



# Learning to cooperate in the shadow of the law

Roberto Galbiati<sup>1</sup> · Emeric Henry<sup>1</sup> · Nicolas Jacquemet<sup>2</sup> 

Received: 13 May 2023 / Revised: 7 November 2023 / Accepted: 27 December 2023 /

Published online: 19 March 2024

© The Author(s), under exclusive licence to Economic Science Association 2024

## Abstract

Formal enforcement punishing defectors can sustain cooperation by changing incentives. In this paper, we introduce a second effect of enforcement: it can also affect the capacity to learn about the group's cooperativeness. Indeed, in contexts with strong enforcement, it is difficult to tell apart those who cooperate because of the threat of fines from those who are intrinsically cooperative types. Whenever a group is intrinsically cooperative, enforcement will thus have a negative dynamic effect on cooperation because it slows down learning about prevalent values in the group that would occur under a weaker enforcement. We provide theoretical and experimental evidence in support of this mechanism. Using a lab experiment with independent interactions and random rematching, we observe that, in early interactions, having faced an environment with fines in the past decreases current cooperation. We further show that this results from the interaction between enforcement and learning: the effect of having met cooperative partners has a stronger effect on current cooperation when this happened in an environment with no enforcement. Replacing one signal of deviation without fine by a signal of cooperation without fine in a player's history increases current cooperation by 10%; while replacing it by a signal of cooperation with fine increases current cooperation by only 5%.

**Keywords** Enforcement · Social values · Cooperation · Learning · Spillovers · Repeated games · Experiments

---

✉ Nicolas Jacquemet  
nicolas.jacquemet@univ-paris1.fr

Roberto Galbiati  
roberto.galbiati@sciences-po.fr

Emeric Henry  
emeric.henry@sciences-po.fr

<sup>1</sup> Department of Economics, CNRS-Sciences Po and CEPR, 28 rue des Saints-Pères, 75007 Paris, France

<sup>2</sup> Paris School of Economics and University Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne, 106 Bd de l'hôpital, 75013 Paris, France

**JEL Classification** C91 · C73 · D02 · K49 · P16 · Z1

## 1 Introduction

Why does the level of cooperation vary across societies and organizations? A natural answer is that rules, and the strength of their enforcement, might differ. One expects high levels of cooperation where formal enforcement punishes defectors, for instance by the means of high fines. When strong formal enforcement is absent, cooperation can be sustained if cooperative values are prevalent enough in the group. For this second driver of cooperation, learning about the cooperativeness of the group becomes essential.

In this paper we study the interaction between these two drivers of cooperation. We explore a simple intuition: formal enforcement does not only affect the individual decisions to cooperate, it also impacts on the capacity to learn about the group's cooperativeness. In contexts with high fines for those who do not cooperate, it is difficult to tell apart people who cooperate because of the threat of fines from those who are intrinsically cooperative types. The shadow of the law thus affects learning about the group and hence future cooperation. Consider for instance the situation of a taxpayer who needs to decide whether to truthfully report income or evade taxes. Taxpayers are disciplined by fines (Bérgolo et al., 2021) but at the same time high fines elicit tax compliance behavior that do not reveal the general intrinsic honesty of the population, which matters for future decisions.<sup>1</sup> In this paper, we provide evidence, both theoretical and experimental, on this interaction between fines and learning.

We rely on a lab experiment where participants play a series of indefinitely repeated prisoner's dilemma. At the beginning of each game, it is randomly determined whether a formal enforcement in the form of a fine will be imposed in all rounds of the game when a participant chooses to deviate rather than cooperate.<sup>2</sup> At the end of the game, each participant is re-matched with a new one and it is randomly determined whether the new game is played with fines. The design ensures that each participant (i) has a different history of exposure to fines and of past behavior of partners, and that this history both (ii) does not depend on self-selection into particular environments, and (iii) is independent from the current environment faced by each individual. Each experimental subject thus faces a different history of past cooperation observed in different enforcement environments.

Our first main result shows that, in early games, past enforcement negatively affects current cooperation. We argue that the interaction between cooperation-enforcing institutions and learning can potentially explain such a pattern. Consider

<sup>1</sup> Learning about the honesty of others matters not only if tax compliance relies on social interactions (Fortin et al., 2007) but more generally in all daily decisions that are not legally enforceable.

<sup>2</sup> In our setting when fines are present they are exogenously enforced. In this sense our setting differs from other studies like Acemoglu and Jackson (2017) or Zasu (2007) studying complementarities in the enforcement of cooperation between laws and social norms.

the case where the population is fairly non cooperative (i.e., less cooperative than expected). In this case, experiencing a fine can speed up learning the bad news, since observing deviation in an environment with fines is a strong indicator that the partner is non-cooperative. Conversely, learning will be slow within a cooperative group in an environment with fines. In such a context, it will not be possible to learn whether cooperation is driven by fines or by partners' willingness to cooperate. This interaction between fines and learning is the key driving force of our model, whose findings are confirmed in the data.

The finding that enforcement in the past decreases current cooperation in early games contrasts with what would be expected based on the literature documenting the behavioral spillovers of enforcement institutions—i.e., the fact that enforcement institutions faced either in the past (e.g., Peysakhovich & Rand 2016; Duffy & Fehr, 2018; Galizzi & Whitmarsh, 2019, for a survey) or in other games (Engl et al., 2021), affect the current willingness to cooperate through behavioral channels.<sup>3</sup> Galbiati et al. (2018) show that enforcement institutions foster future cooperation through indirect spillovers—fines increase the likelihood that the current partner cooperates, which in turn induces more cooperation in the future through indirect reciprocity (Nowak & Roch, 2007).<sup>4</sup> This previous study uses the same experiment as in the current paper, but focuses only on games occurring late in the experiment under the assumption that learning has converged. In this paper, we focus on the interaction between fines and rational learning. In early games, there is uncertainty about intrinsic values in the group (i.e., about whether the average partner is of a cooperative or a non-cooperative type). In such circumstances, our theoretical model shows that partners' behavior in previous games brings information about how cooperative the group is and thus affects current behavior. From an empirical point of view, observing learning becomes challenging due to the simultaneous effect of behavioral spillovers. We propose two identification strategies to identify learning separately from spillovers.

The main identification strategy exploits the idea that behavioral spillovers do not last (as shown in Galbiati et al. 2018) while learning is cumulative: whether cooperation was observed one or two periods ago does not matter for learning, as the information it delivers remains the same, but will matter for spillovers if they decay over time. Thanks to the assumption that spillovers are short-lived, we can disentangle the two by regressing current cooperation levels on variables that are history dependent (spillovers) and history independent (learning). Our results show that replacing in the history one signal of deviation without fine by a signal of cooperation without fine, increases current cooperation by 10%; while replacing it

<sup>3</sup> Another strand of recent literature highlights the existence of possible unexpected effects of naive policy interventions in the presence of social norms by focusing on how the incentives introduced by policies interact with the endogenous emergence of these norms (see for instance Dutta et al. 2021).

<sup>4</sup> This paper relies on the same data as Galbiati et al. (2018), who focus on games occurring late in the experiment, when learning about norms prevalent in the group has converged. Herein, we rather analyze the entire dataset, including early games.

by a signal of cooperation with fine increases current cooperation only by 5%. This is coherent with rational learning dynamics.

As a robustness check, we also provide a structural analysis (provided in the “[Alternative identification strategy](#)” in Appendix 2) which identifies learning in early games conditional on behavioral spillover parameters, under the assumption that learning has converged in late games. We first generalize our theoretical model to the case in which individual values evolve over time as a result of past experience. This extended model allows us to express the probability of cooperation as a function of both learning and spillover parameters that we can estimate with our experimental data. The results confirm that lab participants behave in accordance with the learning dynamics described by the model: cooperation by the partner in the previous game, if it was played with a fine, has a smaller positive effect than if this cooperation took place in a game without fines. This learning effect may imply that current fines negatively affect future cooperation. If the group is non cooperative, fines may speed up learning, since more individuals will be observed deviating in a coercive environment.

By documenting the dynamic interaction between enforcement and learning about group values, our study provides several contributions to the existing literature.<sup>5</sup> Acemoglu and Jackson (2015), study how norms of cooperation can emerge in an environment where current generations learn from observed cooperation in past generations. They do not consider, however, the effect of institutional variations in the past and their interactions with learning. A recent literature shows that formal rules (Sliwka, 2007; Van Der Weele, 2009; Deffains & Fluet, 2020) or principals’ interventions (Friebel & Schnedler, 2011; Galbiati et al., 2013) can convey information on their own about either the distribution of preferences or values in a group, or the type of the principal (Falk & Kosfeld, 2006; Bowles, 2008), thus leading to ambiguous contemporaneous effects of sanctions. The main focus of our study is rather on the information delivered by agents actions’ depending on the enforcing institutions, as in Benabou and Tirole (2011). In their setup, individuals care about their social-image, which is based on inferences made by other group members about their types. Institutions shape equilibrium behaviors and thus the inference induced by different actions. In our study, we rather focus on the informativeness of actions about cooperative types under different enforcement environments. We show that differential learning due to enforcement institutions lead to countervailing spillover effects on future cooperation as long as learning is in progress: strong enforcement weakens the signal of cooperativeness sent by cooperative types, and slows down future cooperation.

Our results are also informative on the impact of enforcement on learning, and could thus be very relevant for the performance of young organizations where members have not yet learned about the cooperativeness of the others. Such an

<sup>5</sup> In a related work, Dal Bó and Dal Bó (2014) show that explicit information about moral values affect cooperation in a standard voluntary contribution game. In their setting however the information is provided by the experimenter and does not allow for dynamic learning about the distribution of prevalent types in the lab.

effect of enforcement on the extent to which cooperative decisions reveal intrinsic cooperativeness also echoes Ali and Bénabou (2020)'s social image model. In their framework, a benevolent principal needs to learn about the values prevalent in a group of agents who care about their social-image. Values evolve over time and the principal needs to decide on the optimal level of transparency to learn the prevalent values. A key result of the model is that some privacy is needed to achieve this goal. A high level of transparency leads to pro-social decisions that are mainly driven by the social-image motives and not representative of individual values. Our results provide empirical support to this key insight on the interaction between learning and enforcement. Last, we also contribute to the experimental literature on repeated games. Our results show how learning generates interdependence across games even when subjects are randomly re-matched. This suggests that independence across games is not granted even in settings with random matching and incentive compatible choices within each game. This point is coherent with previous findings showing that learning about the properties of the group matters for subjects' choices (Dal Bó & Fréchette, 2011; Gill & Rosokha, 2020)<sup>6</sup> and with the theoretical results of Azrieli et al. (2018), showing how uncertainty about the population can generate failure of incentive compatibility of the random incentive system.

## 2 Descriptive experimental evidence

### 2.1 Experimental design

The design of the baseline experiment closely follows the experimental literature on infinitely repeated games and in particular Dal Bó and Fréchette (2011). Subjects in the experiment play infinitely repeated games implemented through a random continuation rule. At the end of each round, the computer randomly determines whether or not another round is to be played in the current repeated game ("match"). This probability of continuation is fixed at  $\delta = 0.75$  and is independent of any choices players make during the match. Participants therefore play a series of matches of random length, with expected length of 4 rounds. At the end of each match, players are randomly and anonymously reassigned to a new partner to play the next match. This corresponds to a quasi-stranger design since there is a non-zero probability of being matched more than once with the same partner during the experiment. The experiment terminates once the match being played at the 15th minute ends.

The stage-game in all interactions is a prisoner's dilemma. Enforcing institutions are randomly assigned: at the beginning of each match, the computer randomly determines whether the match is played with a fine imposed in case of defection (payoffs in Table 1b) or without (Table 1a); the two events occur with equal probability. The result from this draw applies to both players of the current match, and

<sup>6</sup> More precisely Dal Bó and Fréchette (2011) document that the behavior of the partner in the previous match affects the subjects' behavior.

