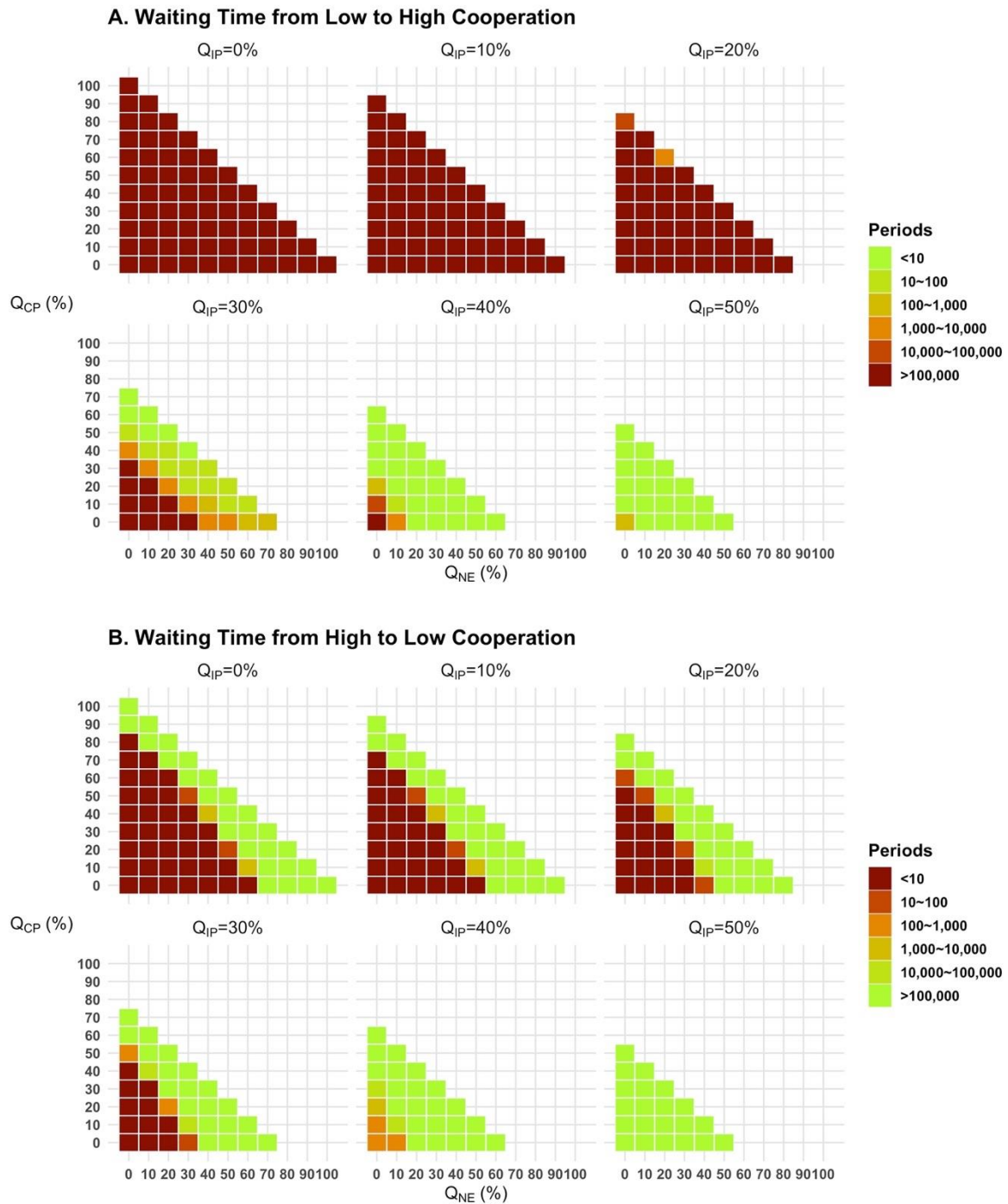
**Fig. S7.** Distributions of punishment strategies in each treatment, excluding participants with more than 10 Quiz Fails (A and B) and those more than 3 Quiz Fails (C and D), respectively. These plots suggest that the distribution of strategies remains stable when we consider different subsets of participants based on their number of attempts on the control questions.



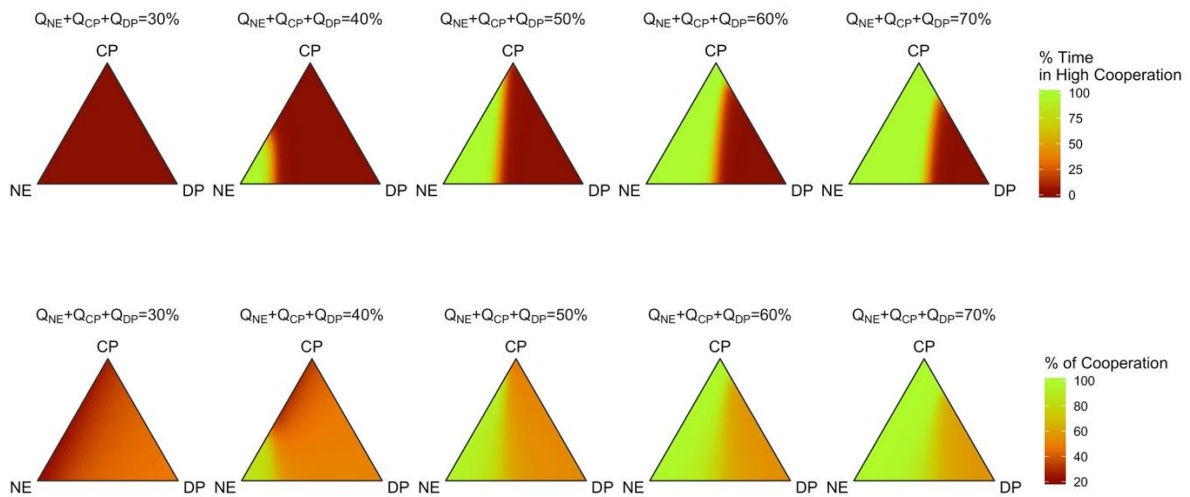
**Fig. S8.** Effects of conditional punishment strategies on the emergence and breakdown of cooperation. Lines show the cumulative probability of cooperation to rise above 75 percent (A) or fall below 25 percent (B), as a function of time. Time is shown on a logarithmic scale, and each line represents 500 simulation runs. Across both panels, we hold fixed the frequency of independent punishers at 30%.  $Q_{NE}$  is the frequency of norm enforcement, and  $Q_{CP}$  is the frequency of conformist punishment. Frequencies of these conditional strategies were chosen such that—according to our analytical results—cooperation would emerge (Panel A) or break down (panel B) in the long run. Initial beliefs regarding cooperation and punishment levels start low in Panel A ( $b_c = b_p = 0.25$ ), and high in Panel B ( $b_c = b_p = 0.75$ ; see Methods for details). Each simulation runs for 100,000 ( $10^5$ ) periods. Further simulation settings:  $n = 100$ ,  $m = 10$ ,  $u = 0.5$ ,  $\epsilon=0.05$ .



**Fig. S9. Frequency of cooperation for different population compositions and different initial conditions.** Colors (and number in the cells) indicate average percentages of cooperation across all periods of the simulation (see right hand side for color key). Each cell represents 30 simulation runs. Initial beliefs regarding cooperation and punishment start low in Panel A ( $b_c = b_p = 0.25$ ), and high in Panel B ( $b_c = b_p = 0.75$ ).  $Q_{IP}$  is the frequency of independent punishment,  $Q_{NE}$  is the frequency of norm enforcement, and  $Q_{CP}$  is the frequency of conformist punishment. Each simulation runs for 100,000 ( $10^5$ ) periods. Further simulation settings:  $n = 100$ ,  $m = 10$ ,  $u = 0.5$ ,  $\epsilon=0.05$ .



**Fig S10. Waiting time for transitions between low and high cooperation given different population compositions.** The color of each cell shows the median number of waiting periods across 30 simulation runs for cooperation to rise above 75% (Panel A) or fall below 25% (Panel B). Initial beliefs regarding cooperation and punishment start low in Panel A ( $b_c = b_p = 0.25$ ), and high in Panel B ( $b_c = b_p = 0.75$ ).  $Q_{IP}$  is the frequency of independent punishment,  $Q_{NE}$  is the frequency of norm enforcement, and  $Q_{CP}$  is the frequency of conformist punishment. The plots rely on the same simulation data used for Fig S9.



**Fig S11. Effects of 'Decreasing Punishment' (DP) (cf. Fig 1A, blue bar) on cooperation dynamics.** Agents with this strategy punish free riders if they believe cooperation rates are lower than 50% ( $b_c < 0.5$ ). Triangles show outcomes of simulations that vary the relative proportion of norm enforcement, conformist punishment, and decreasing punishment, with independent punishment fixed at 30%. The top row shows the percentage of periods for which cooperation was higher than 75% for each combination of strategies, whereas the bottom row shows the frequency of cooperation over all periods.  $Q_{DP}$  is the frequency of decreasing punishment,  $Q_{NE}$  is the frequency of norm enforcement, and  $Q_{CP}$  is the frequency of conformist punishment. Results are the average outcome of simulations where cooperation either started high ( $b_c = b_p = 0.75$ ) or low ( $b_c = b_p = 0.25$ ). In particular, for each possible combination of  $Q_{NE}$ ,  $Q_{CP}$ , and  $Q_{DP}$  averages are based on 10 simulations (5 with high and 5 with low initial beliefs). Each simulation runs for 10,000 ( $10^4$ ) periods. Further simulation settings:  $n = 100$ ,  $m = 10$ ,  $u = 0.5$ ,  $\epsilon = 0.05$ .