**Table 1** Stage-game payoff matrices

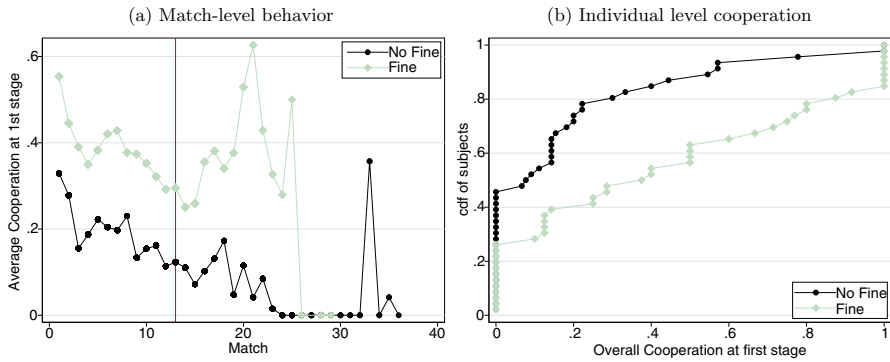
(a) Baseline game		
	C	D
C	40 ; 40	12 ; 60
D	60 ; 12	35 ; 35
(b) With fine		
	C	D
C	40 ; 40	12 ; 60-F
D	60-F ; 12	35-F ; 35-F

to all its rounds. The fine when imposed is set at  $F = 10$  so that the resulting stage-game payoff matrix is isomorphic to Dal Bó and Fréchette (2011)  $\{\delta = 3/4; R = 40\}$  treatment, in which cooperation is a sub-game perfect and risk dominant action. When matched with a new partner, subjects are not provided with any information about the partner's history. Players however receive full feedback at the end of each round about the actions taken within the current match.

## 2.2 Experimental data

Our data come from three sessions of the experiment conducted at Ecole Polytechnique experimental laboratory. The 46 participants are both students (85% of the experiment pool) and employees of the university (15%). Individual earnings are computed as the sum of all tokens earned during the experiment, with an exchange rate equal to 100 tokens for 1 Euro. At the end of the experiment, participants are asked to answer a socio-demographic questionnaire about their gender, age, level of education, labor market status (student/worker/unemployed) as well as the Dohmen et al. (2011) self-reported measure of risk-aversion. Participants earned on average 12.1 Euros from an average of 20 matches, each featuring 3.8 rounds. This data delivers 934 game observations, 48% of which are played with no fine.

All our analysis in this paper will be based on the action chosen in the first round of each match. While this first round decision captures the effect of past history of play on individual behavior, the decisions made within the course of a match are a mix of this component and of the strategic interaction with the current partner, and would thus be noisy measures of learning. To render this restriction meaningful, and also to be consistent with the model we introduce in Sect. 3, we thus restrict the sample to player-game observations for which the first round decision summarizes the future history (“[Replication of the statistical analysis on the full sample](#)” in Appendix 3 provides a replication of our statistical analysis on the full sample). As explained in more detail in the “[Data description](#)” in Appendix 1, if subjects choose among the following repeated-game strategies, Always Defect (AD), Tit-For-tat (TFT) or Grim Trigger (GT), the first round decision is a sufficient statistic for the future sequence of play. While AD dictates defection at the first round, TFT and

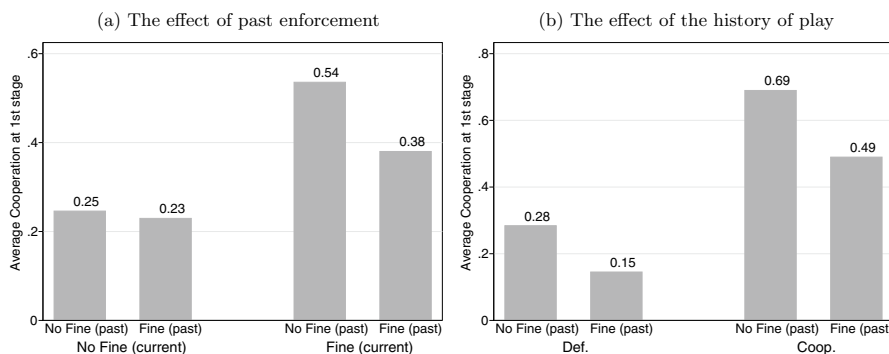


**Fig. 1** The disciplining effect of current enforcement. *Note:* Cooperation observed at first round of each match in the working sample as a function of the current fine. *Left-hand side:* evolution of the average rate of cooperation among players over the number of matches played. The vertical line identifies the point in time beyond which we no longer observe a balanced panel. *Right-hand side:* cumulative distribution of individual cooperation rate at first round of all matches played respectively with and without a fine

GT induce cooperation at the first round and are both observationally equivalent if the partner chooses within the set restricted to these three strategies and give rise to the same expected payoff. The resulting working sample is made of 785 games among which 50.3% are played with a fine. Our outcome variable of interest is the first round decision made by each player in each of these matches. Importantly, all lagged variables are computed according to actual past experience: one's own cooperation at previous match, partner's decision and whether the previous match was played with a fine are all defined according to the match played just before the current one, whether or not this previous match belongs to the working sample.

### 2.3 Learning to cooperate: descriptive evidence

Figure 1 provides an overview of the cooperation rate observed in each of the two institutional environments. The overall average cooperation rate is 32%, with a strong gap depending on whether a fine enforces cooperation: the average cooperation rate jumps from 19% in the baseline, to 46% with a fine. This is clear evidence of a strong disciplining effect of current enforcement. Figure 1a documents the time trend of cooperation over matches. The vertical line identifies the point in time beyond which we no longer observe a balanced panel—the number of matches played within the duration of the experiment is individual specific, since it depends on game lengths. Time trends beyond this point are to a large extent driven by the size of the sample. Focusing on the balanced panel, our experiment replicates in both environments the standard decrease in cooperation rates: from 15% at the initial match in the baseline, 69% with a fine, to 11% and 41% at the 13th game. The time trends are parallel between the two conditions. Note that since the history of past enforcement is both individual specific and random, it is statistically the same for the two curves for any match number.



**Fig. 2** Observed dynamics of cooperation in early games *Note:* Cooperation at first stage in the working sample in early games (see Footnote 7) according to individual history. In each figure, the data is split according to whether the previous match was played with No fine (“No fine (past)”) or with a fine (“Fine (past)”). *Left-hand side:* each subpanel refers to current enforcement; *Right-hand side* each sub-panel refers to the partner’s decision experienced at previous match

Figure 1b reorganizes the same data but at the individual level, and displays the cumulative distribution of cooperation in a given environment. We observe variations in both the intensive and the extensive margin of cooperation in the adjustment to the fine—resulting in first order stochastic dominance of the distribution of cooperation with no fine. First, regarding the extensive margin, we observe a switch in the mass probability of subjects who always choose the same first round response: 45% never cooperate with a fine, while only 26% do so with a fine, and the share of subjects who always cooperate raises from 4 to 17% when a fine is implemented. More than half the difference in mass at 0 thus moves to 1. Turning to the intensive margin, the distribution of cooperative decisions with no fine is more concentrated towards the left: 70% of individuals who switch between cooperation and defection cooperate less than 30% of the time with no fine, while it is the case of only 40% of individuals who switch from one match to the other in the fine environment.

We now turn to the main focus of the paper. To present the evidence graphically, we restrict to early games where the uncertainty about group cooperativeness is large.<sup>7</sup> Fig. 2a documents the surprising effect of fines experienced in previous matches on current cooperation. Comparing the two left-hand side bars to the right-hand side ones unambiguously demonstrates that current enforcement has a strong disciplining effect. For instance, restricting to matches where no fine was experienced in the past, the average rate of cooperation increases from 0.25 to 0.54 in environments with enforcement (bars 1 and 3). On the contrary, enforcement in the past induces a fall in current cooperation. For instance, comparing the two bars

<sup>7</sup> We distinguish early and late games by splitting the matches in three groups, in such a way that one third of the observed decisions are classified as “early”, and one third as “late”. Observed matches are accordingly defined as “early” up to the 7th, as late after the 13th, in line with the definitions used in the “Alternative identification strategy” in Appendix 2. Note that the matches we exclude from the working sample all appear in late games. As a result, Fig. 2a is not affected by the choice of the working sample.



on the right hand side, corresponding to matches where a fine is currently in place, having played the previous match with fines decreases cooperation from an average of 0.54 to 0.38 (bars 3 and 4). Such an effect of past enforcement is puzzling, since one would expect that past fines are either neutral or exert a positive effect on current cooperation through behavioral channels (e.g., Peysakhovich & Rand, 2016).

The interaction between cooperation-enforcing institutions and learning can potentially explain such a pattern. Consider the case where the news are bad (i.e., the population is less cooperative than expected; as seems to be the case according to the evolution of cooperation over time shown in Fig. 1). In this case, experiencing a fine can speed up learning the bad news, since observing deviation in an environment with fines is a strong indicator that the partner is non-cooperative. This interaction between enforcement and learning is presented in Fig. 2b, which reports the level of cooperation depending on whether cooperation (right panel) or a deviation (left panel) has been observed in the previous match in an environment with or without a fine. Comparing the two left hand side bars to the right hand side ones shows that cooperation by the partner in the previous match increases cooperation in the current match, consistent with the idea that experimental subjects learn about the willingness to cooperate of their partners thanks to their decisions. However this learning is clearly affected by the institutional environment. When the cooperative action was taken in an environment without fines, it leads to higher levels of current cooperation. For instance, comparing the two bars on the right hand side, corresponding to matches where the partner cooperated in the previous match, if that cooperation was observed in an environment with no enforcement, the average level of cooperation is 0.69 while it falls to 0.49 if the previous match was played with fines (bars 3 and 4).

Changes in cooperation according to the history of institutional exposure however combine the effect of learning as well as the direct effect of past enforcement on cooperation behavior. To clarify the link between learning and enforcement institutions, we now turn to a theoretical model that formalizes the interaction between the institutional environment and learning about group cooperativeness.

### 3 A theoretical model of cooperation dynamics

In each match (we use index  $t$  for the match number), the players simultaneously choose between actions  $C$  and  $D$  to maximize their payoff in the current match. At the end of the match, players observe the partner's decision. In the case where a match is a repeated prisoner's dilemma, as is the case in the experiment, this requires the first period action in a match to fully summarize strategies. To ease exposition, we denote  $i$  the player under consideration and  $j_t$  the partner of  $i$  in match  $t$ . Whether player  $\ell$  experience a fine in match  $t$  is tracked by the variable  $F_{\ell,t} \in \{0, 1\}$  and the action of player  $\ell$  in match  $t$  is denoted  $a_{\ell,t} \in \{C, D\}$ .

The payoff of player  $i$  from playing  $a_{it} \in \{C, D\}$  in match  $t$  is denoted  $U_{it}^a$  and is given by<sup>8</sup>:

$$\begin{cases} U_{it}^C(F_{it}, p_{it}) = V_{it}^C(F_{it}, p_{it}) + \beta_i, \\ U_{it}^D(F_{it}, p_{it}) = V_{it}^D(F_{it}, p_{it}), \end{cases}$$

where  $V_{it}^a(F_{it}, p_{it})$  is the material payoff player  $i$  expects from choosing action  $a$  in match  $t$ . This expected payoff depends in particular on the beliefs player  $i$  holds on the probability that the partner  $j_i$  cooperates,  $p_{it}$ , and on whether the current match is played with a fine,  $F_{it}$ . Note that  $p_{it}$  is in fact a function of  $F_{it}$ , since the presence of a fine affects the probability that the partner cooperates.<sup>9</sup>

The parameter  $\beta_i$  measures player  $i$ 's intrinsic values, i.e., the individual propensity to cooperate.<sup>10</sup> We suppose there is uncertainty on the set of group's values, i.e., the set of individual values  $\beta_i$ . We consider two possible states of the world. With probability  $q_0$  the state is high and  $\beta_i$  is drawn from the normal distribution  $\Phi(\mu_H, \sigma^2)$ , while with probability  $1 - q_0$ ,  $\beta_i$  is drawn from  $\Phi(\mu_L, \sigma^2)$ , with  $\mu_L < \mu_H$ . The value attached to cooperation by society is higher in the high state.

### 3.1 Benchmark model

First consider a benchmark model with no uncertainty on values ( $q = 1$ ). We now use the specific payoffs corresponding to the prisoner's dilemma to explicitly describe the impact of fines on payoffs. Denote  $\pi_{a_{it}, a_{jt}}$  the monetary payoff of  $i$  in a match where  $a_{it}$  is played against  $a_{jt}$ . Individual  $i$ , with beliefs  $p_{it}$  that her partner will cooperate, chooses action  $C$  if and only if the following condition is satisfied<sup>11</sup>:

$$\begin{aligned} p_{it} \frac{1}{1-\delta} \pi_{C,C} + (1-p_{it}) \left[ \pi_{C,D} + (\pi_{D,D} - F \times \mathbb{1}_{\{F_{it}=1\}}) \frac{\delta}{1-\delta} \right] + \beta_i \\ \geq p_{it} \left[ (\pi_{D,C} - F \times \mathbb{1}_{\{F_{it}=1\}}) + (\pi_{D,D} - F \times \mathbb{1}_{\{F_{it}=1\}}) \frac{\delta}{1-\delta} \right] \\ + (1-p_{it}) \frac{1}{1-\delta} (\pi_{D,D} - F \times \mathbb{1}_{\{F_{it}=1\}}). \end{aligned}$$

This condition can be re-expressed as

$$\beta_i \geq \beta^*(F_{it}) \equiv \Pi_1 - F \times \mathbb{1}_{\{F_{it}=1\}} + p_{it} \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \times \mathbb{1}_{\{F_{it}=1\}} + \Pi_3) \right], \quad (1)$$

<sup>8</sup> This is a specific functional form of the more general function in, e.g., Kartal and Müller (2018).

<sup>9</sup> We drop this dependency of  $p_{it}$  on  $F_{it}$  in the notation.

<sup>10</sup> In the generalized model presented in the "The dynamics of learning with behavioral spillovers" in Appendix 2, we introduce spillovers by assuming that the values  $\beta_{it}$  depend on  $t$  and in particular can be affected by past experiences.

<sup>11</sup> We explicitly use the fact that players are restricted to choosing between Grim Trigger, Tit For Tat and Always Defect.

with the parameters defined as  $\Pi_1 \equiv \pi_{D,D} - \pi_{C,D} > 0$ ,  $\Pi_2 \equiv (\pi_{D,C} - \pi_{D,D}) - (\pi_{C,C} - \pi_{C,D})$  and  $\Pi_3 \equiv \pi_{C,C} - \pi_{D,D} > 0$ .<sup>12</sup>

Condition (1) implies that the decision to cooperate follows a cutoff rule, such that an individual  $i$  cooperates if and only if she attaches a sufficiently strong value to cooperation  $\beta_i \geq \beta^*(F_{it})$ , where the cutoff  $\beta^*$  depends on whether the current match is played with a fine. Since there is no uncertainty, and thus no learning, all players share the same belief over the probability that the partner cooperates, given by  $p_{it}(F_{it}) = P[\beta_j \geq \beta^*(F_{it})] = 1 - \Phi_H[\beta^*(F_{it})]$ . The cutoff value  $\beta^*(F_{it})$  is thus defined by the indifference condition:

$$\beta^*(F_{it}) = \Pi_1 - F \times \mathbb{1}_{\{F_{it}=1\}} + [1 - \Phi_H[\beta^*(F_{it})]] \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \times \mathbb{1}_{\{F_{it}=1\}} + \Pi_3) \right]. \quad (2)$$

We show in Proposition 1 below that there always exists at least one equilibrium, and this equilibrium is of the cutoff form. There could exist multiple equilibria, but all stable equilibria share the intuitive property that individuals are more likely to cooperate in an environment with fines.

**Proposition 1** *In an environment with no uncertainty on values ( $q = 1$ ), there exists at least one equilibrium. Furthermore all equilibria are of the cutoff form, i.e., individuals cooperate if and only if  $\beta_i \geq \beta^*(F_{it})$  and, in all stable equilibria,  $\beta^*$  decreases with  $F$  and with  $\mu_H$ .*

**Proof** See “Appendix 5: Proofs”. □

The benefit of cooperation is increasing in the probability that the partner cooperates. There exist equilibria where cooperation is prevalent, which indeed makes cooperation individually attractive. On the contrary there are equilibria with low levels of cooperation which makes cooperation unattractive. These equilibria can be thought of as different norms of cooperativeness in the group, driven by complementarities in cooperation.

### 3.2 Learning in the shadow of the law

We now consider the more general formulation with uncertainty about the group's values. We denote  $q_{it}$  the belief held by player  $i$  at match  $t$  that the state is  $H$ . All group members initially share the same beliefs  $q_{i0} = q_0$ . They gradually learn about the group's values when observing the decisions of partners in previous matches and we show how fines impact learning.

First consider the initial match,  $t = 1$ . All members of the group share the same belief  $q_0$  that the state is  $H$ . The equilibrium is defined by a single cutoff value  $\beta^*(F_{it})$  as in the benchmark model,

<sup>12</sup> In the case of the experiment,  $\Pi_1 = 23$ ,  $\Pi_2 = -3$  and  $\Pi_3 = 5$ .

$$\beta^*(F_{it}) = \Pi_1 - F \times \mathbb{1}_{\{F_{it}=1\}} + p_1^*(F_{it}) \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \times \mathbb{1}_{\{F_{it}=1\}} + \Pi_3) \right].$$

The only difference with the benchmark model is that the probability that the partner cooperates takes into account the uncertainty about the group's values:

$$p_1^*(F_{it}) = q_0 [1 - \Phi_H[\beta^*(F_{it})]] + (1 - q_0) [1 - \Phi_L[\beta^*(F_{it})]].$$

We now consider how beliefs about the state of the world are updated following the initial match. The updating following this initial match provides all the intuitions for the more general updating process. The update depends on the action of the partner and whether the match was played with or without a fine. The general notation we use is  $q_{it}(F_{it-1}, a_{j_{t-1}}, q_{it-1})$ . For the update following the first match, we can drop the dependence on  $q_{it-1}$ , since all individuals initially share the same belief.

Clearly, the belief that the state is  $H$  decreases if the partner chose  $D$ , while it increases if the choice was  $C$ . The update however depends as well on whether the previous match was played with a fine or not. If the partner cooperated in presence of a fine, it is a less convincing signal that society is cooperative than if he cooperated in the absence of the fine— $q_{i2}(0, C) > q_{i2}(1, C)$ . Similarly, deviation in the presence of a fine decreases particularly strongly the belief that the state is high— $q_{i2}(1, D) < q_{i2}(0, D)$ . This is summarized in the following lemma<sup>13</sup>:

**Lemma 1** *In any stable equilibrium, beliefs following the first period actions are updated in the following way:*

$$\begin{aligned} q_{i2}(0, C) &> q_{i2}(1, C) > q_0, \\ q_{i2}(1, D) &< q_{i2}(0, D) < q_0. \end{aligned}$$

**Proof** See “Appendix 5: Proofs”. □

We show in Proposition 2 that this updating property is true in general for later matches. The beliefs on how likely it is that the partner cooperates in match  $t$ ,  $p_t^*(F_{it}, q_{it})$ , depends both on player  $i$ 's history, but also on the beliefs about the partner's history. For instance if the partner faced a lot of cooperation in previous games, she becomes more likely to cooperate. The general problem requires to keep track of the higher order beliefs. However if a stationary equilibrium exists, with the property that  $\beta^*(0, q) > \beta^*(1, q)$  for all beliefs  $q$ , then the updating property of Lemma 1 is preserved. Furthermore, in the “Appendix 5: Proofs”, we show existence of such a stationary equilibrium, under a natural restriction on higher order beliefs, i.e., if we assume that a player who had belief  $q_{it}$  in match  $t$  believes that players in the preceding match had the same beliefs  $q_{i',t-1} = q_{it}$ .

<sup>13</sup> Stability guarantees that when the current match is played with a fine, the probability of cooperation increases.

**Proposition 2** (Learning) *In an environment with spillovers and learning, if an equilibrium exists, all equilibria are of the cutoff form, i.e., individuals cooperate if and only if  $\beta_i \geq \beta^*(F_{it}, q_{it})$ . Furthermore, if in equilibrium  $\beta^*(0, q) > \beta^*(1, q)$  for all beliefs  $q$ , then the beliefs are updated in the following way following the history in the previous interaction:*

$$\begin{aligned} q_{it}(0, C, q_{it-1}) &> q_{it}(1, C, q_{it-1}) > q_{it-1}, \\ q_{it}(1, D, q_{it-1}) &< q_{it}(0, D, q_{it-1}) < q_{it-1}. \end{aligned}$$

**Proof** The “Appendix 5: Proofs”, proves the result in the more general case with spillovers.  $\square$

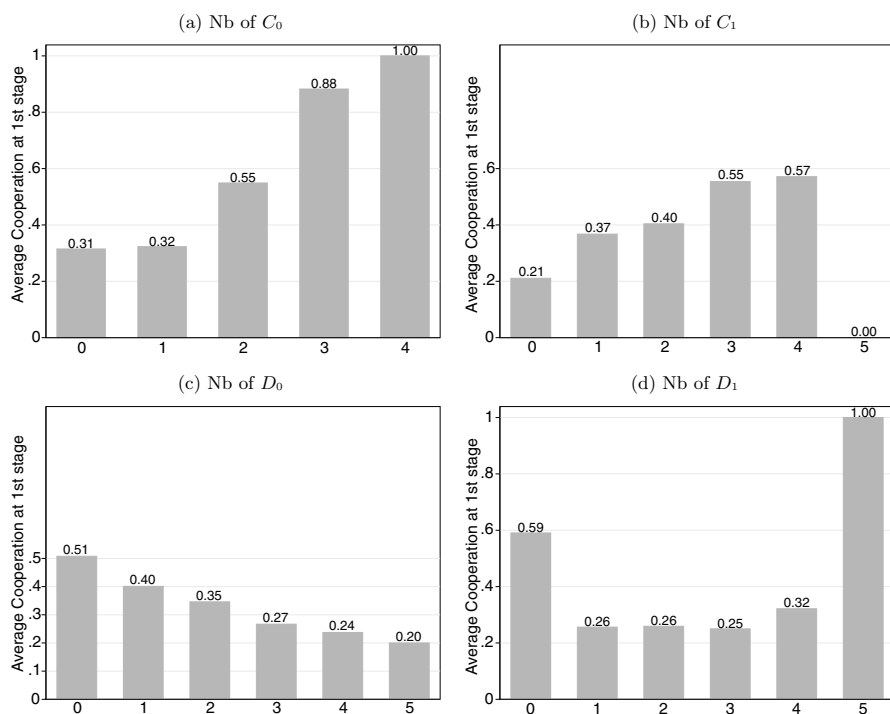
Lemma 1 and Proposition 2 show how enforcing institutions affect learning. These results imply that having fines in the previous match can potentially decrease average cooperation in the current match. If the state is low, a fine can accelerate learning if, on average, sufficiently many people deviate in the presence of a fine. This in turn decreases cooperation in the current match.

## 4 Results

We now study empirically the interaction between enforcement and learning highlighted in the model. The descriptive evidence provided in Fig. 2b (Sect. 2.3), suggests that the pattern of cooperation observed in the current match is consistent with the ranking of posterior beliefs predicted in Lemma 1 and Proposition 2,  $q_{it}(0, C) > q_{it}(1, C) > q_{it}(0, D) > q_{it}(1, D)$ .<sup>14</sup>

The identification of learning effects is however complicated by the fact that both enforcing institutions and cooperation by the partner in previous matches can also create spillovers on current cooperation. Two types of such spillovers of past enforcing institutions can be at stake: direct spillovers, according to which the fine experienced in the immediate past directly affects preferences and increases current cooperation; and indirect spillovers, according to which fines in the past increase cooperation of the previous partner, which in turn increases current cooperation. If such spillovers exist, they both interfere with the identification of learning effects. On the one hand, cooperation by the previous partner affects current cooperation both because it provides information on the cooperativeness of the group, but also because of indirect spillovers. On the other hand, a fine in the previous period similarly impacts learning as explained in the model, but also gives rise to direct spillovers. Galbiati et al. (2018) show that these spillovers are short-living: cooperation by the partner two matches ago has a much weaker effect on current cooperation than cooperation by the partner in the previous match.

<sup>14</sup>  $q_{it}(0, C)$  corresponds to the third bar in Fig. 2b,  $q_{it}(0, C)$  to the fourth,  $q_{it}(0, D)$  to the first and  $q_{it}(1, D)$  to the second.