## Supplementary Tables

**Table S1.** Determinants of punishment.

	CC treatment (i)	CP treatment (ii)	Pooled (iii)
Reference_cooperation (cooperators in the reference group; 1 unit = 10% increase)	0.051** (0.016)		
Reference_punishment (mean deduction points of the reference group)		0.156*** (0.016)	
Cooperator (1 if cooperate in stage 1; 0 otherwise)	1.239*** (0.249)	1.776*** (0.241)	1.510*** (0.023)
Reference_cooperation × Cooperator			
Reference_punishment × Cooperator			
Reference_behavior			0.051** (0.016)
Treatment_CP			-0.689*** (0.184)
Reference_behavior × Treatment_CP			0.105** (0.023)
Male	0.199 (0.257)	0.212 (0.260)	0.193 (0.183)
Age	-0.012 (0.013)	0.019 (0.012)	0.004 (0.009)
Constant	1.907*** (0.495)	-0.272 (0.473)	1.157** (0.357)
Observations	5,434	5,478	10,912
Number of participants	494	498	992
R-square	0.038	0.095	0.064

Notes: The table displays results from OLS regression analyses of the number of deduction points that participants assigned to their defecting partners. Reference\_cooperation (Reference\_punishment) ranges from 0 to 10, corresponding to the 11 situations presented in stage 2 of the CC (CP) treatment. In the pooled regression (vii), Reference\_behavior equals to Reference\_cooperation for the CC treatment and Reference\_punishment for the CP treatment. Treatment\_CP is a dummy variable that takes the value of 1 if the participant was in the CP treatment, and 0 otherwise. Robust standard errors (in parentheses) are corrected for clustering at the participant level. Asterisks denote statistical significance. Asterisks denote statistical significance: \* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001.

Regression (i) shows that in the CC treatment, an increase of 10% of cooperators in the reference group leads to an average increase of 0.051 deduction points assigned to a defecting interaction partner (F-test:  $P < 0.01$ ). Regression (ii) shows that in the CP treatment, an average increase of one deduction point assigned in the reference group leads to an average increase of about 0.157 deduction points (F-test:  $P < 0.001$ ). Regression (iii) pools the data of both treatments. The coefficient of the interaction term 'Reference\_behavior  $\times$  Treatment\_CP is significantly positive (F-test:  $P < 0.001$ ). This suggests that, on average, participants' punishment of their interaction partner was more strongly influenced by punishment than by cooperation in the reference group. Additionally, we find that participants who cooperated themselves punish their defecting interaction partner more severely, on average, deduction points assigned by cooperators are 1.2 to 1.8 points higher than those assigned by defectors.



# Supplementary Analysis

## Contents

1. Setting
2. Strategies
3. Short-run (Nash) equilibrium
4. Long-run equilibrium

Here we use analytical methods to evaluate our model, addressing how the experimentally identified punishment strategies interact to shape the dynamics of cooperation in the long run. Section 1 describes and formalizes the interaction setting. Section 2 describes the strategies we consider. Sections 3 and 4 analyze the effects of conditional punishment strategies on cooperation in the short run and in the long run, respectively.

## 1. Setting

We consider the following decision setting, which is similar to the task used in the experiment. Two agents, A and B, are randomly drawn from a large population to play a two stage game. In Stage 1, they can either *cooperate* or *defect*. Table S2 shows how the Stage 1 material payoffs for both agents depend on their choices.

**Table S2.** Payoffs from stage 1.

		Agent B	
		<i>Cooperate</i>	<i>Defect</i>
Agent A	<i>Cooperate</i>	$a, a$	$d, e$
	<i>Defect</i>	$e, d$	$c, c$

Note: we consider a prisoner's dilemma, which is characterized by  $e > a > c > d$ . In the experiment, the values used were:  $e = 25, a = 18, c = 16, d = 9$ .

In Stage 2, each agent can punish their partner, if their partner defected in Stage 1. We depart from the experiment by considering binary punishment decisions (rather than choosing integers on a 0-10 scale). Punishment incurs a cost  $k > 0$  to the punisher and a loss  $l > 0$  to

the defector. As consistent with our experiment, our model considers punishment of defectors and ignores antisocial punishment of cooperators. The final payoffs from the game are the payoffs from Stage 1 minus the costs of conducting punishment and losses from being punished in Stage 2.

We focus on binary punishment decisions for the sake of exposition and tractability. Compared with the task in our experiment, focusing on binary punishment decisions in our model is not without loss of generality. Binary punishment excludes the possibility that an individual's punishment is not weakly monotonic—i.e., that it is neither independent, nor weakly increasing or weakly decreasing—in response to increasing cooperation rate or punishment rate in the population. Our experimental results, however, suggest that non-monotonic punishment behavior is much less common than independent punishment, norm enforcement, and conformist punishment (Fig. 1 in the main text). Furthermore, the group of participants who show non-monotonic punishment behavior becomes very small (less than 10 percent) if we exclude participants who had difficulty answering the nine compulsory control questions (Fig. S7), suggesting that such non-monotonic behavior is likely to be the result of inattentive choice behavior, rather than a real preference.

## 2. Strategies

**Cooperation.** We assume that an agent's choice to cooperate or defect depends on which choice generates the highest expected material payoffs. Let  $b_c \in [0,1]$  denote an agent's belief about the cooperation rate in the population, and  $b_p \in [0,1]$  the punishment rate. From Table S2 we can see that the expected payoff from choosing *cooperate* is

$$(1) \quad b_c a + (1 - b_c) d.$$

The expected payoff from choosing *defect* is

$$(2) \quad b_c e + (1 - b_c) c - b_p l.$$

An agent cooperates if and only if (1)  $\geq$  (2) (assuming they cooperate if expected payoffs are the same). Rearranging the terms leads to the condition

$$(3) \quad b_p \geq \theta_c \equiv \frac{1}{l} [b_c(e - a) + (1 - b_c)(c - d)].$$

This shows that an agent cooperates if and only if their beliefs of being punished if they defect ( $b_p$ ) exceeds a threshold. This threshold is linearly increasing in the temptation to defect  $b_c(e - a) + (1 - b_c)(c - d)$ , and decreasing in the loss from being punished  $l$ . In the analysis

presented in the main text, we assume  $\theta_C = 0.5$ . Here, we consider the general case of arbitrary threshold values.

**Punishment.** Our implementation of punishment strategies is informed by our experimental results. We consider four distinct ‘types’ of agents: *i) independent punishers* who punish independently of  $b_c$  and  $b_p$ , *ii) norm enforcers* who punish if and only if  $b_c$  is high enough, and *iii) conformist punishers* who punish if and only if  $b_p$  is high enough; and *iv) non-punishers*, who never punish. For simplicity and ease of illustration, we assume that the four strategies above are mutually exclusive and stable: each individual has a unique strategy that doesn’t change over time.

The frequencies of punishment types in the population—*independent punishers* ( $Q_{IP}$ ), *norm enforcers* ( $Q_{NE}$ ), *conformist punishers* ( $Q_{CP}$ ) and *non-punishers* ( $Q_0$ )—sum up to 1. Agents do not know the punishment strategy of their interaction partners. Norm enforcers punish if and only if  $b_c \geq \theta_{NE}$ . Conformist punishers punish if and only if  $b_p \geq \theta_{CP}$ . In the main text, we assume  $\theta_{NE} = \theta_{CP} = 0.5$ . Here we will consider the case of arbitrary threshold values.

We characterize short-run (Nash) equilibria and long-run equilibria of the game. First, we show that there are often two Nash equilibria: one in which all agents cooperate, and another in which all agents defect. Second, we analyze how the relative frequencies of the different punishment strategies ( $Q_{IP}, Q_{NE}, Q_{CP}, Q_0$ ) affect the likelihood of either equilibrium to emerge and persist in the long run. Finally, we construct a stochastic model to examine how dynamics of cooperation and punishment may evolve over time.

### 3. Short-run (Nash) equilibrium

Proposition 1 shows the conditions under which cooperation can be sustained in the short run. Agents do not know the type of agent with whom they are matched. As is standard in the economic literature, we assume that agents have a common prior on the population composition, which corresponds to  $(Q_{IP}, Q_{NE}, Q_{CP}, Q_0)$ . In the next section ‘Long-run equilibrium’, we will address the problem of how agents form and update beliefs over time. Exogenous payoff parameters of the game determine the equilibria through their effects on the thresholds  $\theta_C, \theta_{NE}$ , and  $\theta_{CP}$ .

#### Proposition 1. (Nash equilibrium)

1. If  $Q_0 > 1 - \frac{n-1}{n} \theta_C$ , then in every Nash equilibrium all agents defect.

2. *If  $Q_{IP} \geq \frac{1}{n} + \frac{n-1}{n}\theta_C$ , or if  $Q_{IP} > \theta_{CP}$  and  $Q_{IP} + Q_{CP} \geq \frac{1}{n} + \frac{n-1}{n}\theta_C$ , then in every Nash equilibrium all agents cooperate.*
3. *If  $Q_0 < 1 - \frac{1}{n} - \frac{n-1}{n}\max\{\theta_C, \theta_{CP}\}$ , then there exists a Nash equilibrium in which all agents cooperate, and all agents (except non-punishers) punish defectors.*
4. *If  $Q_{IP} < \frac{n-1}{n}\min\{\theta_C, \theta_{CP}\}$ , then there exists a Nash equilibrium in which all agents defect, and only those punishing independently punish defectors.*

The proof of the proposition is provided at the end of this section. The proposition states that, first, if there are many agents who do not punish ( $Q_0$  exceeds a critical threshold), then all agents defect in equilibrium. Second, it states that if there is a high enough level of independent punishment such that cooperation is the payoff maximizing choice for all individuals ( $Q_{IP} \geq \frac{1}{n} + \frac{n-1}{n}\theta_C$ ), then all agents cooperate in equilibrium. Similarly, if there are enough independent punishers such that their behavior triggers punishment by conformist punishers ( $Q_{IP} > \theta_{CP}$ ), and their joint number is high enough to make cooperation the payoff maximizing choice for all agents ( $Q_{IP} + Q_{CP} \geq \frac{1}{n} + \frac{n-1}{n}\theta_C$ ), then all agents cooperate in equilibrium as well.

The third statement of the proposition says that if there are not sufficiently many non-punishers ( $Q_0 < 1 - \frac{1}{n} - \frac{n-1}{n}\max\{\theta_C, \theta_{CP}\}$ ), then there exists a Nash equilibrium in which all agents cooperate and all agents (apart from non-punishers) punish defectors. The fourth statement says that if there are not sufficient independent punishers to either make cooperation the payoff maximizing choice (either by themselves, or in unison with conformist punishers), then there exists a Nash equilibrium in which all agents defect.

Together, the third and the fourth statement imply that when both independent punishers and non-punishers occur at intermediate frequencies, cooperation and defection can both emerge as Nash equilibria. In the following section, and in the simulations reported in the main text, we focus on situations where there are multiple equilibria, and examine how (conditional and unconditional) punishment strategies affect which of these equilibria is selected in the long run.

### **Proof of Proposition 1.**

- (1) By contradiction: Suppose there is a (Nash) equilibrium in which some agent cooperates. Then for this agent,  $b_p \geq \theta_C$  where  $b_p$  is the proportion of those who punish among the other agents. Note  $b_p \leq \frac{(1-Q_0)n}{n-1}$ , where  $(1 - Q_0)n$  is an upper bound

on the number of agents who punish, and  $n - 1$  is the number of all other agents. Given  $Q_0 > 1 - \frac{n-1}{n}\theta_C$ , however, we have  $\frac{(1-Q_0)n}{n-1} < \theta_C$ , contradicting with  $b_p \geq \theta_C$ .

(2) We show the contrapositive: Suppose there is an equilibrium in which an agent defects.

Then for the agent,  $b_p < \theta_C$ , where  $b_p$  is at least  $\frac{Q_{IP}n-1}{n-1}$ . Hence  $\frac{Q_{IP}n-1}{n-1} < \theta_C$ , implying  $Q_{IP} < \frac{1}{n} + \frac{n-1}{n}\theta_C$ .

Next, suppose  $Q_{IP} > \theta_{CP}$ . Then both independent punishers and conformist punishers punish. Hence, for each agent, the proportion of those who punish among the others is at least  $\frac{(Q_{IP}+Q_{CP})n-1}{n-1}$ . By  $Q_{IP} + Q_{CP} \geq \frac{1}{n} + \frac{n-1}{n}\theta_C$ , we have  $\frac{(Q_{IP}+Q_{CP})n-1}{n-1} \geq \theta_C$ . Thus every agent cooperates.

(3) Consider the strategy profile such that all agents cooperate, and all agents (except non-punishers) punish defectors. To show that this is an equilibrium, we check the best-response of each agent. First, consider each agent's cooperation decision. The specified condition implies  $b_p \geq \frac{(Q_{IP}+Q_{NE}+Q_{CP})n-1}{n-1} \geq \theta_C$  (a). Hence it is each agent's best response to cooperate. Second, given that everyone cooperates, it is each norm enforcer's best response to punish any defector. And by definition, each independent punisher also punishes. Third, it is each conformist punisher's best response to punish if  $\frac{(Q_{IP}+Q_{NE}+Q_{CP})n-1}{n-1} \geq \theta_{CP}$  (b). The condition  $Q_0 < 1 - \frac{1}{n} - \frac{n-1}{n}\max\{\theta_C, \theta_{CP}\}$  implies both (a) and (b). This establishes the statement.

(4) Consider the strategy profile such that all agents defect, and no agent (except independent punishers) punishes defectors. We prove the statement by checking each agent's best responses. First,  $\theta_{IP} < \frac{n-1}{n}\theta_C$  implies  $\frac{\theta_{IP}n}{n-1} < \theta_C$ . Hence it is each agent's best response to defect. Second, that all agents defect implies that it is each norm enforcer's best response to not punish defectors. Third,  $\theta_{IP} < \frac{n-1}{n}\theta_{CP}$  implies  $\frac{\theta_{IP}n}{n-1} < \theta_{CP}$ . Hence it is each conformist punisher's best response to not punish. This completes the proof.

*Q.E.D.*

### 3. Long-run equilibrium

In this section we examine the long-run effects of conditional and unconditional punishment strategies on cooperation. We aim to delineate the conditions under which conditional punishment strategies (norm enforcement and conformist punishment) will, in the long run,

cause the population to be in or around the cooperation equilibrium for most of the time. Our analysis builds on (1–3).

We consider discrete time periods:  $t = 0, 1, 2, \dots, T$ . In each period, agents are randomly matched and interact in the two-stage game described in Section 1 above. An agent's punishment strategy and the population composition  $(Q_{IP}, Q_{NE}, Q_{CP}, Q_0)$  are fixed over time, but agents may update their cooperation and punishment decisions as their beliefs  $b_c$  and  $b_p$  change. In each period agents react to their beliefs 'myopically' to maximise their expected payoffs in that period.

To be more precise, each period involves two subsequent classes of events:

- I. *Updating beliefs.* In each period  $t \geq 1$ , each agent updates their beliefs with probability  $u$ , with  $0 < u < 1$ . Belief updating works as follows. The agent randomly samples  $m$  agents from the population, with  $0 < m \leq n$ . She counts how many agents in the sample cooperated and would punish according to their strategies in the previous period, and divide the counts by  $m$ . The results become their beliefs  $b_c$  and  $b_p$  in the current period.
- II. *Responding myopically to beliefs.* An agent cooperates in a period if and only if they have belief  $b_p \geq \theta_c$ . Punishment decisions are determined according to the agents' types (as specified in Section 2 above).

With a high probability, an agent's decisions are implemented according to the rules stated above. With small probability  $\varepsilon \geq 0$ , however, an agent makes a mistake ("tremble"). A mistake implies that the agent randomly selects a cooperative action or a punishment action. We assume that mistakes are independent across periods, agents, and across cooperation and punishment decisions. Following (1–3), we refer to the dynamic with  $\varepsilon > 0$  as *the stochastic dynamic*, and the dynamic with  $\varepsilon = 0$  as *the best-response dynamic*.

We first analyze the stochastic dynamic in the case of  $m = n$ ,  $T \rightarrow \infty$ , and  $\varepsilon \rightarrow 0$  using analytical methods. As previous studies of the same class of stochastic dynamics show (Kandori et al. 1993; Young 1993, 1998), whether  $m < n$  or  $m = n$  does not affect stationary distributions of the dynamics. Later, we also conduct simulations to explore the cases of small sample size  $m$ , finite  $T$ , and non-negligible  $\varepsilon$ .

Our analytical results aim to characterize the set of long-run equilibria. These are the equilibria that have a positive frequency in the stationary distribution of the stochastic dynamic when the probability of mistakes is vanishingly small. A long run equilibrium is formally defined as follows. Let  $s$  be a population state specifying the cooperation decision and punishment decision of each agent in the population. Let  $S$  denote the set of all population states. Let  $P^\varepsilon \in \Delta(S)$

denote the stationary distribution of the stochastic dynamic under  $\varepsilon > 0$  and  $m = n$ . The stochastic dynamic is an irreducible Markov chain on the finite state space  $S$ . Hence  $P^\varepsilon$  exists and is unique for each  $\varepsilon$ . We obtain  $P^\varepsilon$  by taking  $T \rightarrow \infty$ . Let  $P \equiv \lim_{\varepsilon \rightarrow 0} P^\varepsilon$  denote the limit distribution as  $\varepsilon$  approaches zero. A state  $s$  is a *long-run equilibrium* if  $P(s) > 0$  (1–3). If a state is a unique long-run equilibrium for sufficiently large  $n$ , then it is a *generically unique long-run equilibrium*.

For the sake of exposition and analytical tractability, we restrict our attention to the parameter ranges specified by the Assumptions below.

**Assumption 1.**

- 1)  $Q_{IP} < \min\{\theta_C, \theta_{CP}\}$ ,  $\theta_{NE} < 1$ , and  $Q_0 < 1 - \max\{\theta_C, \theta_{CP}\}$ ;
- 2) either (i)  $Q_{NE} \geq |\theta_C - \theta_{CP}|$  and  $Q_{CP} \geq |\theta_C - \theta_{CP}|$ , or (ii)  $Q_{NE} \leq |\theta_C - \theta_{CP}|$  and  $Q_{CP} \leq |\theta_C - \theta_{CP}|$ ;
- 3) either (i)  $\theta_{CP} \leq Q_{IP} + Q_{NE}$  and  $\theta_{CP} \leq Q_{IP} + Q_{CP}$ , or (ii)  $\theta_{CP} \geq Q_{IP} + Q_{NE}$  and  $\theta_{CP} \geq Q_{IP} + Q_{CP}$ .