**Fig. 3** Current cooperation as a function of history in the previous 5 games. *Note:* Each panel reports the average level of cooperation in  $t$  as a function of the number of decisions  $a_{j,t-s}$ ,  $s = 1, \dots, 5$  in the current history. The abscissa is  $\sum_{s=1}^5 C_{t-s}^0$  in panel **a**,  $\sum_{s=1}^5 C_{t-s}^1$  in panel **b**,  $\sum_{s=1}^5 D_{t-s}^0$  in panel **c** and  $\sum_{s=1}^5 D_{t-s}^1$  in panel **d**. For both  $C_0$  and  $D_1$ , observed histories only range from 0 to 4. For  $D_1$ , the observed level of cooperation when 5 of them belong to history is 0%

We use these findings to identify separately learning from spillovers. To illustrate the idea, compare two situations, with identical institutional histories: the first where the partner in the previous match cooperated while the one two matches ago did not and the second, with the opposite behavior, the partner in the previous match deviated and the one two matches ago did not. From the point of view of learning, both situations are equivalent since the information obtained is identical: one of the two previous partners did cooperate. However, in terms of spillovers, the first situation should lead to higher levels of current cooperation: if spillovers decay over time, facing cooperation two periods ago has a smaller spillover effect than cooperation one period ago.

We exploit this idea in Fig. 3, where we examine the effect of the history, in terms of fines and behavior of the partner, in the five previous matches independently of the order in which this history occurred. Figure 3a for instance displays how average cooperation is affected by the number of matches without fines where the partner cooperated (an outcome we denote  $C_0$ ). An increase in the number of  $C_0$  has a very strong effect on current cooperation, with an average rate of cooperation of 0.29 when it never occurred in the 5 previous matches

**Table 2** Statistical evidence on the interaction between enforcement and learning

	(1)		(2)		(3)	
	Coef.	Marg. eff.	Coef.	Marg. eff.	Coef.	Marg. eff.
Constant	0.040 (0.274)		0.147 (0.306)		0.179 (0.152)	
$\mathbb{1}_{\{F_t=1\}}$	1.356*** (0.270)	0.305*** (0.052)	1.361*** (0.271)	0.303*** (0.051)	1.355*** (0.268)	0.302*** (0.047)
$\bar{C}_0 = \sum_{s=1}^5 C_{t-s}^0$	0.410*** (0.061)	0.092*** (0.013)	0.406*** (0.056)	0.091*** (0.011)	0.433*** (0.083)	0.097*** (0.017)
$\bar{C}_1 = \sum_{s=1}^5 C_{t-s}^1$	0.210*** (0.016)	0.047*** (0.006)	0.164*** (0.009)	0.037*** (0.003)	0.167* (0.092)	0.037* (0.020)
$\bar{D}_1 = \sum_{s=1}^5 D_{t-s}^1$	-0.123*** (0.046)	-0.028*** (0.008)	-0.180*** (0.038)	-0.040*** (0.006)	-0.195*** (0.063)	-0.043*** (0.014)
$\mathbb{1}_{\{F_{t-1}=1\}}$			0.230*** (0.070)	0.051*** (0.019)	0.199 (0.275)	0.044 (0.061)
$\mathbb{1}_{\{a_{j,t-1}=C\}}$			-0.019 (0.037)	-0.004 (0.008)	0.073 (0.111)	0.016 (0.025)
$a_{jt} = C$ in a row					-0.059 (0.037)	-0.013* (0.008)
$F_{it}$ in a row					0.020 (0.152)	0.004 (0.034)
$N$	599	—	599	—	599	—
$\sigma_u$	1.196	—	1.201	—	1.208	—
$\rho$	0.588	—	0.591	—	0.593	—
LL	-220.466	—	-219.610	—	-219.454	—

Probit models with individual random effects on the decision to cooperate at first stage estimated on the working sample. Standard errors (in parenthesis) are clustered at the session level. All specifications include control variables for gender, age, whether participant is a student, whether a fine applies to the first match, the decision to cooperate at first match, the length of the previous game and match number. Marginal effects are computed at sample mean, assuming random effects are 0. *Significance levels:* \*10%, \*\*5%, \*\*\*1%.

to full cooperation when it occurred 4 times. Another striking feature visible in Fig. 3, is that the rate of increase in cooperation is much faster as a function of the number of  $C_0$  than as a function of the number of  $C_1$  (cooperation of the partner in an environment with fines). More specifically, Fig. 3b shows that the average rate of cooperation increases from 0.15 when  $C_1$  never occurred in the 5 previous matches to 0.53 when it occurred 4 times. This reflects the idea, highlighted in the model, that cooperation in the absence of fine is a much stronger signal of intrinsic cooperativeness than cooperation in the presence of fines. The behavior for  $D_0$  and  $D_1$  is similar. The rate of cooperation tends to decrease more sharply with the number of  $D_1$  than with the number of  $D_0$ , even though the pattern is less striking than in the case of cooperation.

We confirm these graphical results in Table 2 where we estimate a Probit model on  $C_{it} = \mathbb{1}_{\{a_{it}=C\}} \in \{0, 1\}$ , the observed decision to cooperate of participant  $i$  in the first round of match  $t$  in the experiment. All estimated models control for current enforcement. Current fines have a very strong disciplining effect on current cooperation, increasing the probability of cooperation by more than 30%. In model (1), we do not account for spillovers and examine the effect of the history in the five previous matches.<sup>15</sup> The ranking of the effect is perfectly coherent with the results of Proposition 2: the signal  $C_0$  (variable  $\bar{C}_0$  in the table) has a positive significant effect compared to  $D_0$  (the reference) and the effect is larger than  $C_1$  (variable  $\bar{C}_1$ ). Similarly,  $D_1$  (variable  $\bar{D}_1$ ) decreases cooperation relative to  $D_0$ . In terms of magnitudes, replacing in the history one signal  $D_0$  by a signal  $C_0$  increases the probability of cooperation by 10% while replacing it by a signal  $C_1$  only increases the probability of cooperation by 5%. Replacing in the history one signal  $D_0$  by a signal  $D_1$  decreases the probability of cooperation by 3%.

In models (2) and (3), we control for potential spillovers. Model (2) introduces short-living spillovers by controlling for whether the previous match was played with a fine and whether the partner cooperated in the previous match. As explained previously, identification here relies on the assumption that spillovers are short lived, whereas learning is cumulative. Controlling for spillovers does not change the ordering of histories and only marginally affects magnitudes. Finally, in model (3), we relax the identifying assumption and allow spillovers to be longer lasting. We add a control for the number of matches in a row where partners cooperated, as well as the number of fines in a row in all previous matches. Identification here relies on the assumption that learning does not depend on the order in which signals were received, while it affects the strength of the effect of spillovers. None of these controls impact the results on learning, which still strongly affects how current cooperation react to past enforcement and behavior.

We provide an alternative identification strategy in the “[Alternative identification strategy](#)” in Appendix 2, where we model explicitly the interaction between learning and spillovers. To that end, we extend the model in Sect. 3 to the assumption that the taste for cooperation,  $\beta_i$ , is directly affected by the history of partner’s behavior and institutional settings. Proposition D shows that updated beliefs obeys the same ranking as in Proposition 2. This model explicitly shows that the learning and spillovers parameters cannot be separately identified when both affect cooperation simultaneously. The empirical analysis provided in “[The dynamics of cooperation with learning and spillovers](#)” in Appendix 2 relies on the assumption that learning has converged in games occurring late in the experiment so as to achieve separate identification of both kinds of parameters—i.e., estimate learning parameters conditional on the estimated spillovers. We test the predictions of the model and

<sup>15</sup> The specification of the empirical model is the same as the one used in the “[Alternative identification strategy](#)” in Appendix 2. All results are robust to alternative definitions of the number of previous matches included in the past history. The results are available from the authors upon request. The “[Replication of the main results with bootstrapped standard errors](#)” in Appendix 4, provides the results from a robustness exercise with bootstrapped standard errors.



confirm in Table 3 the ranking predicted in Proposition D. The similarity of the results between the two identification strategies confirm (i) that learning about values is a transitory issue that no longer affects cooperation once enough interactions has taken place and (ii) that the number of games implemented in our experiment gives enough room to learning so that only spillovers affect cooperative behavior in late games.

## 5 Conclusion

This paper studies cooperative behavior in a setting in which individuals interact without knowing each others' propensity to cooperate. In these situations, exogenous enforcement of cooperation may affect individuals' capacity to make inferences about the prevalent types in the society and, as a consequence, their propensity to cooperate.

We analyze this setting through the lens of a theoretical model tailored to interpret the results from an experiment where individuals play a series of infinitely repeated games with random re-matching. We rely on two different identification strategies to disentangle institution-specific learning from the effect of past enforcement on one's own willingness to cooperate (i.e., behavioral spillovers). The first relies on the fact that institution-specific learning, in contrast with spillovers, does not depend on the order in which a given history of cooperation occurred. The second, presented in the "Appendix 1", relies on the structure of the model and the (untestable) assumption that learning has converged in games occurring late in the experiment. The results provide strong support for the main behavioral insights of the model. The presence or absence of cooperation enforcing institutions affects the dynamics of learning about others' likely behavior: cooperation from partners faced in the past fosters cooperation today (with different partners) differently according to the institutional environment of past interactions. Past cooperation is more informative about other's cooperativeness when it is observed under weak enforcement institutions. Similarly, defection is more detrimental to cooperation when it was observed in an environment with strong enforcement.

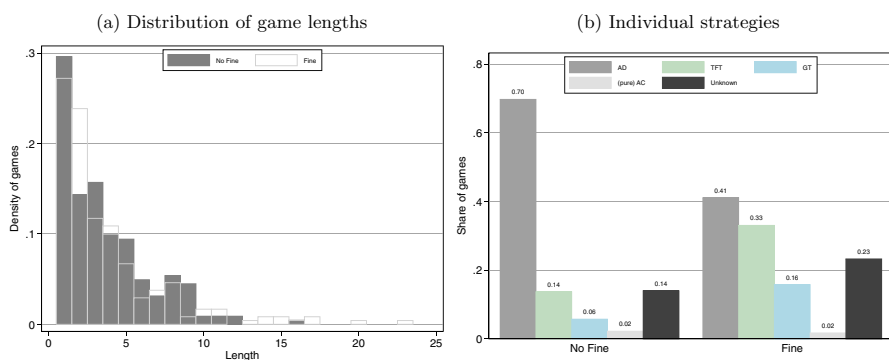
These results show that the choice of an institutional setting must be fine-tuned to the prevalent values in the target group. Strong enforcement aims at providing incentives to cooperate, which aren't necessary if the cooperative standards are high in the group to which enforcement applies. Our results show that such a mismatch between the institutional arrangement and the prevalent values comes with a cost whenever there is imperfect information about these values. Providing incentives to cooperate will slow down the virtuous dynamic in cooperation that would result from learning about the group values thanks to cooperative behavior observed without enforcement. Similarly, weak enforcement within an intrinsically non-cooperative group hinders the rate of learning that would occur from observing deviations under a stronger (and better suited) enforcement policy. These countervailing effects of the mismatch between values and enforcement typically applies to situations in which learning has not yet converged. In young organizations, in which the members of the

group need to learn about each others, offering incentives that best fit the underlying values in the group is key to the early success of the organization.

From a methodological point of view, our results show that games played in a sequence are related to one another even under random rematching for two reasons: first, under imperfect information about other players' preferences, the action observed in past matches provide information about the prevalent preferences in the population. Second, behavioral spillovers induce path-dependence in the willingness to cooperate across games. This suggests in particular that identifying spillovers, the focus of a large recent literature (see Galizzi & Whitmarsh, 2019, for a survey), can be challenging when the group members are also learning about prevalent values. This might lead to an underestimation of the size of spillovers in the case where the group has a low level of cooperation, since having fines might speed up learning and thus initially have a negative effect on cooperation. The similarity in the results between our two identification strategies however confirms that learning is transitory, and enough repetitions allows learning to converge so that path-dependence only results from behavioral spillovers in interactions that happen late enough in the sequence.