Assumption (1) restricts our attention to cases where the following Nash equilibria both exist (see Proposition 1): *the defection equilibrium* in which all agents defects and only independent punishers punish defectors; and *the cooperation equilibrium* in which all agents cooperate and all agents (except non-punishers) punish defectors. The remaining two assumptions greatly reduce the number of cases we need to consider, but still allow us to obtain the key intuitions from the model. Specifically, assumption (2) holds that the proportions of conditional punishers ( $Q_{NE}$  and  $Q_{CP}$ ) are both either high or low. Assumption (3) holds that the value of  $\theta_{CP}$  is either high or low, compared to the number of punishers.

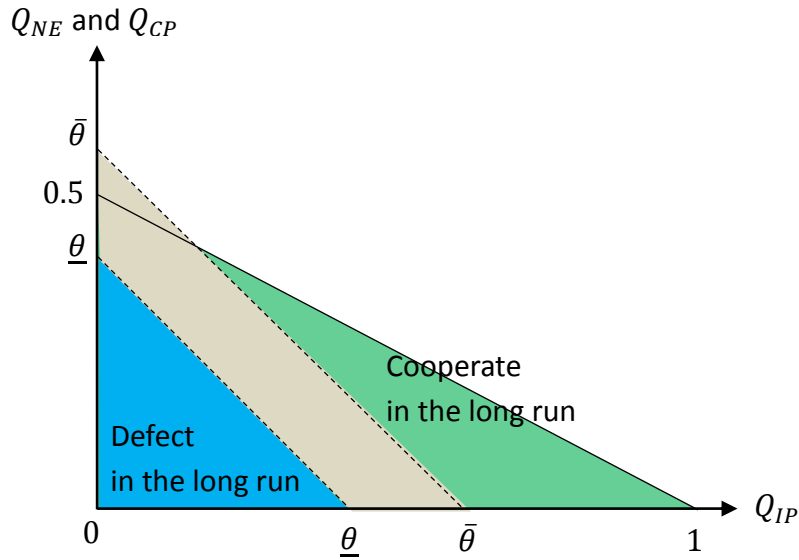
Now we can state the proposition about the long-run equilibrium of the stochastic dynamic. It shows how norm enforcement and conformist punishment interact with independent punishment to affect cooperation in the long run.

**Proposition 2. (Long-run equilibrium)** Suppose  $m = n$  and Assumption 1 hold. Let  $\underline{\theta} \equiv$

$\min\{\frac{1}{2}(\theta_C + \theta_{CP}), 2\theta_{NE} + \theta_{CP} - 1\}$  and  $\bar{\theta} \equiv \max\{\frac{1}{2}(\theta_C + \theta_{CP}), 2\theta_{NE} + \theta_{CP} - 1\}$ .

1. If  $Q_{IP} + Q_{NE} > \bar{\theta}$  and  $Q_{IP} + Q_{CP} > \bar{\theta}$ , then the cooperation equilibrium is the generically unique long-run equilibrium.
2. If  $Q_{IP} + Q_{NE} < \underline{\theta}$  and  $Q_{IP} + Q_{CP} < \underline{\theta}$ , then the defection equilibrium is the generically unique long-run equilibrium.

The proofs for Proposition 2 are provided at the end of this section. Fig. S12 illustrates the proposition, which says that, together with the independent punishment, conditional punishment can support cooperation as the generically unique long-run equilibrium. If the frequencies of independent punishment and conditional punishment are both low, then the cooperation equilibrium cannot be sustained in the long run.



**Fig. S12.** Illustration of Proposition 2. When  $Q_{IP} + Q_{NE} > \bar{\theta}$  and  $Q_{IP} + Q_{CP} > \bar{\theta}$ , we expect to see the cooperation equilibrium in the long run. When  $Q_{IP} + Q_{NE} < \underline{\theta}$  and  $Q_{IP} + Q_{CP} < \underline{\theta}$ , we expect to see the defection equilibrium in the long run. When  $\theta_c = \theta_{NE} = \theta_{CP} = 0.5$ , we have  $\underline{\theta} = \bar{\theta} = 0.5$ .

Nevertheless, the characterization by Proposition 2 is incomplete. It is silent about the case of  $\underline{\theta} < Q_{IP} + Q_{NE} < \bar{\theta}$  or  $\underline{\theta} < Q_{IP} + Q_{CP} < \bar{\theta}$ . Proposition 3 below provides precise cut-off conditions for the long-run equilibrium for the special case where  $\theta_c = \theta_{NE} = \theta_{CP} = \frac{1}{2}$ , which is also the set of parameters we use in our simulations presented in the main text. The assumption that all thresholds are equal to a half is somewhat arbitrary. As stated in Section 2 of this Supplement, the threshold  $\theta_c$  is determined by exogenous payoff parameters. Hence, setting it equal to a half comes down to considering a subset of the potential payoff space. For  $\theta_{NE}$  and  $\theta_{CP}$ , however, a threshold of a half makes intuitive sense. For norm enforcement, it is in line with the idea that people will judge the more common behavior as the more moral one, and act to enforce it (4). For conformist punishment, it states that these agents follow the behavior of the majority. Furthermore, our focus here is not on the comparative statics with respect to these thresholds, but rather on how the population composition (with respect to punishment strategies) affects cooperation dynamics. In this regard, the proposition below is illuminating.



**Proposition 3.** Suppose  $m = n$  and  $\theta_C = \theta_{NE} = \theta_{CP} = \frac{1}{2}$ . Then

1. If  $Q_{IP} + \frac{1}{2}(Q_{NE} + Q_{CP}) > \frac{1}{2}$ , then the cooperation equilibrium is the generically unique long-run equilibrium;
2. If  $Q_{IP} + \frac{1}{2}(Q_{NE} + Q_{CP}) < \frac{1}{2}$ , then the defection equilibrium is the generically unique long-run equilibrium.

The proof for Proposition 3 is provided at the end of this section. When  $\theta_C = \theta_{NE} = \theta_{CP} = \frac{1}{2}$ , we already have  $\underline{\theta} = \bar{\theta} = \frac{1}{2}$  in the conditions specified in Proposition 2. Proposition 3 further reveals that: if and only if independent punishers and the average number of norm enforcers and conformist punishers together add up to over a half of the population, then the cooperation equilibrium will be the only state that occurs with positive probability  $P(s) > 0$  in the long run. That is, the average frequency of norm enforcement and conformist punishment is important to support cooperation in the long run; it is as important as the role played by independent punishment.

**Remarks.** Economists have used myopic best-response stochastic dynamics to study bargaining norms (5), customs in economic contracts (6), evolution of altruism (7), the selection of coordination actions in social networks (8, 9), diffusion of innovations (10, 11), and the evolution of cooperation strategies in repeated games (12). In particular, (1–3) show that many details of these dynamics do not affect their stationary distributions when  $\varepsilon \rightarrow 0$ . In particular, the stationary distribution is not affected by the value of the updating probability  $u$  as long as  $0 < u < 1$ , or the sample size  $m$  as long as  $m$  does not become too small to affect the tipping thresholds, or the probability distribution used to pick actions when making mistakes.

Assuming  $u < 1$  means that it will not occur that all agents update simultaneously in a period. If  $u = 1$  and  $\varepsilon$  is small, then besides the cooperation equilibrium and the defection equilibrium, the population can also be trapped in a loop of jumping back and forth between two states: in one, all agents defect and all punish defectors except for the non-punishers; in the other, all agents cooperate but no one would punish defectors except for the independent punishers. We exclude this possibility to focus on the transitions between the cooperation equilibrium and the defection equilibrium characterized by Proposition 1.

### **Proof of Proposition 2.**

**Preliminaries.** First, we introduce necessary terminology for our proof (see, e.g., Young (1998) for a more extensive discussion). An *absorbing set* (of the best-response dynamic) is a subset

of states  $X \subset S$  such that (i) if the the best-response dynamic starts from a state in  $X$  then it stays within  $X$  with probability 1, and (ii) for any  $s, s' \in X$ , there is a positive probability of transiting from  $s$  to  $s'$  within a finite number of periods. If an absorbing set contains only one state, then we call the state an *absorbing state*. A *transition path from  $s$  to  $s'$*  is a finite sequence of states,  $s_1, s_2, \dots, s_K \in S$ , with  $s_1 = s, s_K = s'$ , and  $s_k \neq s_{k+1}$  for each  $1 \leq k < K$ . The *cost* of a transition path, denoted by  $cost(s_1, s_2, \dots, s_K)$ , is the number of mistakes (choices that are not best responses) that occur along the path.

We use stochastic trees to represent minimum transition costs between absorbing sets. A *stochastic tree* is a directed tree with each absorbing set as a vertex. The directed edges in a stochastic tree represent transitions among absorbing sets. Each edge is weighted by the minimum number of mistakes required to transit from one absorbing set to another. An absorbing set is said to be at the root of a stochastic tree if there is no edge (with positive weight) leading from it to other absorbing sets in the tree. The cost of a stochastic tree is the sum of the weights of all its edges. Our proof applies the following theorem:

**Theorem** (Young (2)).

1. A state is a long-run equilibrium only if it is contained in an absorbing set.
2. If an absorbing state is at the root of the stochastic tree that strictly minimizes the cost among all stochastic trees, then the state is the unique long-run equilibrium.

The best-response dynamic in our model has only two absorbing sets: one consisting of defection equilibrium, and the other consisting of the cooperation equilibrium. With abuse of notation, we denote them by  $D$  and  $C$ , respectively. By Young's theorem,  $D$  and  $C$  are the only candidates for a long-run equilibrium.

We can construct two stochastic trees:  $D \rightarrow C$  (a directed line with  $D$  and  $C$  as its two vertices connected by a unique edge leading from  $D$  to  $C$ ) and  $C \rightarrow D$ . Let  $M_{C \rightarrow D}$  denote the minimum number of mistakes required to transit from  $C$  to  $D$ . More precisely,  $M_{C \rightarrow D}$  is the minimum value of  $cost(s_1, s_2, \dots, s_K)$  among the set of all paths  $s_1, s_2, \dots, s_K$  with  $s_1 = C$  and  $s_K = D$ . Likewise,  $M_{C \rightarrow D}$  is the minimum value of  $cost(s_1, s_2, \dots, s_K)$  among the set of all paths  $s_1, s_2, \dots, s_K$  with  $s_1 = D$  and  $s_K = C$ . By Young's theorem, it suffices to compare  $M_{C \rightarrow D}$  with  $M_{C \rightarrow D}$  to determine the long-run equilibrium.

**Transition paths.** Now we examine transition paths with minimum costs between  $D$  and  $C$ . Three paths are relevant to determine the minimum cost of transitions from  $D$  to  $C$ :

- Path E1 ('E' for Emergence of cooperation): Starting from  $D$  at time  $t = 0$ , if  $\theta_C - Q_{IP} < Q_0$ , then let  $[(\theta_C - Q_{IP})n]$  non-punishers punish defectors *by mistake* at  $t = 1$  (for any real

number  $x$ ,  $\lceil x \rceil$  is the lowest integer equal to or greater than  $x$ ). If  $\theta_C - Q_{IP} \geq Q_0$ , then let all non-punishers and  $\lceil (\theta_C - Q_{IP} - Q_0)n \rceil$  conformist punishers punish by *mistake* at  $t = 1$ . At  $t = 2$ , let all agents update their cooperation decision. Then they all cooperate (for brevity, if we do not explicitly mention that agents update their cooperation or punishment decision, then the agents do not update from the last period, and do not make any mistakes). At  $t = 3$ , let all norm enforcement agents update their punishment decision. Then they all punish. It follows that following  $t = 3$ , all agents cooperate, and  $\lceil (\theta_C + Q_{NE})n \rceil$  agents punish defectors. For a real number  $x$ , we write  $x^+ \equiv \max\{0, x\}$ . At  $t = 4$ , let  $\Delta_1^E \equiv \lceil (\theta_{CP} - \theta_C - Q_{NE})^+ n \rceil$  agents who do not punish at  $t = 3$  start to punish by *mistake*. Finally, at  $t = 5$ , let all conformist punishers update their punishment decision. Then by requiring all agents to update both cooperation and punishment decisions at  $t = 6$ , we reach  $C$ . Counting the number of mistakes, we obtain the cost of path E1:  $cost(E1) = \lceil (\theta_C - Q_{IP})n \rceil + \Delta_1^E$ , where  $\Delta_1^E = 0$  if  $\min\{Q_{NE}, Q_{CP}\} \geq |\theta_C - \theta_{CP}|$ .

- Path E2: Starting from  $D$  at  $t = 0$ , if  $\theta_{CP} - Q_{IP} < Q_0 + Q_{NE}$ , let  $\lceil (\theta_{CP} - Q_{IP})n \rceil$  non-punishers or norm enforcers punish by *mistake* at  $t = 1$ . If  $\theta_{CP} - Q_{IP} \geq Q_0 + Q_{NE}$ , then let all non-punishers, all norm enforcers, and  $\lceil (\theta_{CP} - Q_{NE} - Q_0)n \rceil$  conformist punishers punish by *mistake* at  $t = 1$ , resulting in  $\lceil \theta_{CP}n \rceil$  agents punishing. At  $t = 2$ , let all conformist punishers update their punishment decision. Then all conformist punishers punish. At  $t = 3$ , let  $\Delta_2^E \equiv \lceil (\theta_C - \theta_{CP} - Q_{CP})^+ n \rceil$  agents who do not punish at  $t = 2$  start to punish by *mistake*. At  $t = 4$ , let all agents update their cooperation decision. Then all agents cooperate. At  $t = 5$ , let all norm enforcers update their punishment decision and start to punish. By requiring all agents to update both cooperation and punishment decisions at  $t = 6$ , we reach  $C$ . The cost of path E2 is  $cost(E2) = \lceil (\theta_{CP} - Q_{IP})n \rceil + \Delta_2^E$ , with  $\Delta_2^E = 0$  if  $\min\{Q_{NE}, Q_{CP}\} \geq |\theta_C - \theta_{CP}|$ .
- Path E3: Starting from  $D$  at  $t = 0$ , let  $\lceil \theta_{NE}n \rceil$  agents cooperate by *mistakes* at  $t = 1$ . At  $t = 2$ , let all norm enforcers update their punishment decision. Then all norm enforcers will punish defectors. At  $t = 3$ , let  $\Delta_3^E \equiv \lceil (\theta_{CP} - Q_{IP} - Q_{NE})^+ n \rceil$  agents who did not punish in the last period start to punish by *mistakes*. At  $t = 4$ , let all conformist punishers update their punishment decision. Then all conformist punishers will punish. At  $t = 5$ , let all agents update their cooperation decision. Then by requiring all agents to update both cooperation and punishment decisions in  $t = 6$ , we reach  $C$ . The cost of this is path is  $cost(E3) = \lceil \theta_{NE}n \rceil + \Delta_3^E$ , where  $\Delta_3^E = 0$  if  $Q_{IP} + Q_{NE} \geq \theta_{CP}$ .

Correspondingly, the following three paths are relevant to compute the minimum cost of transiting from  $C$  to  $D$ :

- Path B1 ('B' for Breakdown of cooperation): Starting from  $C$  at  $t = 0$ , if  $[(1 - \theta_C - Q_0)n] \leq Q_{IP}$ , let  $[(1 - \theta_C - Q_0)n]$  independent punishers stop punishing by *mistake* at  $t = 1$ . If  $[(1 - \theta_C - Q_0)n] > Q_{IP}$ , let all independent punishers and  $[(1 - \theta_C - Q_0 - Q_{IP})n]$  conformist punishers stop punishing by *mistakes* at  $t = 1$ . At  $t = 2$ , let all agents update their cooperation decision. Then all agents now defect. At  $t = 3$ , let all norm enforcement punishers update punishment decisions and stop punishing. At  $t = 4$ , let all  $\Delta_1^B \equiv [(\theta_C - \theta_{CP} - Q_{NE})^+ n]$  agents who punish in  $t = 3$  stop punishing by *mistake*. At  $t = 5$ , let all conformist punishers update punishment decisions and stop punishing. Then by requiring all agents to update cooperation decisions as well as punishment decisions at  $t = 6$ , we reach  $D$ . The cost of this path is  $cost(B1) = [(1 - \theta_C - Q_0)n] + \Delta_1^B$ , where  $\Delta_1^B = 0$  when  $\min\{Q_{NE}, Q_{CP}\} \geq |\theta_C - \theta_{CP}|$ .
- Path B2: Starting from  $C$  at  $t = 0$ , let  $[(1 - \theta_{CP} - Q_0)n]$  agents who punish at  $t = 0$  stop punishing by *mistake* at  $t = 1$ . At  $t = 2$ , let all conformist punishers update punishment decision. They all stop punishing. At  $t = 3$ , let  $\Delta_2^B \equiv [(\theta_{CP} - \theta_C - Q_{CP})^+ n]$  agents who punish at  $t = 2$  stop punishing by *mistake*. At  $t = 4$ , let all agents update cooperation decisions and start to defect. At  $t = 5$ , let all norm enforcers update punishment decisions and stop punishing. Then we reach  $D$  by requiring all agents update both decisions at  $t = 6$ . The cost of this path is  $cost(B2) = [(1 - \theta_{CP} - Q_0)n] + \Delta_2^B$ , where  $\Delta_2^B = 0$  when  $\min\{Q_{NE}, Q_{CP}\} \geq |\theta_C - \theta_{CP}|$ .
- Path B3: Starting from  $C$  at  $t = 0$ , let  $[(1 - \theta_{NE})n]$  agents defect by *mistake* at  $t = 1$ . At  $t = 2$ , let all norm enforcers update punishment decisions and stop punishing. At  $t = 3$ , let  $\Delta_3^B \equiv [(1 - \theta_{CP} - Q_0 - Q_{NE})^+ n]$  agents who punish at  $t = 2$  stop punishing by *mistake*. Note  $\Delta_3^B$  can also be expressed by  $\Delta_3^B = [(Q_{IP} + Q_{CP} - \theta_{CP})^+ n]$ . By requiring all agents to update both decisions at  $t = 3$ , we reach  $D$ . The cost of this path is  $cost(B3) = [(1 - \theta_{NE})n] + \Delta_3^B$ , where  $\Delta_3^B = 0$  if  $Q_0 + Q_{NE} \geq 1 - \theta_{CP}$ .