## Appendix 1: Data description

Our data delivers 934 game observations, 48% of which are played with no fine. Figure 4a displays the empirical distribution of game lengths in the sample split according to the draw of a fine. With the exception of two-rounds matches, the distributions are very similar between the two environments. This difference in the share of two-rounds matches mainly induces a slightly higher share of matches longer than 10 rounds played with a fine. In both environments, one third of the matches we observe lasts one round, and one half of the repeated matches last between 2 and 5



**Fig. 4** Sample characteristics: distribution of game lengths and repeated-game strategies. *Note:* *Left-hand side:* empirical distribution of game lengths in the experiment, split according to the draw of the fine. *Right-hand side:* distribution of repeated-games strategies observed in the experiment. One-round matches are excluded. AD: Always Defect; AC: Always Cooperate; TFT: Tit-For-Tat; GT: Grimm Trigger

rounds. A very small fraction of matches (less than 5% with a fine, less than 2% with no fine) feature lengths of 10 rounds or more.

As explained in the text, Sect. 2.2, for matches that last more than one round (2/3 of the sample), we thus reduce the observed outcomes to the first round decision in each match, consistently with the theory. The first round decision is a sufficient statistic for the future sequence of play if subjects choose among the following repeated-game strategies: Always Defect (AD), Tit-For-tat (TFT) or Grim Trigger (GT). While AD dictates defection at the first round, both TFT and GT induce cooperation and are observationally equivalent if the partner chooses within the set restricted to these three strategies and give rise to the same expected payoff. Figure 4b displays the distribution of strategies we observe in the experiment (excluding games that last one round only). Decisions are classified in each repeated game and for each player based on the observed sequence of play. For instance, a player who starts with *C* and switches forever to *D* when the partner starts playing *D* will be classified as playing GT. In many instances, TFT and GT cannot be distinguished (so that the classes displayed in Fig. 4b overlap): it happens for instance for subjects who always cooperate against a partner who does the same (in which case, TFT and GT also include Always Cooperate, AC), or if defection is played forever by both players once it occurs. Last, the Figure also reports the share of Always Cooperate that can be distinguished from other match strategies—when AC is played against partners who do defect at least once.

All sequences of decisions that do not fall in any of these strategies cannot be classified—this accounts for 14% of the games played without a fine, and 24% of those played with fine. The three strategies on which we focus are thus enough to summarize a vast majority of match decisions: AD accounts for 70% of the repeated-game observations with no fine, and 41% with a fine, while TFT and GT account for 14% and 34% of them.

## Appendix 2: Alternative identification strategy

The empirical evidence presented in the paper relies on the insights from the model to provide a reduced-form statistical analysis of the interaction between learning and enforcement institutions. As a complement to this evidence, this section provides a structural analysis which takes into account both learning and behavioral spillovers. To that end, we first generalize the model presented in Sect. 3.2 to the case in which individual values evolve over time as a result of past experience. We then estimate the parameters of the model. This provides an alternative strategy to separately estimate learning and spillovers.

### The dynamics of learning with behavioral spillovers

Consistent with Galbiati et al. (2018), we allow both for past fines and past behaviors of the partners to affect values:

$$\beta_{it} = \beta_i + \phi_F \mathbb{1}_{\{F_{it-1}=1\}} + \phi_C \mathbb{1}_{\{a_{j_{t-1},t-1}=C\}}. \quad (3)$$

According to this simple specification, personal values evolve through two channels. First, direct spillovers increase the value attached to cooperation in the current match if the previous one was played with a fine, as measured by parameter  $\phi_F$ . Second, indirect spillovers, measured by  $\phi_C$ , increase the value attached to cooperation if in the previous match the partner cooperated.<sup>16</sup>

### Introducing behavioral spillovers in the benchmark

We start by introducing spillovers in the benchmark model. Under the assumption that values follow the process in (3) and  $\phi_F > 0$ ,  $\phi_C > 0$ , the indifference condition (1) remains unchanged,<sup>17</sup> but now  $\beta_{it}$  is no longer constant and equal to  $\beta_i$  since past shocks affect values. In this context, individual  $i$  cooperates at  $t$  if and only if:

$$\beta_{it} \geq \Pi_1 - F \times \mathbb{1}_{\{F_{it}=1\}} + p_{it} \Pi_2 - \frac{\delta}{1-\delta} p_{it} (F \times \mathbb{1}_{\{F_{it}=1\}} + \Pi_3).$$

The cutoff value is defined in the same way as before:

$$\beta_t^*(F_{it}) = \Pi_1 - F \times \mathbb{1}_{\{F_{it}=1\}} + p_t^*(F_{it}) \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \times \mathbb{1}_{\{F_{it}=1\}} + \Pi_3) \right]. \quad (4)$$

The main difference with the benchmark model is in the value of  $p_t^*(F_{it})$ . There is now a linkage between the values of the cutoffs at match  $t$ ,  $\beta_t^*$ , and the values of the cutoffs  $\beta_{t'}^*$  in all the preceding matches  $t' < t$  through  $p_t^*(F_{it})$ . Indeed, when an individual evaluates the probability that her current partner in  $t$ , player  $j_t$ , will cooperate, she needs to determine how likely it is that she received a direct and/or an indirect spillover from the previous period. The probability of having a direct spillover is given by  $P[F_{j_{t-1}} = 1] = 1/2$  and is independent of any equilibrium decision. By contrast, the probability of having an indirect spillover is linked to whether the partner of  $j_t$  in her previous match cooperated or not. This probability in turn depends on the cutoffs in  $t-1$ ,  $\beta_{t-1}^*$ , which also depends on whether that individual himself received indirect spillovers, i.e. on the cutoff in  $t-2$ . Overall, these cutoffs in  $t$  depend on the entire sequence of cutoffs.

<sup>16</sup> The model can easily be extended to allow for longer histories to impact values. For instance, the effect on past institutions on values could be extended to:

$$\beta_{it} = \beta_i + \sum_{\tau=1}^T \phi_{F\tau} \mathbb{1}_{\{F_{it-\tau}=1\}} + \sum_{j=1}^T \phi_{C\tau} \mathbb{1}_{\{a_{j_{t-\tau},t-\tau}=C\}},$$

with  $\phi_{F\tau}$  and  $\phi_{C\tau}$  increase in  $\tau$ , in other words the more recent history having more impact. This could be introduced at the cost of added complexity.

<sup>17</sup> As stated above, we work under the assumption that players are myopic and choose between C and D to maximize their payoff in the current match. Without this assumption, when spillover are introduced, a player would need to take into account that her current action would influence her partner's future actions and thus influence the partner's future partners. An alternative would be to assume that players are negligible enough so that current actions cannot influence future beliefs.

In the remaining, we focus on stationary equilibria, such that  $\beta^*$  is independent of  $t$ . We show in Proposition C that such equilibria do exist.

**Proposition C** (Spillovers) *In an environment with spillovers ( $\phi_F > 0$  and  $\phi_C > 0$ ) and no uncertainty on values, there exists a stationary equilibrium. Furthermore all equilibria are of the cutoff form, i.e individuals cooperate if and only if  $\beta_{it} \geq \beta^*(F_{it})$ .*

**Proof** See “Appendix 5: Proofs”.  $\square$

Proposition C proves the existence of an equilibrium and presents the shape of the cutoffs. The Proposition also allows to express the probability that a random individual cooperates as:

$$1 - \Phi_H \left[ \Lambda_1 - \phi_F \mathbb{1}_{\{F_{it-1}=1\}} - \phi_C \mathbb{1}_{\{a_{j,t-1,t-1}=C\}} - \Lambda_2 \mathbb{1}_{\{F_{it}=1\}} \right], \quad (5)$$

where

$$\begin{aligned} \Lambda_1 &\equiv \beta^*(0) = \Pi_1 + p^*(0) \left[ \Pi_2 - \frac{\delta}{1-\delta} \Pi_3 \right], \\ \Lambda_2 &\equiv \beta^*(1) - \beta^*(0) = F + [p^*(0) - p^*(1)] \left[ \Pi_2 - \frac{\delta}{1-\delta} \Pi_3 \right] + \frac{\delta}{1-\delta} p^*(1) F. \end{aligned}$$

### The dynamics of cooperation with learning and spillovers

We now solve the full model with uncertainty about the group's values and with spillovers. As in the main text, we denote  $q_{it}$  the belief held by player  $i$  at match  $t$  that the state is  $H$ .

In this expanded model, the beliefs on how likely it is that the partner cooperates in match  $t$ ,  $p_t^*(F_{it}, q_{it})$ , depends on the probability that the partner experienced spillovers. In addition, the probability that the partner  $j$  had an indirect spillover itself depends on whether his own partner  $k$  in the previous match did cooperate, and thus depends on the beliefs  $q_{kt-1}$  of that partner in the previous match. The general problem requires to keep track of the higher order beliefs. The proof of the following Proposition shows the existence of such a stationary equilibrium, under a natural restriction on higher order beliefs, i.e if we assume that a player who had belief  $q_{it}$  in match  $t$  believes that players in the preceding match had the same beliefs  $q_{j',t-1} = q_{it}$ .

**Proposition D** *In an environment with spillovers and learning, if an equilibrium exists, all equilibria are of the cutoff form, i.e individuals cooperate if and only if  $\beta_i \geq \beta^*(F_{it}, q_{it})$ . Furthermore, if in equilibrium  $\beta^*(0, q) > \beta^*(1, q)$  for all beliefs  $q$ , then the beliefs are updated in the following way given the history in the previous interaction:*

$$\begin{aligned} q_{it}(0, C, q_{it-1}) &> q_{it}(1, C, q_{it-1}) > q_{it-1}, \\ q_{it}(1, D, q_{it-1}) &< q_{it}(0, D, q_{it-1}) < q_{it-1}. \end{aligned}$$

**Proof** See “Appendix 5: Proofs”. □

Proposition D derives a general property of equilibria. The Proposition also allows to express the probability of cooperation for given belief  $q_{it-1}$  as:

$$\begin{aligned} 1 - \Phi_H \left[ \Lambda_3 - \phi_F \mathbb{1}_{\{F_{it-1}=1\}} - \phi_C \mathbb{1}_{\{a_{j_{t-1},j-1}=C\}} - \Lambda_4 \mathbb{1}_{\{F_{it}=1\}} \right. \\ \left. - \sum_{j,k \in \{0,1\}, l \in \{C,D\}} \Lambda_{k,l}^j \mathbb{1}_{\{F_{it}=j, F_{it-1}=k, a_{j_{t-1},j-1}=l\}} \right]. \end{aligned} \quad (6)$$

where

$$\begin{aligned} \Lambda_3 &\equiv \Pi_1 + p^*(0, 0, D, q) \left[ \Pi_2 - \frac{\delta}{1-\delta} \Pi_3 \right], \\ \Lambda_4 &\equiv F - [p^*(1, 0, D, q) - p^*(0, 0, D, q)] \\ &\quad \left[ \Pi_2 - \frac{\delta}{1-\delta} \Pi_3 \right] + \frac{\delta}{1-\delta} p^*(1, 0, D, q) F > 0, \\ \Lambda_{k,l}^0 &\equiv [p^*(0, k, l, q) - p^*(0, 0, D, q)] \left[ \Pi_2 - \frac{\delta}{1-\delta} \Pi_3 \right], \\ \Lambda_{k,l}^1 &\equiv -[p^*(0, k, l, q) - p^*(0, 0, D, q)] \left[ \Pi_2 - \frac{\delta}{1-\delta} (F + \Pi_3) \right]. \end{aligned}$$

Note that the parameters  $\Lambda_3$ ,  $\Lambda_4$  and  $\Lambda_{k,l}^j$  in Eq. (6) depend on  $q_{it-1}$ . Compared to the case without learning, there are 6 additional parameters, reflecting the updating of beliefs. According to the result in Proposition D, these parameters, both in the case where the current match is played with a fine and when it is not, are such that:

$$\begin{aligned} \Lambda_{0,C}^1 &> \Lambda_{1,C}^1 > 0 > \Lambda_{1,D}^1, \\ \Lambda_{0,C}^0 &> \Lambda_{1,C}^0 > 0 > \Lambda_{1,D}^0. \end{aligned} \quad (7)$$

Overall, having fines in the previous match can potentially decrease average cooperation in the current match. There are two countervailing effects. On the one hand, a fine in the previous match increases the direct and indirect spillovers and thus increases cooperation. On the other hand, if the state is low, a fine can accelerate learning if, on average, sufficiently many people deviate in the presence of a fine. This then decreases cooperation in the current match.