**Simplifying observations.** We need to determine the path with minimum cost among the six paths above. This requires solving a set of linear inequalities. Two observations simplify our calculations. First, since we are only concerned with generically unique long-run equilibria, it is both sufficient and necessary for the minimum cost path to have *strictly* lower cost than all transition paths of the opposite direction for *infinitely many*  $n$ . A sufficient and necessary condition for this is that there is a finite  $n$  under which all relevant inequalities for pairwise cost comparisons hold strictly. This condition is equivalent to having all relevant inequalities holding strictly when we ignore all “[.]” brackets, i.e., by ignoring the “least integer greater than” operator. To see the equivalence, first, suppose  $[xn] < [yn]$  for some positive  $n$ . Then

obviously  $xn < yn$ . Conversely, suppose  $xn < yn$  for some positive  $n$ . Then  $x < y$ , and there is some large enough integer  $r$  such that  $(y - x)r > 1$ , implying  $xr + 1 < yr$ . Thus  $\lceil xr \rceil < \lceil yr \rceil$ , and  $\lceil xn \rceil < \lceil yn \rceil$  for all  $n > r$ .

Second, after removing all “[.]” brackets, the costs of the paths listed above are all multiplications of  $n$ . Taking the two observations together, it suffices to consider their *relative costs*  $\widehat{cost}(\cdot) \equiv cost(\cdot)/n$  and ignore all “[.]” operators. Henceforth we will focus on  $\widehat{cost}(\cdot)$  and remove all “[.]” operators. The six transition paths and their relative costs  $\widehat{cost}$  are summarized in Table S4 below.

**Table S4.** Transition paths and relative costs.

	Relative costs $\widehat{cost}$
Transition paths from $D$ to $C$	
E1	$\theta_C - Q_{IP} + \Delta_1^E$ , where $\Delta_1^E = (\theta_{CP} - \theta_C - Q_{NE})^+$
E2	$\theta_{CP} - Q_{IP} + \Delta_2^E$ , where $\Delta_2^E = (\theta_C - \theta_{CP} - Q_{CP})^+$
E3	$\theta_{NE} + \Delta_3^E$ , where $\Delta_3^E = (\theta_{CP} - Q_{IP} - Q_{NE})^+$
Transition paths from $C$ to $D$	
B1	$1 - \theta_C - Q_0 + \Delta_1^B$ , where $\Delta_1^B = (\theta_C - \theta_{CP} - Q_{NE})^+$
B2	$1 - \theta_{CP} - Q_0 + \Delta_2^B$ , where $\Delta_2^B = (\theta_{CP} - \theta_C - Q_{CP})^+$
B3	$1 - \theta_{NE} + \Delta_3^B$ , where $\Delta_3^B = (Q_{IP} + Q_{CP} - \theta_{CP})^+$

**Final steps.** Three final steps pin down the minimum cost path:

**Step 1:** If  $Q_{IP} + \min\{Q_{NE}, Q_{CP}\} > 2\theta_{NE} + \theta_{CP} - 1$  then  $\widehat{cost}(E3) < \widehat{cost}(B3)$ ; if  $Q_{IP} + \max\{Q_{NE}, Q_{CP}\} < 2\theta_{NE} + \theta_{CP} - 1$  then  $\widehat{cost}(E3) > \widehat{cost}(B3)$ .

First, suppose  $\theta_{CP} \leq Q_{IP} + \min\{Q_{NE}, Q_{CP}\}$ . Then  $\Delta_3^E = 0$ , and  $\Delta_3^B \geq 0$ . Thus,  $\widehat{cost}(E3) = \theta_{NE}$ , and  $\widehat{cost}(B3) = 1 - \theta_{NE} + Q_{IP} + Q_{CP} - \theta_{CP}$ . It follows that

$$\widehat{cost}(E3) \leq \widehat{cost}(B3) \Leftrightarrow Q_{IP} + Q_{CP} \geq 2\theta_{NE} + \theta_{CP} - 1.$$

Second, suppose  $\theta_{CP} \geq Q_{IP} + \max\{Q_{NE}, Q_{CP}\}$ . Then  $\Delta_3^E \geq 0$  and  $\Delta_3^B = 0$ . Thus,  $\widehat{cost}(E3) = \theta_{NE} + \theta_{CP} - Q_{IP} - Q_{NE}$ , and  $\widehat{cost}(B3) = 1 - \theta_{NE}$ . Then

$$\widehat{cost}(E3) \leq \widehat{cost}(B3) \Leftrightarrow Q_{IP} + Q_{NE} \geq 2\theta_{NE} + \theta_{CP} - 1.$$

Taking together, we have the claimed properties.

In the remaining two steps, we write  $r_E \equiv \min\{\widehat{\text{cost}}(E1), \widehat{\text{cost}}(E2)\}$  and  $r_B \equiv \min\{\widehat{\text{cost}}(B1), \widehat{\text{cost}}(B2)\}$ .

**Step 2:** In the case of  $\min\{Q_{NE}, Q_{CP}\} \geq |\theta_C - \theta_{CP}|$ , we have  $r_E \leq r_B$  if and only if  $2Q_{IP} + Q_{NE} + Q_{CP} \geq \theta_C + \theta_{CP}$ .

In this case,  $\Delta_1^E = \Delta_2^E = \Delta_1^B = \Delta_2^B = 0$ . Hence,  $r_E = \min\{\theta_C, \theta_{CP}\} - Q_{IP}$ , and  $r_B = 1 - \max\{\theta_C, \theta_{CP}\} - Q_0$ . It follows that  $r_E \leq r_B \Leftrightarrow 2Q_{IP} + Q_{NE} + Q_{CP} \geq \theta_C + \theta_{CP}$ .

**Step 3:** In the case of  $\max\{\theta_C, \theta_{CP}\} \leq |\theta_C - \theta_{CP}|$ , if  $Q_{IP} + \min\{Q_{NE}, Q_{CP}\} > \frac{1}{2}(\theta_C + \theta_{CP})$ , then  $r_E < r_B$ ; if  $Q_{IP} + \max\{Q_{NE}, Q_{CP}\} < \frac{1}{2}(\theta_C + \theta_{CP})$ , then  $r_E > r_B$ .

In this case, first, suppose  $\theta_{CP} \geq \theta_C$ . Then  $\Delta_1^E \geq 0$ ,  $\Delta_2^E = 0$ ,  $\Delta_1^B = 0$ , and  $\Delta_2^B \geq 0$ . Thus  $r_E = \theta_{CP} - Q_{IP} - Q_{NE}$  and  $r_B = 1 - \theta_C - Q_0 - Q_{CP}$ . Hence,

$$(8) \quad r_E \leq r_B \Leftrightarrow 2Q_{IP} + 2Q_{NE} \geq \theta_C + \theta_{CP}.$$

Second, suppose  $\theta_{CP} < \theta_C$ . Then  $\Delta_1^E = 0$ ,  $\Delta_2^E \geq 0$ ,  $\Delta_1^B \geq 0$ , and  $\Delta_2^B = 0$ . Hence  $r_E = \theta_C - Q_{IP} - Q_{CP}$  and  $r_B = 1 - \theta_C - Q_0 - Q_{NE}$ . Therefore,

$$(9) \quad r_E \leq r_B \Leftrightarrow 2Q_{IP} + 2Q_{CP} \geq \theta_C + \theta_{CP}.$$

Collecting (8) and (9), we establish the claim.

To complete the proof, take together Steps 1 to 3. Then we know that if  $Q_{IP} + \min\{Q_{NE}, Q_{CP}\} > 2\theta_{NE} + \theta_{CP} - 1$  and  $Q_{IP} + \min\{Q_{NE}, Q_{CP}\} > \frac{1}{2}(\theta_C + \theta_{CP})$ , then  $\widehat{\text{cost}}(E3) < \widehat{\text{cost}}(B3)$  and  $r_E < r_B$ , so that by Young's theorem,  $C$  is the generically unique long-run equilibrium. Conversely, if  $Q_{IP} + \max\{Q_{NE}, Q_{CP}\} > 2\theta_{NE} + \theta_{CP} - 1$  and  $Q_{IP} + \max\{Q_{NE}, Q_{CP}\} < \frac{1}{2}(\theta_C + \theta_{CP})$ , then  $\widehat{\text{cost}}(B3) < \widehat{\text{cost}}(E3)$  and  $r_B < r_E$ , so that  $D$  is the generically unique long-run equilibrium.

*Q.E.D.*

**Proof of Proposition 3.** From Table S4, we know the relative costs of transitions between  $C$  and  $D$ . Let  $\theta_C = \theta_{NE} = \theta_{CP} = 0.5$ . Then  $\Delta_1^E = \Delta_2^E = \Delta_1^B = \Delta_2^B = 0$ , and

$$r_E \equiv \min\{\widehat{\text{cost}}(E1), \widehat{\text{cost}}(E2)\} = \frac{1}{2} - Q_{IP}$$

$$r_B \equiv \min\{\widehat{\text{cost}}(B1), \widehat{\text{cost}}(B2)\} = \frac{1}{2} - Q_0.$$

Therefore,  $r_E \leq r_B$  if and only if  $Q_{IP} \geq Q_0$ , which is equivalent to  $2Q_{IP} + Q_{NE} + Q_{CP} \geq 1$ .

Observe that  $r_E, r_B \leq \frac{1}{2}$ , but  $\widehat{cost}(E3), \widehat{cost}(B3) \geq \frac{1}{2}$ . Hence,  $E3$  and  $B3$  are never the paths with strictly minimum costs. Therefore, by Young's theorem,  $Q_{IP} > Q_0$ , implying  $r_E < r_B$ , is both necessary and sufficient for  $C$  to be the generically unique long-run equilibrium. And  $Q_{IP} < Q_0$ , implying  $r_E > r_B$ , is necessary and sufficient for  $D$  to be the generically unique long-run equilibrium.