### Statistical implementation of the model

The main behavioral insights from the model are summarized by Eq. (6), which involves both learning and spillover parameters. As the equation clearly shows, exogenous variations in legal enforcement are not enough to achieve separate identification of learning and spillover parameters—an exogenous change in

any of the enforcement variables, or past behavior of the partner, involves both learning and a change in the values  $\beta_{it}$ . In the main text, identification relies on the assumption that spillovers are short-living in the sense that their effect on behavior is smaller the earlier they happen in one's own history—while learning should not depend on the order in which a given sequence of actions happens. In this section, we report the results from an alternative identification strategy that relies on the assumption that learning has converged once a large enough number of matches has been played. Under this assumption, in late games, behavior is described by Eq. (5), which involves only spillover parameters. Exogenous variations in enforcement thus provide identification of spillover parameters in late games, which in turn allows to identify learning parameters in early ones.

To that end, as explained in the text, we split the matches in three groups, in such a way that one third of the observed decisions are classified as “early”, and one third as “late”. We use matches, rather than periods, as a measure of time since we focus on games for which the first stage decision summarizes all future actions within the current repeated game—hence ruling out learning within a match. Observed matches are accordingly defined as “early” up to the 7th, as late after the 13th—we disregard data coming from intermediary stages.<sup>18</sup> Denote  $\mathbb{1}_{\{\text{Early}\}}$  the dummy variable equal to 1 in early games and to 0 in late games. Under the identifying assumption that learning has converged in late games, the model predicts that behavior in the experiment is described by:

$$\begin{aligned}
 P(C_{it} = 1) = & 1 - \Phi_H \left[ \Lambda_1 + (\Lambda_3 - \Lambda_1) \mathbb{1}_{\{\text{Early}\}} - \phi_F \mathbb{1}_{\{F_{it-1}=1\}} \right. \\
 & - \phi_C \mathbb{1}_{\{a_{j_{t-1},j-1}=C\}} - \Lambda_2 \mathbb{1}_{\{F_{it}=1\}} \\
 & - (\Lambda_4 - \Lambda_2) \mathbb{1}_{\{F_{it}=1\}} \times \mathbb{1}_{\{\text{Early}\}} \\
 & \left. - \sum_{j,k \in \{0,1\}, l \in \{C,D\}} \Lambda_{k,l}^j \mathbb{1}_{\{F_{it}=j, F_{it-1}=k, a_{j_{t-1},j-1}=l\}} \times \mathbb{1}_{\{\text{Early}\}} \right]
 \end{aligned}$$

which is the structural form of a Probit model on the individual decision to cooperate. This probability results from equilibria of the cutoff form involving the primitives of the model. Denoting  $\varepsilon_{it}$  observation specific unobserved heterogeneity,  $\theta$  the vector of unknown parameters embedded in the above equation,  $\mathbf{x}_{it}$  the associated set of observables describing participant  $i$  experience up to  $t$  and  $C_{it}^*$  the latent function generating player  $i$  willingness to cooperate at match  $t$ , observed decisions inform about the model parameters according to:

$$C_{it} = \mathbb{1}[C_{it}^* = \mathbf{x}_{it}\theta + \varepsilon_{it} > 0]$$

<sup>18</sup> All results are robust to alternative definitions of these thresholds. The results are available from authors upon request.

**Table 3** Learning and spillovers arising from past enforcement

Variable	Model	(1)	(2)	(3)	(4)	(5)
	Parameter					
Constant	$(-A_1)$	-1.986*** (0.392)	-2.020*** (0.362)	-2.290*** (0.298)	-2.302*** (0.317)	-2.291*** (0.220)
$\mathbb{1}_{\{F_{it}\}}$	$(A_2)$	1.448*** (0.164)	1.454*** (0.149)	1.480*** (0.150)	1.473*** (0.142)	1.472*** (0.137)
Early	$(-A_3 + A_1)$	0.285 (0.411)	0.292 (0.415)	0.348 (0.440)	0.460 (0.383)	0.453 (0.377)
Early $\times \mathbb{1}_{\{F_{it}\}}$	$(A_4 - A_2)$	-0.698*** (0.243)	-0.698*** (0.245)	-0.646** (0.258)	-0.644** (0.261)	-0.643** (0.271)
$\mathbb{1}_{\{F_{it-1}\}}$	$(\phi_F)$		0.049 (0.120)	0.306** (0.140)	0.094 (0.193)	0.085 (0.306)
$\mathbb{1}_{\{a_{j,t-1}=C\}}$	$(\phi_C)$				0.693*** (0.169)	0.674*** (0.121)
Early $\times C_0$	$(A_{0,C})$			1.066*** (0.186)	0.430 (0.363)	0.448* (0.246)
Early $\times C_1$	$(A_{1,C})$			0.233 (0.168)	-0.228** (0.096)	-0.230** (0.094)
Early $\times D_1$	$(A_{1,D})$			-0.876** (0.423)	-0.631* (0.329)	-0.621* (0.334)
$C_1$						0.029 (0.380)
$N$		553	553	553	553	553
$\sigma_u$		1.063	1.064	1.063	1.060	1.060
$\rho$		0.531	0.531	0.530	0.529	0.529
LL		-234.677	-234.624	-224.416	-220.033	-220.031

Probit models with individual random effects on the decision to cooperate at first stage estimated on the working sample restricted to early (before the 7th) and late (beyond the 13th) games. Standard errors (in parenthesis) are clustered at the session level. All specifications include control variables for gender, age, whether participant is a student, whether a fine applies to the first match, the decision to cooperate at first match, the length of the previous game and match number. *Significance levels*: \*10%, \*\*5%, \*\*\*1%.

The structural parameters govern the latent equation of the model. Our empirical test of the model is thus based on estimated coefficients,  $\theta = \frac{\partial C^*}{\partial x_{it}}$ , rather than marginal effects,  $\frac{\partial C}{\partial x_{it}} = \theta \frac{\partial \Phi(x_{it}\theta)}{\partial x_{it}}$ .

In the set of covariates, both current ( $\mathbb{1}_{\{F_{it}=1\}}$ ) and past enforcement ( $\mathbb{1}_{\{F_{it-1}=1\}}$ ) are exogenous by design. The partner's past decision to cooperate,  $C_{jt-1}$ , is exogenous to  $C_{it}$  as long as player  $i$  and  $j$  have no other player in common in their history. Moreover, due to the rematching of players from one match to the other, between subjects correlation arises over time within an experimental session. We address these concerns in three ways. We include the decision to cooperate at the first



stage of the first match in the set of control variables, as a measure of individual unobserved *ex ante* willingness to cooperate. To further account for the correlation structure in the error of the model, we specify a panel data model with random effects at the individual level, control for the effect of time thanks to the inclusion of the match number, and cluster the errors at the session level to account in a flexible way for within sessions correlation.

Table 3 reports the estimation results from several specifications, in which each piece of the model is introduced sequentially. The parameters of interest are the learning parameters  $A_{k,l}$ ,  $k \in \{0, 1\}$ ,  $l \in \{C, D\}$ .<sup>19</sup> Columns (1) and (2) focus on the effect of past and current enforcement. While we do not find any significant change due to moving from early to late games *per se* (the Early variable is not significant), the effect of current enforcement on the current willingness to cooperate is much weaker in early games. This is consistent with participants becoming less confident that the group is cooperative, thus less likely to cooperate, as time passes—i.e., prior belief over-estimate the average cooperativeness of the group. The disciplining effect of current fines is thus stronger in late games.

Column (3) introduces learning parameters. As stressed above, the learning parameters play a role before beliefs have converged. They are thus estimated in interaction with the Early dummy variable. Once learning is taken into account, enforcement spillovers turn-out significant. More importantly, the model predicts that learning is stronger when observed decisions are more informative about societal values, which in turn depends on the enforcement regime under which behavior has been observed: cooperation is more informative about cooperativeness under weak enforcement, while defection is stronger signal of non-cooperative values under strong enforcement. This results in a clear ranking between learning parameters—see Eq. (7). We use defection under weak enforcement as a reference for the estimated learning parameters. The results show that cooperation under weak enforcement ( $\text{Early} \times C_0$ ) leads to the strongest increase in the current willingness to cooperate. Observing this same decision but under strong enforcement institutions rather than weak ones ( $\text{Early} \times C_1$ ) has almost the same impact as observing defection under strong institutions (the reference): in both cases, behavior is aligned with the incentives implemented by the rules and barely provides any additional insights about the distribution of values in the group. Last, defection under strong institutions ( $\text{Early} \times D_0$ ) is informative about a low willingness to cooperate in the group, and results in a strongly significant drop in current cooperation.

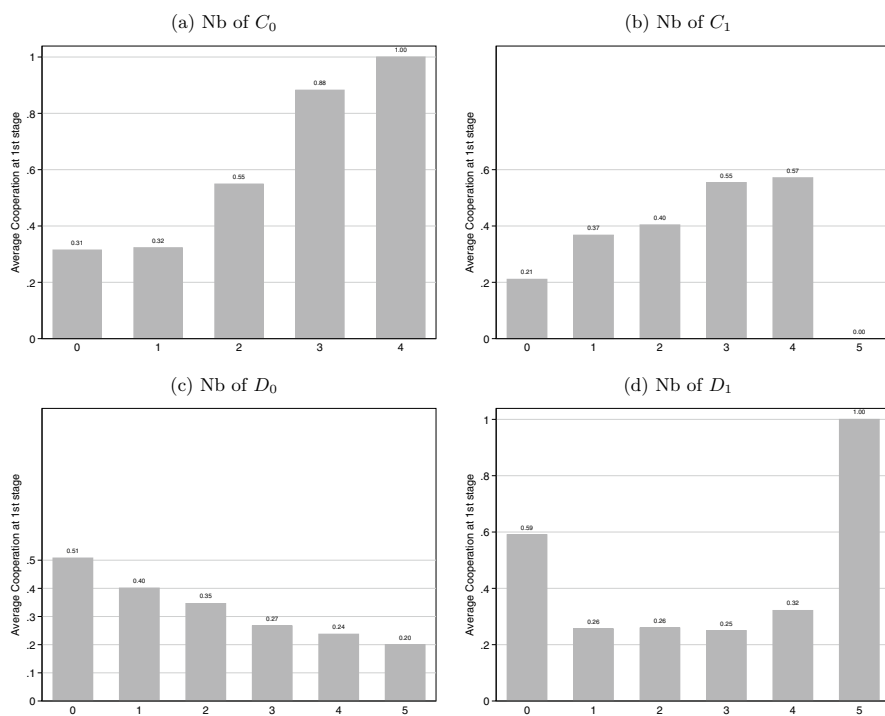
Column (4) adds indirect spillovers, induced by the cooperation of the partner in the previous game. The identification of learning parameters in this specification is quite demanding since both past enforcement and past cooperation are included as dummy variables. We nevertheless observe a statistically significant effect of learning in early games, with the expected ordering according to how informative the signal delivered by a cooperative decision is, with the exception of  $C_1$ —i.e., when cooperation has been observed under fines. Finally, column (5) provides a

<sup>19</sup> Note that we do not separately estimate these parameters according to the current enforcement environment, but rather estimate weighted averages  $A_{k,l} = \mathbb{1}_{\{F_{it}=0\}} A_{k,l}^0 + \mathbb{1}_{\{F_{it}=1\}} A_{k,l}^1$ .

robustness check for the reliability of the assumption that learning has converged in late games. To that end, we further add the interaction between observed behavior from partner in the previous game and the enforcement regime. Once learning has converged, past behavior is assumed to affect the current willingness to cooperate through indirect spillovers only. Absent learning, this effect should not interact with the enforcement rule that elicited this behavior. As expected, this interaction term is not significant: in late games, it is cooperation per se, rather than the enforcement regime giving rise to this decision, that matters for current cooperation.

### Appendix 3: Replication of the statistical analysis on the full sample

In this section, we replicate the results on the full sample of observations: instead of restricting the analysis to the working sample made of decisions consistent with the subset of repeated-game strategies described in Sect. 2.2, we include all available observations. As already mentioned in the text, the matches we exclude from the working sample all appear in late games so that Fig. 2a is not affected by the choice of the working sample. Figure 5 below replicates Fig. 3 in the paper; and Table 4 replicates Table 2. In both instances, the data is more noisy but the qualitative conclusions all remain the same.



**Fig. 5** Current cooperation as a function of history in the previous 5 games, computed on the full sample

**Table 4** Replication of Table 2 on the full sample

	(1) Enforcement	(2) Spillovers	(3)	(4)	(5)	(6) est6
Constant	− 0.029 (0.736)		0.037 (0.733)		0.126 (0.847)	
$\mathbb{1}_{\{F_t=1\}}$	1.172*** (0.279)	0.310*** (0.034)	1.168*** (0.278)	0.307*** (0.034)	1.165*** (0.265)	0.306*** (0.031)
$\bar{C}_0 = \sum_{s=1}^5 C_{t-s}^0$	0.347** (0.142)	0.092*** (0.032)	0.321** (0.149)	0.084** (0.034)	0.361* (0.185)	0.095** (0.040)
$\bar{C}_1 = \sum_{s=1}^5 C_{t-s}^1$	0.196** (0.090)	0.052** (0.022)	0.122 (0.086)	0.032 (0.022)	0.101 (0.148)	0.027 (0.037)
$\bar{D}_1 = \sum_{s=1}^5 D_{t-s}^1$	− 0.049 (0.036)	− 0.013* (0.008)	− 0.090* (0.052)	− 0.024** (0.011)	− 0.133** (0.060)	− 0.035** (0.015)
$\mathbb{1}_{\{F_{t-1}=1\}}$			0.195** (0.083)	0.051*** (0.017)	0.089 (0.232)	0.023 (0.059)
$\mathbb{1}_{\{a_{jt-1}=C\}}$			0.146** (0.074)	0.039* (0.022)	0.258* (0.137)	0.068** (0.031)
$a_{jt} = C$ in a row					− 0.079 (0.081)	− 0.021 (0.019)
$F_{it}$ in a row					0.071 (0.109)	0.019 (0.030)
$N$	694	−	694	−	694	−
$\sigma_u$	1.099	−	1.100	−	1.113	−
$\rho$	0.547	−	0.547	−	0.553	−
LL	− 300.754	−	− 298.972	−	− 298.235	−

Marginal effects are computed at sample mean, assuming random effects are 0. *Significance levels:* \*10%, \*\*5%, \*\*\*1%.

## Appendix 4: Replication of the main results with bootstrapped standard errors

As explained in the main text, the statistical analysis presented in Table 2 clusters the standard errors at the session level, so as to take into account in a flexible way the possible correlations between subjects due to random rematching of subjects within pairs. While this approach is conservative (since it does not impose any structure on the correlation between subjects over time), the number of clusters is small and there is a risk of small sample downward bias in the estimated standard errors. Table 5 provides the results from a robustness exercise replicating Table 2 in the text with bootstrapped standard error based a delete-one jackknife procedure (see Bell & McCaffrey, 2002; Cameron et al. 2008).

**Table 5** Replication of Table 2 using bootstrapped errors

	(1) Enforcement	(2) Spillovers	(3)	(4)	(5)	(6) est6
Main constant	0.040 (0.167)		0.147 (0.217)		0.179 (0.258)	
$\mathbb{1}_{\{F_i=1\}}$	1.356** (0.281)	0.305** (0.056)	1.361** (0.280)	0.303** (0.055)	1.355** (0.285)	0.302** (0.052)
$\bar{C}_0 = \sum_{s=1}^5 C_{t-s}^0$	0.410** (0.078)	0.092** (0.015)	0.406** (0.074)	0.091** (0.014)	0.433** (0.095)	0.097** (0.018)
$\bar{C}_1 = \sum_{s=1}^5 C_{t-s}^1$	0.210*** (0.018)	0.047** (0.007)	0.164*** (0.006)	0.037*** (0.003)	0.167 (0.111)	0.037 (0.024)
$\bar{D}_0 = \sum_{s=1}^5 D_{t-s}^1$	-0.123 (0.048)	-0.028* (0.009)	-0.180** (0.041)	-0.040** (0.007)	-0.195 (0.073)	-0.043 (0.016)
$\mathbb{1}_{\{F_{t-1}=1\}}$			0.230* (0.060)	0.051* (0.017)	0.199 (0.323)	0.044 (0.072)
$\mathbb{1}_{\{a_{jt-1}=C\}}$			-0.019 (0.041)	-0.004 (0.009)	0.073 (0.140)	0.016 (0.031)
$a_{jt} = C$ in a row					-0.059 (0.053)	-0.013 (0.011)
$F_{it}$ in a row					0.020 (0.182)	0.004 (0.041)
$N$	599	599	599	599	599	599
$\sigma_u$	1.196	1.196	1.201	1.201	1.208	1.208
$\rho$	0.588	0.588	0.591	0.591	0.593	0.593
LL	-220.466	-220.466	-219.610	-219.610	-219.454	-219.454

Marginal effects are computed at sample mean, assuming random effects are 0. *Significance levels:* \*10%, \*\*5%, \*\*\*1%.

## Appendix 5: Proofs

### Proof of Proposition 1

As derived in the main text, if an equilibrium exists, it is necessarily such that players use cutoff strategies. Reexpressing characteristic Eq. (2), we can show that the cutoffs are determined by the equation  $g(\beta^*[F_{it}]) = 0$ , where  $g$  is given by

$$g(x) = -x + \Pi_1 - F\mathbb{1}_{\{F_{it}=1\}} + (1 - \Phi_H(x)) \left[ \Pi_2 - \frac{\delta}{1 - \delta} (F\mathbb{1}_{\{F_{it}=1\}} + \Pi_3) \right].$$

The function  $g$  has the following properties:  $g(x) > 0$  when  $x$  converges to  $-\infty$  and  $g(x) < 0$  when  $x$  converges to  $+\infty$ . Thus, since  $g$  is continuous, there is at least one solution to the equation  $g(\beta^*[F_{it}]) = 0$ . At least one equilibrium exists.

If  $g$  is non monotonic, there could exist multiple equilibria. However, in all stable equilibria,  $\beta^*$  is such that  $g$  is decreasing at  $\beta^*$ , i.e.,

$$-1 - \phi_H[\beta^*] \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \mathbb{1}_{\{F_{ii}=1\}} + \Pi_3) \right] < 0. \quad (8)$$

Using the implicit theorem we have:

$$\frac{\partial \beta^*}{\partial F} = -\frac{\partial g}{\partial F} / \frac{\partial g}{\partial \beta^*} = -\frac{-1 - (1 - \Phi_H[\beta^*]) \frac{\delta}{1-\delta}}{-1 - \phi_H[\beta^*] \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \mathbb{1}_{\{F_{ii}=1\}} + \Pi_3) \right]},$$

where  $\phi_H$  is the density corresponding to distribution  $\Phi_H$ . For stable equilibria, the denominator is negative as shown in (8), so that overall

$$\frac{\partial \beta^*}{\partial F} < 0.$$

Similarly,

$$\frac{\partial \beta^*}{\partial \mu_H} = -\frac{-\frac{\partial \Phi_H[\beta^*]}{\partial \mu_H} \left[ \Gamma_2 - \frac{\delta}{1-\delta} (F \mathbb{1}_{\{F_{ii}=1\}} + \Pi_3) \right]}{-1 - \phi_H[\beta^*] \left[ \Pi_2 - \frac{\delta}{1-\delta} (F \mathbb{1}_{\{F_{ii}=1\}} + \Pi_3) \right]}.$$

Again, in stable equilibria the denominator is negative by (8). Furthermore we have  $\frac{\partial \Phi_H[\beta^*]}{\partial \mu_H} < 0$  since an increase in the mean of the normal distribution decreases  $\Phi_H[x]$  for any  $x$ . Overall we get

$$\frac{\partial \beta^*}{\partial \mu_H} < 0.$$

### Proof of Lemma 1

We first show the result:  $q_{i2}(1, D) < q_{i2}(0, D) < q_0$ . According to Baye's rule, the belief that the state is  $H$  following a deviation by the partner in the first match (which has been played with a fine) is:

$$\begin{aligned} q_{i2}(1, D) &= \frac{q_0 P[D|F_{i1} = 1, s = H]}{q_0 P[D|F_{i1} = 1, s = H] + (1 - q_0) P[D|F_{i1} = 1, s = L]} \\ &= \frac{q_0 \Phi_H[\beta^*(1)]}{q_0 \Phi_H[\beta^*(1)] + (1 - q_0) \Phi_L[\beta^*(1)]} \\ &= \frac{1}{1 + \frac{1-q_0}{q_0} \frac{\Phi_L[\beta^*(1)]}{\Phi_H[\beta^*(1)]}}. \end{aligned} \quad (9)$$

Furthermore, since  $\Phi_L[\beta^*(1)] > \Phi_H[\beta^*(1)]$ , we have  $q_{i2}(1, D) < q_0$ . Similarly we have:

$$q_{i2}(0, D) = \frac{1}{1 + \frac{1-q_0}{q_0} \frac{\Phi_L[\beta^*(0)]}{\Phi_H[\beta^*(0)]}} > q_0. \quad (10)$$

Thus,

$$q_{i2}(1, D) < q_{i2}(0, D) \Leftrightarrow \frac{\Phi_L[\beta^*(0)]}{\Phi_H[\beta^*(0)]} < \frac{\Phi_L[\beta^*(1)]}{\Phi_H[\beta^*(1)]}.$$

Using the fact that  $\frac{\Phi_L[x]}{\Phi_H[x]}$  is decreasing in  $x$  as shown in Property 1 below, and the fact that in stable equilibria, we have  $\beta^*(1) \leq \beta^*(0)$ , as shown in Proposition 1, implies directly that  $q_{i2}(1, D) < q_{i2}(0, D)$ . The proof that  $q_{i2}(0, C) > q_{i2}(1, C) > q_0$  follows similar lines.

**Property 1**  $\frac{\Phi_H[x]}{\Phi_L[x]}$  is increasing in  $x$ .

**Proof** Denote  $\phi_H$  (resp.  $\phi_L$ ) the density of  $\Phi_H$  (resp.  $\Phi_L$ ). Given that  $\phi_H$  (resp.  $\phi_L$ ) is the density of a normal distribution with standard deviation  $\sigma$  and mean  $\mu_H$  (resp.  $\mu_L$ ), it is the case that:

$$\begin{aligned} \frac{\phi_H[x]}{\phi_L[x]} &= \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu_H)^2/\sigma^2}}{\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu_L)^2/\sigma^2}} \\ &= e^{-(x-\mu_H)^2/\sigma^2 + (x-\mu_L)^2/\sigma^2} \\ &= e^{\frac{1}{\sigma^2}(\mu_H - \mu_L)(2x - \mu_L - \mu_H)}. \end{aligned}$$

Thus  $\frac{\phi_H}{\phi_L}$  is increasing in  $x$ . In particular for  $y < x$ , we have:  $\phi_H[y]\phi_L[x] < \phi_L[y]\phi_H[x]$ . By definition,  $\Phi_s(x) = \int_{-\infty}^x \phi_s(y)dy$ . Integrating with respect to  $y$  between  $-\infty$  and  $x$  thus yields:

$$\Phi_H[x]\phi_L[x] < \Phi_L[x]\phi_H[x]. \quad (11)$$

Consider now the function  $\frac{\Phi_H}{\Phi_L}$ . The derivative of this function is given by  $\frac{\phi_H\Phi_L - \Phi_L\phi_H}{\Phi_L^2}$ , which is positive by Eq. (11). This establishes Property 1 that  $\frac{\Phi_H[x]}{\Phi_L[x]}$  is increasing in  $x$ .  $\square$

### Proof of Proposition D (that generalizes Proposition 2)

In the first part of the proof we assume a stationary equilibrium exists and is such that the equilibrium cutoffs are always higher without a fine in the current match  $\beta^*(0, q) > \beta^*(1, q)$  for any given belief  $q$ . We then derive the property on updating of beliefs. In the second part of the proof we show existence under a natural restriction on beliefs.