*Q.E.D.*

## **Experimental Procedures and Materials**

Participants were recruited from Amazon Mechanical Turk (MTurk), which has been shown to provide good quality data in various settings (13–15), social dilemma games with punishment (16). After reading instructions, participants were placed in a 'lobby' until another participant arrived. Once two participants were in the lobby, they were matched and directed to the first decision stage of the experiment. In case no match could be made within 5 minutes, participants could choose to leave and receive a fixed bonus payment of \$1.00, or to wait for another 2 minutes for a possible matching partner (as in (16)). Participants were informed that from the point of reaching the lobby onwards, they did not have to make any further decisions.

Below we show on-screen instructions as displayed to participants. We start with the CC treatment in which participants could condition punishment of their interaction partner on descriptive norms of cooperation. Then we show the CP treatment, in which participants could condition punishment of their interaction partner on descriptive norms of punishment. The experiment was programmed in LIONESS Lab (17). Participants could not navigate the experimental pages at will. Each time they pressed a button, the browser history was automatically overwritten.

## 1. Instructions for the conditional cooperation (CC) treatment

### Welcome

This HIT is different from HITs that you might be used to completing via MTurk.

You will be participating in an interactive task with another MTurker.

As you are completing this task at the same time, it is important that you complete this HIT *without interruptions*.

Including the time for reading these instructions, the HIT will take about 15 minutes to complete.

Do not close this window or leave the HIT's web pages in any other way during the HIT.

If you close your browser or leave the HIT, you will NOT be able to re-enter the HIT and we will NOT be able to pay you!

You can earn **Points** in this HIT. You will have **25 Points to start with**.

Depending on your choices and the choices of the other MTurker in the task, you may earn additional Points.

At the end of the task, your Points will be converted into real money according to the exchange rate:

**20 Points = 1 Dollar.**

You will receive a code to collect your payment via MTurk upon completion.

I understand



# Instructions

You and **another MTurker** will form a pair and participate in this task at the same time.

We will refer to the MTurker in your pair as **your partner**.

You and your partner have received these same instructions. You two will play a game.

The game has two stages: **Stage 1** and **Stage 2**.

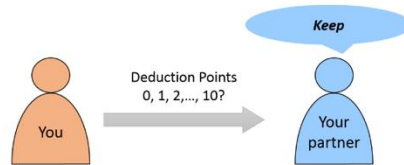
Continue

## Stage 2

(Page 2/3)

Once you and your partner have each submitted your choice for Stage 1, you proceed to Stage 2.

If your partner chose **Keep** in Stage 1, you can choose to assign **up to 10 Deduction Points** to your partner.

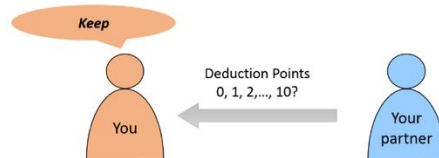


For each Deduction Point you assign, **1 Point** will be deducted from **your earnings**, and **3 Points** will be deducted from **your partner's earnings**.



Similarly, if you chose **Keep** in Stage 1, your partner can choose to assign up to 10 Deduction Points to you.

For each Deduction Point your partner chooses to assign, **1 Point** will be deducted from your partner's earnings and **3 Points** will be deducted from your earnings.



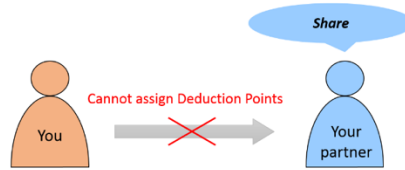
Previous

Continue

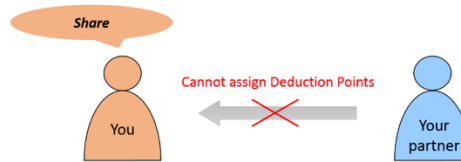
## Stage 2

(Page 3/3)

If your partner chose **Share** in Stage 1, you **cannot** assign any Deduction Points to your partner.



Similarly, if you chose **Share** in Stage 1, your partner **cannot** assign Deduction Points to you.



Previous

Continue

## Here is an example:

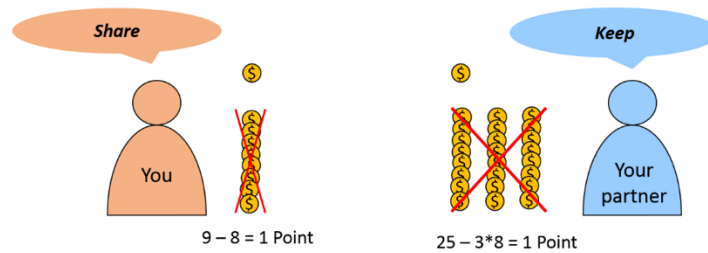
Suppose that, in Stage 1, you choose **Share** and your partner chooses **Keep**.

Then at the end of Stage 1, **you** will have **9 Points** and **your partner** will have **25 Points**.

In Stage 2, suppose that you assign **8 Deduction Points** to your partner. Then at the end of the game:

Your earnings will be **1 Point** (9 points from Stage 1 *minus* 8 points deducted from Stage 2).

Your partner's earnings will be **1 Point** (25 points from Stage 1 *minus* 3 x 8 points deducted from Stage 2).



Previous

Continue

## Summary

In Stage 1, you and your partner choose between **Share** and **Keep**.

If you choose:	and your partner chooses:	then you get:	and your partner gets:
<b>Keep</b>	<b>Keep</b>	16	16
<b>Share</b>	<b>Share</b>	18	18
<b>Keep</b>	<b>Share</b>	25	9
<b>Share</b>	<b>Keep</b>	9	25

If your partner chooses **Keep**, you can assign up to 10 Deduction Points to your partner in Stage 2. For each Deduction Point you assign, 1 Point will be deducted from your earnings, and 3 Points will be deducted from your partner's earnings.

If your partner chooses **Share** in Stage 1, you CANNOT assign Deduction Points to your partner in Stage 2.

Your partner faces the same choices as you do.

## Questions to check whether you understand the task

1. Suppose that in Stage 1, you choose **Share** and your partner chooses **Keep**.

(a) How many Points would you earn from Stage 1?

(b) How many Points would your partner earn from Stage 1?

2. Subsequently in Stage 2, you assign 5 Deduction Points to your partner.

(a) How many Points would you have at the end of the game?

(b) How many Points would your partner have at the end of the game?

Previous

Continue

## Stage 1

You answered the questions correctly.

Now click the button below to start Stage 1.

Continue

## Your choice for Stage 1



Remember:

If you choose:	and your partner chooses:	then you get:	and your partner gets:
<i>Keep</i>	<i>Keep</i>	16	16
<i>Share</i>	<i>Share</i>	18	18
<i>Keep</i>	<i>Share</i>	25	9
<i>Share</i>	<i>Keep</i>	9	25

Now please make your choice:



After you enter your decision, press "Submit" to proceed to Stage 2.  
Once you press "Submit", you **cannot** go back and change your choice for Stage 1.

## Stage 2

(Page 1/3)

Your choice for Stage 1 has been recorded. Stage 2 begins now.

Before we inform you about your partner's choice, imagine that

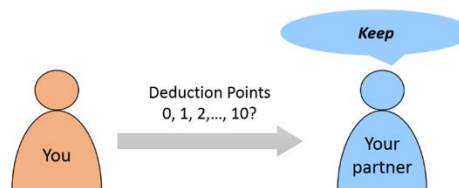
**your partner has chosen *Keep*.**

Hence:

You can assign up to 10 Deduction Points to your partner.

For each Deduction Point you assign, 1 Point will be deducted from your earnings, and 3 Points will be deducted from your partner's earnings.

If your partner indeed chose ***Keep***, the Deduction points you assign will be implemented.  
If your partner in fact chose ***Share***, the Deduction points you assign will NOT be implemented.



## Stage 2

(Page 2/3)

Over 200 MTurkers from the USA have participated in this study before.  
They were also paired and made choices between **Share** and **Keep**.

Before you make your decision, we will **randomly select fifty MTurkers (NOT your partner)** from the ones who participated in this study before.

We will tell you how many of the randomly selected **previous MTurkers (NOT your partner)** chose **Share**.



Previous

Continue

## Stage 2

(Page 3/3)

We distinguish between **11 possible situations** regarding how many of the selected **previous MTurkers** (NOT your partner) chose **Share**.  
The eleven possible situations are:

Less than 5 percent of them chose **Share**;  
Between 5 and 15 percent of them chose **Share**;  
Between 15 and 25 percent of them chose **Share**;  
Between 25 and 35 percent of them chose **Share**;  
Between 35 and 45 percent of them chose **Share**;  
Between 45 and 55 percent of them chose **Share**;  
Between 55 and 65 percent of them chose **Share**;  
Between 65 and 75 percent of them chose **Share**;  
Between 75 and 85 percent of them chose **Share**;  
Between 85 and 95 percent of them chose **Share**;  
More than 95 percent of them chose **Share**.

Before we inform you about how many of the randomly selected previous MTurkers (NOT your partner) chose **Share**, we ask you to consider EACH of the situations above.

You need to indicate, for **EACH** of the situations above, how many Deduction Points you assign to your partner if your partner chose **Keep**.

After you submit your decisions for all eleven situations, we will let you know which situation actually occurred.

The **Deduction points** you assign in the **actual situation** will be used to calculate the earnings of you and your partner. Since you do not know yet which situation is the actual one when you make your decisions, this means that you need to consider each of the situations above seriously.