**Part 1:** We first derive the properties on updating. We have

$$\begin{aligned}
 & q_{it}(F_{it-1}, a_{j_{t-1}, t-1}, q_{it-1}) \\
 &= \frac{q_{it-1}P[a_{j_{t-1}, t-1}|F_{it-1}, s = H]}{q_{it-1}P[a_{j_{t-1}, t-1}|F_{it-1}, s = H] + (1 - q_{it-1})P[a_{j_{t-1}, t-1}|F_{it-1}, s = L]}. \quad (12)
 \end{aligned}$$

We can express the probability that the partner  $j_{t-1}$  in match  $t - 1$  cooperated, by considering all the possible environments this individual might have faced in the past, in particular what his partner in match  $t - 2$ , individual  $k_{t-1}$  chose:

$$\begin{aligned}
 P[a_{j_{t-1}, t-1} = D | F_{it-1}, s = H] \\
 &= \sum_{F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}} \Phi_H \left[ \beta^*(1, q_{j_{t-1}, t-1}) - \phi_F \mathbb{I}_{\{F_{j_{t-1}, t-2} = 1\}} \right. \\
 &\quad \left. - \phi_C \mathbb{I}_{\{a_{k_{t-1}, t-2} = C\}} \right] \\
 &\quad \times P[F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}].
 \end{aligned}$$

Denote

$$\begin{aligned}
 & \gamma^*(F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}) \\
 &= \beta^*(1, q_{j_{t-1}, t-1}) - \phi_F \mathbb{I}_{\{F_{j_{t-1}, t-2} = 1\}} - \phi_C \mathbb{I}_{\{a_{k_{t-1}, t-2} = C\}}.
 \end{aligned}$$

and

$$R(x) \equiv \frac{\sum_{F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}} \Phi_H [\gamma^*(x, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1})] P[F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}]}{\sum_{F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}} \Phi_L [\gamma^*(x, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1})] P[F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}]}.$$

Using expression (12), we have:

$$q_{it}(1, D, q_{it-1}) < q_{it}(0, D, q_{it-1}) \Leftrightarrow R(1) \geq R(0).$$

We then use all possible values of the vector  $(F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1})$  in turn. Take such a value  $v$  for this vector and denote

$$\begin{aligned}
 a &\equiv \sum_{(F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}) \neq v} \Phi_H [\gamma^*(x, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1})] \\
 &\quad P[F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}], \\
 b &\equiv \sum_{(F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}) \neq v} \Phi_L [\gamma^*(x, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1})] \\
 &\quad P[F_{j_{t-1}, t-2}, a_{k_{t-1}, t-2}, q_{j_{t-1}, t-1}].
 \end{aligned}$$

We clearly have  $a < b$ . Furthermore, we can write

$$R(x) \equiv \frac{a + \Phi_H[\gamma^*(v)]P[v]}{b + \Phi_L[\gamma^*(v)]P[v]}$$

We have  $\beta^*(0, q) > \beta^*(1, q)$  imply that  $\gamma^*(0, a_{k't-2}, q_{j't-1}) > \gamma^*(1, a_{k't-2}, q_{j't-1})$ . Thus, using Property 2 below, it implies that  $R(1) \geq R(0)$  and thus  $q_{it}(1, D, q_{it-1}) < q_{it}(0, D, q_{it-1})$ .

**Property 2**  $\frac{a+p\Phi_H[x]}{b+p\Phi_L[x]}$  where  $b > a$  is increasing in  $x$ .

**Proof** The derivative of the ratio is given by

$$\frac{\phi_H(b + p\Phi_L) - \phi_L(a + p\Phi_H)}{(b + p\Phi_L)^2} \quad (13)$$

We showed in the proof of Property 1 that:  $\phi_H\Phi_L - \phi_L\Phi_H > 0$ . Furthermore, we also showed that  $\frac{\phi_H}{\phi_L}$  is increasing and since  $a < b$  this implies:  $\phi_H b - \phi_L a > 0$ . Combining these two results in condition (13) establishes Property 2.  $\square$

**Part 2:** We show that an equilibrium exists if we assume that a player who has belief  $q_{it}$  in match  $t$  believes that other players in match  $t$  and  $t - 1$  shared the same belief  $q_{j,t} = q_{k,t} = q_{it}$ .

If a stationary equilibrium exists, it is necessarily such that players use cutoff strategies where the cutoff is defined by:

$$\beta_t^*(F_{it}, q_{it}) = \Pi_1 - F\mathbb{1}_{\{F_{it}=1\}} + p_t^*(F_{it}, q_{it}) \left[ \Pi_2 - \frac{\delta}{1-\delta} (F\mathbb{1}_{\{F_{it}=1\}} + \Pi_3) \right]$$

We have:

$$\begin{aligned} p_t^*(F_{it}, q_{it}) &= P[a_{j,t} = C | F_{it}, q_{it}] \\ &= q_{it} \sum_{(F_{j,t-1}, a_{k,t-1}, q_{j,t-1})} \left[ 1 - \Phi_H[\beta^*(F_{j,t}, q_{j,t-1}) \right. \\ &\quad \left. - \phi_F\mathbb{1}_{\{F_{j,t-1}=1\}} - \phi_C\mathbb{1}_{\{a_{k,t-1}=C\}} \right] \\ &\quad \times P[F_{j,t-1}, a_{k,t-1}, q_{j,t-1} | s = H] \\ &\quad + (1 - q_{it}) \sum_{(F_{j,t-1}, a_{k,t-1}, q_{j,t-1})} \left[ 1 - \Phi_L[\beta^*(F_{j,t}, q_{j,t-1}) \right. \\ &\quad \left. - \phi_F\mathbb{1}_{\{F_{j,t-1}=1\}} - \phi_C\mathbb{1}_{\{a_{k,t-1}=C\}} \right] \\ &\quad \times P[F_{j,t-1}, a_{k,t-1}, q_{j,t-1} | s = L] \end{aligned}$$

Furthermore, we have

$$\begin{aligned} P[F_{j,t-1}, a_{k,t-1}, q_{j,t-1} | s = H] &= P[F_{j,t-1}] \\ P[a_{k,t-1} | F_{j,t-1}, s = H] &= f_t(q_{j,t-1} | s = H) \\ &= \frac{1}{2} P[a_{k,t-1} | F_{j,t-1}, s = H] f_t(q_{j,t-1} | s = H) \end{aligned}$$



we assumed that a player who had belief  $q_{it-1}$  in match  $t-1$  believes that all other players in that match share the same belief  $q_{it-1}$ . Under this restriction, we have  $f_t(q_{j't-1}|s = \cdot) = \mathbb{1}_{(q_{j't-1}=q_{it-1})}$

$$P[1, D, q|s = H] = \frac{1}{2}p_t^*(1, q)$$

We get a similar expression as in the proof of Proposition 2:

$$\begin{aligned} p^*(F_{it}, q_{it}) &= [1 - \Phi_H[\beta^*(F_{it}, q_{it}) - \phi_F - \phi_C]] \\ &\quad \frac{1}{2}p^*(1, q_{it}) + [1 - \Phi_H[\beta^*(F_{it}, q_{it}) - \phi_F]] \frac{1}{2}[1 - p^*(1, q_{it})] \\ &\quad + [1 - \Phi_H[\beta^*(F_{it}, q_{it}) - \phi_C]] \frac{1}{2}p^*(0, q_{it}) + \\ &\quad [1 - \Phi_H[\beta^*(F_{it}, q_{it})]] \frac{1}{2}[1 - p^*(0, q_{it})]. \end{aligned} \quad (14)$$

This implies that for each belief  $q$ , there is a system of equation equivalent to system A in the proof of Proposition 2. We thus have a solution of this system for each value  $q$ .

**Acknowledgements** This paper supersedes “Learning, Spillovers and Persistence: Institutions and the Dynamics of Cooperation”, CEPR DP n° 12128. We thank Bethany Kirkpatrick for her help in running the experiment, and Gani Aldashev, Maria Bigoni, Frédéric Koessler, Bentley McLeod, Nathan Nunn, Jan Sonntag, Sri Srikanthan and Francisco Ruiz Aliseda as well as participants to seminars at ENS-Cachan, ECARES, Middlesex, Montpellier, PSE and Zurich, and participants to the 2018 Behavioral Public Economic Theory workshop in Lille, the 2019 Behavioral economics workshop in Birmingham, the 2018 Psychological Game Theory workshop in Soletto, the 2016 ASFE conference in Cergy-Pontoise, the 2016 SIOE conference in Paris, the 2017 JMA (Le Mans) and the ESA European Meeting (Dijon) for their useful remarks on earlier versions of this paper. Jacquemet gratefully acknowledges funding from ANR-17-EURE-001.

## References

- Acemoglu, D., & Jackson, M. O. (2015). History, expectations, and leadership in the evolution of social norms. *Review of Economic Studies*, 82(2), 423–456.
- Acemoglu, D., & Jackson, M. O. (2017). Social norms and the enforcement of laws. *Journal of the European Economic Association*, 15(2), 245–295.
- Ali, S. N., & Bénabou, R. (2020). Image versus information: Changing societal norms and optimal privacy. *American Economic Journal: Microeconomics*, 12(3), 116–164.
- Azieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4), 1472–1503.
- Bell, R. M., & McCaffrey, D. F. (2002). Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology*, 28(2), 169–179.
- Benabou, R., & Tirole, J. (2011). *Laws and norms*. NBER WP 17579.
- Bérgolo, M., Ceni, R., Cruces, G., Giacobasso, M., & Pérez-Truglia, R. (2021). Tax audits as scarecrows. Evidence from a large-scale field experiment. *American Economic Journal: Economic Policy*, 2021, 1.
- Bowles, S. (2008). Policies designed for self-interested citizens may undermine the moral sentiments: Evidence from economic experiments. *Science*, 320(5883), 1605–1609.

- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3), 414–427.
- Dal Bó, E., & Dal Bó, P. (2014). ‘Do the right thing’: The effects of moral suasion on cooperation. *Journal of Public Economics*, 117, 28–38.
- Dal Bó, P., & Fréchette, G. R. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, 101(1), 411–429.
- Deffains, B., & Fluet, C. (2020). Social norms and legal design. *Journal of Law, Economics, and Organization*, 36(1), 139–169.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550.
- Duffy, J., & Fehr, D. (2018). Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. *Experimental Economics*, 21(3), 573–600.
- Dutta, R., Levine, D. K., & Modica, S. (2021). Interventions with sticky social norms: A critique. *Journal of the European Economic Association*, 2021, 1.
- Engl, F., Riedl, A., & Weber, R. (2021). Spillover effects of institutions on cooperative behavior. *Preferences, and Beliefs*, *American Economic Journal: Microeconomics*, 13(4), 261–299.
- Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5), 1611–1630.
- Fortin, B., Lacroix, G., & Villevall, M.-C. (2007). Tax evasion and social interactions. *Journal of Public Economics*, 91(11–12), 2089–2112.
- Friebel, G., & Schnedler, W. (2011). Team governance: Empowerment or hierarchical control. *Journal of Economic Behavior and Organization*, 78(1–2), 1–13.
- Galbiati, R., Henry, E., & Jacquemet, N. (2018). Dynamic effects of enforcement on cooperation. *Proceedings of the National Academy of Sciences*, 115(49), 12425–12428.
- Galbiati, R., Schlag, K. H., & Van Der Weele, J. J. (2013). Sanctions that signal: An experiment. *Journal of Economic Behavior and Organization*, 94, 34–51.
- Galizzi, M. M., & Whitmarsh, L. E. (2019). How to measure behavioural spillovers? A methodological review and checklist. *Frontiers in Psychology*, 10, 342.
- Gill, D., & Rosokha, Y. (2020). *Beliefs, learning, and personality in the indefinitely repeated Prisoner's dilemma*. CAGE online working paper series 489.
- Kartal, M., & Müller, W. (2018). *A new approach to the analysis of cooperation under the shadow of the future: Theory and experimental evidence*. Working paper.
- Nowak, M. A., & Roch, S. (2007). Upstream reciprocity and the evolution of gratitude. *Proceedings of the Royal Society B-Biological Sciences*, 274(1610), 605–609.
- Peysakhovich, A., & Rand, D. G. (2016). Habits of virtue: Creating norms of cooperation and defection in the laboratory. *Management Science*, 62(3), 631–647.
- Sliwka, D. (2007). Trust as a signal of a social norm and the hidden costs of incentive schemes. *American Economic Review*, 97(3), 999–1012.
- Van Der Weele, J. J. (2009). The signaling power of sanctions in social dilemmas. *Journal of Law, Economics, and Organization*, 28(1), 103–126.
- Zasu, Y. (2007). Sanctions by social norms and the law: Substitutes or complements? *The Journal of Legal Studies*, 36(2), 379–396.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.