**Note:**

**Your partner is NOT one of these previous MTurkers.**

**You CANNOT assign Deduction Points to any of these previous MTurkers, and the previous MTurkers CANNOT assign Deduction Points to you or your partner.**

**The Deduction Points you assign only affect the earnings of you and your partner.**

Previous

Next

## Questions to check whether you understand the task

1. Is it possible for you to assign Deduction Points to the randomly selected MTurkers (NOT your partner)?

Yes, this is possible.

No, this is NOT possible.

2. Is it possible for the randomly selected MTurkers to assign Deduction Points to your partner?

Yes, this is possible.

No, this is NOT possible.

3. Suppose that, in Stage 1, your partner chose *Keep*. Among the fifty randomly selected MTurkers (NOT your partner), between 75 and 85 percent of them chose *Share*. In this situation, you assign 3 Deduction Points to your partner if your partner chose *Keep*.

(a) How many Points will be deducted from your earnings?

(b) How many Points will be deducted from your partner's earnings?

(c) How many Points will be deducted from the previous MTurkers' earnings because of the Deduction Points you assign?

Previous

Continue

## Stage 2

You answered the questions correctly.

Now click the button below to start Stage 2.

Continue

## Your choice for Stage 2

(Page 1/2)

Now consider each of the eleven situations and make your decisions.

"The others" in the descriptions below are referred to the **fifty randomly selected MTurkers** (NOT your partner).

**Less than 5 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 5 and 15 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 15 and 25 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 25 and 35 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 35 and 45 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

Next page



## Your choice for Stage 2

(Page 2/2)

Now consider each of the eleven situations and make your decisions.

"The others" in the descriptions are referred to **the fifty randomly selected MTurkers** (NOT your partner).

**Between 45 and 55 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 55 and 65 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 65 and 75 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 75 and 85 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**Between 85 and 95 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

**More than 95 percent of the others** (NOT your partner) chose **Share**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

To go back and change your decisions for previous situations, press "Previous page" below.

To submit your final decision, press "Submit".

Previous page

Submit

### Questionnaire

You have now completed the decision making part of this HIT. Please fill out this brief questionnaire.

Once your partner is ready, we will calculate you and your partner's earnings and display them on your screen.

You will then receive a code to collect your earnings on MTurk.

1. What is your age?

2. What is your gender?

3. Could you briefly describe your reasoning for choosing to either *Share* or *Keep*?

remaining characters

4. Could you briefly describe the reasoning you used when allocating Deduction Points? In particular, why did or didn't you make your choices dependent on how many previous MTurkers chose *Share*?

remaining characters

**2. Stage 2 instructions for the conditional punishment (CP) treatment**

### Stage 2

(Page 1/3)

Your choice for Stage 1 has been recorded. Stage 2 begins now.

Before we inform you about your partner's choice, imagine that

**your partner has chosen *Keep*.**

Hence:

You can assign up to 10 Deduction Points to your partner.  
For each Deduction Point you assign, 1 Point will be deducted from your earnings, and 3 Points will be deducted from your partner's earnings.

If your partner indeed chose ***Keep***, the Deduction points you assign will be implemented.  
If your partner in fact chose ***Share***, the Deduction points you assign will NOT be implemented.

### Stage 2

(Page 2/3)

Over 200 MTurkers from the USA have participated in this study before.

They were also paired and could also assign 0,1,2,..., or 10 Deduction Points to their partners when their partners chose ***Keep***.

Before you make your decision, we will **randomly select fifty MTurkers (NOT your partner)** from the ones who participated in this study before.  
We will tell you the **average Deduction Points** that the randomly selected **previous MTurkers (NOT your partner)** assigned to their partners when their partners chose ***Keep***.

Previous

Continue

## Stage 2

(Page 3/3)

Since each player can assign up to 10 Deduction Points, there are **eleven possible situations** regarding the **average Deduction Points** that the selected **previous MTurkers** (NOT your partner) assigned to their partners. Rounding to the nearest integer, the eleven possible situations are:

- On average, they assigned 0 Deduction Points;
- On average, they assigned 1 Deduction Point;
- On average, they assigned 2 Deduction Points;
- On average, they assigned 3 Deduction Points;
- On average, they assigned 4 Deduction Points;
- On average, they assigned 5 Deduction Points;
- On average, they assigned 6 Deduction Points;
- On average, they assigned 7 Deduction Points;
- On average, they assigned 8 Deduction Points;
- On average, they assigned 9 Deduction Points;
- On average, they assigned 10 Deduction Points.

Before we inform you about the average Deduction Points that the randomly selected MTurkers actually assigned, we ask you to consider EACH of the situations above.

You need to indicate, for **EACH** of the situations above, how many Deduction Points you assign to your partner if your partner chose **Keep**.

After you submit your decisions for all eleven situations, we will let you know which situation actually occurred.

The **Deduction points** you assign in the **actual situation** will be used to calculate the earnings of you and your partner.

Since you do not know yet which situation is the actual one when you make your decisions, this means that you need to consider each of the situations above seriously.

**Note:**

**Your partner is NOT one of these previous MTurkers.**

**You CANNOT assign Deduction Points to any of these previous MTurkers, and the previous MTurkers CANNOT assign Deduction Points to you or your partner.**

**The Deduction Points you assign only affect the earnings of you and your partner.**

Previous

Next

## Questions to check whether you understand the task

1. Is it possible for you to assign Deduction Points to the randomly selected MTurkers (NOT your partner)?

Yes, this is possible.

No, this is NOT possible.

2. Is it possible for the randomly selected MTurkers to assign Deduction Points to your partner?

No, this is NOT possible.

Yes, this is possible.

3. Suppose that, in Stage 1, your partner chose **Keep**. The fifty randomly selected MTurkers (NOT your partner) on average assigned 5 Deduction Points to their partners when their partners chose **Keep**. In this situation, you assign 3 Deduction Points to your partner when your partner chooses **Keep**.

(a) How many Points will be deducted from your earnings?

(b) How many Points will be deducted from your partner's earnings?

(c) How many Points will be deducted from the previous MTurkers' earnings because of the Deduction Points you assign?

Previous

Continue

## Stage 2

You answered the questions correctly.

Now click the button below to start Stage 2.

Continue

## Your choice for Stage 2

(Page 1/2)

Now consider each of the eleven situations and make your decisions.

"The others" in the descriptions are referred to **the fifty randomly selected MTurkers** (NOT your partner).

On average, **the others assigned 0 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 1 Deduction Point** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 2 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 3 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 4 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

Next page

## Your choice for Stage 2

(Page 2/2)

Now consider each of the eleven situations and make your decisions.

"The others" in the descriptions are referred to **the fifty randomly selected MTurkers** (NOT your partner).

On average, **the others assigned 5 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 6 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 7 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 8 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 9 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

On average, **the others assigned 10 Deduction Points** when their partners chose **Keep**. Enter the Deduction Points you assign to your partner below if your partner chose **Keep**.

To go back and change your decisions for previous situations, press "Previous page" below.

To submit your final decisions, press "Submit".

Previous page

Submit

## References

1. M. Kandori, G. J. Mailath, R. Rob, Learning, mutation, and long run equilibria in games. *Econometrica* **61**, 29–56 (1993).
2. H. P. Young, The evolution of conventions. *Econometrica* **61**, 57–84 (1993).
3. H. P. Young, *Individual strategy and social structure: an evolutionary theory of institutions* (Princeton University Press, 2001).
4. B. Lindström, S. Jangard, I. Selbing, A. Olsson, The role of a “common is moral” heuristic in the stability and change of moral norms. *J. Exp. Psychol. Gen.* **147**, 228–242 (2018).
5. H. P. Young, Conventional contracts. *Rev. Econ. Stud.* **65**, 773–792 (1998).
6. H. P. Young, M. A. Burke, Competition and custom in economic contracts: a case study of illinois agriculture. *Am. Econ. Rev.* **91**, 559–573 (2001).
7. I. Eshel, L. Samuelson, A. Shaked, Altruists, egoists, and hooligans in a local interaction model. *Am. Econ. Rev.* **88**, 157–179 (1998).
8. M. O. Jackson, A. Watts, The evolution of social and economic networks. *J. Econ. Theory* **106**, 265–295 (2002).
9. S. Goyal, F. Vega-Redondo, Network formation and social coordination. *Games Econ. Behav.* **50**, 178–207 (2005).
10. H. P. Young, Innovation diffusion in heterogeneous populations: contagion, social influence, and social learning. *Am. Econ. Rev.* **99**, 1899–1924 (2009).
11. H. P. Young, The dynamics of social innovation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 21285–91 (2011).
12. H. P. Young, D. Foster, Cooperation in the long-run. *Games Econ. Behav.* **3**, 145–156 (1991).
13. G. Paolacci, J. Chandler, P. G. Ipeirotis, Running experiments on amazon mechanical turk. *Judgm. Decis. Mak.* **5**, 411–419 (2010).
14. T. S. Behrend, D. J. Sharek, A. W. Meade, E. N. Wiebe, The viability of crowdsourcing for survey research. *Behav. Res. Methods* **43**, 800 (2011).
15. R. A. Klein, *et al.*, Investigating variation in replicability. *Soc. Psychol.* **45**, 142–152 (2014).
16. A. A. Arechar, S. Gaechter, L. Molleman, Conducting interactive experiments online. *Exp. Econ.* **21**, 99–131 (2018).
17. M. Giamattei, L. Molleman, K. Seyed Yahosseini, S. Gachter, Lioness Lab -- a free web-based platform for conducting interactive experiments online. *SSRN Electron. J.* (2019).