| aptara | JPET | jpet12090 | Dispatch: February 20, 2014 | CE: AFL |
| --- | --- | --- | --- | --- |
| | Journal | MSP No. | No. of pages: 27 | PE: Ajay Patel |

# Coordinated Punishment and the Evolution of Cooperation

GONZALO OLCINA
*University of Valencia*

Q1

VICENTE CALABUIG
*University of Valencia*

## Abstract

In this paper, we analyze a team trust game with coordinated punishment of the allocator by investors and where there is also a final stage of peer punishment. We study the effect of punishment on the reward and the investment decisions, when the effectiveness and cost of coordinated punishment depend on the number of investors adhering to this activity. The interaction takes place in an overlapping-generations model with heterogeneous preferences and incomplete information. The only long-run outcomes of the dynamics are either a fully cooperative culture (FCC) with high levels of trust and cooperation and fair returns or a non-cooperative culture with no cooperation at all. The basin of attraction of the FCC is larger, the higher the institutional capacity of coordinated punishment, the higher the level of peer pressure and the smaller the individual cost of coordinated punishment.

Gonzalo Olcina, University of Valencia, ERJ-CES (Gonzalo.Olcina@uv.es). Vicente Calabuig, University of Valencia, ERJ-CES.

The authors thank an associate editor and a referee for helpful comments. This paper has also benefited from the comments of participants in the 14th International Symposium on Dynamic Games and Applications (Banff, 2010), Spain, Italy, and Netherland Meeting on Game Theory (Palermo, 2010), 2nd Valencia-Louvain Workshop on Game Theory (Louvain Ia Neuve, 2010), Public Economic Theory PET'11 (Bloomington, Indiana, 2011) and Fifth International Conference on Game Theory and Management (Saint Petersburg, 2011). The authors acknowledge financial support from the Spanish Ministry of Education and Science project EC02008-06035/ECON.

Received July 28, 2011; Accepted December 18, 2012.



*Journal of Public Economic Theory*, 00 (0), 2014, pp. 1–27.

## 1. Introduction

A well-known salient feature of modern market economies is the huge quantity of mutually beneficial transactions that take place regularly in one-shot and anonymous interactions. Field and laboratory experiments show significant levels of trust and cooperative behavior in this class of games even when cooperative behavior is costly.

Both experimental evidence[1] and recent theoretical work[2] from anthropology and evolutionary theory show that individually costly cooperative behavior can be sustained by costly punishment. However, from our point of view there are two important weaknesses in most of the existing models on punishment and in the experimental work performed on this issue. On the one hand, current models and experiments usually assume that punishment is carried out on an individual basis, that is, it is uncoordinated. Nevertheless, this way of modelling punishment is quite unrealistic because most of the punishment exerted in real-life situations is coordinated. For instance, a minimum number of punishers is needed in order to obtain effective punishment in a strike or in a boycott.

On the other hand, the other weakness in most of the experimental literature on the role of punishment is that it has focused on symmetric team or group situations such as public goods games. But there have been very few studies on the impact of punishment on asymmetric economic games based on specialization and on the division of labor such as, for instance, the principal-agent relationship, the hold-up game, or in general any sequential transaction between a seller and a buyer.

A well-known social dilemma that captures these asymmetric economic games in a bilateral situation is the trust or investment game. One player (the investor) has the option of investing or not investing in a project which is administered by the other player (the allocator). Investing results in a higher joint surplus, but the allocator controls the proceeds of investment.

Many real-life economic situations are trust games with a team of investors. Moreover, punishment itself is also a team decision problem. The investors' capacity for punishment in this situation is endogenous, depending on the number of investors adhering to this activity. A prominent example of what we will denote as a team trust game appears in the labor market. In many employment relations, a group of employees is hired by a single employer (the firm). The labor contract in these cases is highly incomplete and it usually assigns significant authority to the employer. This asymmetric distribution of decision rights puts the other side, the employees, in danger of being exploited, leading to inefficiency if they refuse to cooperate.

---

[1] See for example, Fehr and Gächter (2000, 2002), Gächter, Renner, and Sefton (2009), Yamagishi (1986), Hauert *et al.* (2007).

[2] See for example Henrich *et al.* (2006) and Boyd *et al.* (2003).

In this paper, we analyze a team trust game with coordinated punishment of the allocator by the investors and where we also add a final stage of peer punishment. In addition to institutions and the law, the ability of different groups to overcome the collective action problem is at the core of almost any form of coordinated punishment and seems crucial for the effectiveness of punishment. In the previous labor market example punishing the firm is very costly and highly ineffective for an individual or a small group of workers. However, if the workers succeed in reaching a threshold in the number of punishers, the damage inflicted to the firm can be very large and the individual cost of coordinated punishment can be relatively low.

The empirical literature on collective action[3] agrees on two important factors that affect the likelihood of successful collective action: the possibility of peer punishment and the heterogeneity of preferences in the population.

Regarding the first factor, notice that punishment itself is a public good among investors and it is also subject to free-riding behavior. Peer punishment or peer pressure is targeted at those individuals who free-ride in the phase of coordinated punishment of the allocator. Concerning the heterogeneity of preferences, we assume that there are two types of investors: selfish individuals and conditional punishers or reciprocators. The former are only motivated by their absolute material payoff. The latter are willing to punish an unfair return offered by the allocator provided the cost of punishing is low enough and/or to punish the free-riding behavior of their teammate. There are also two types of allocators: selfish (profit-maximizers) and fair-minded allocators who always set a fair return.

To assume a heterogeneity of preferences is quite standard nowadays, but our main assumption is that preferences are endogenous, that is the distribution of preferences in both populations evolves over time. Different forces govern the evolution of preferences in both populations. The dynamics of the allocator population is driven by market forces: profits. But the dynamics in the investor population is governed by a cultural transmission process that combines intentional and costly parental (direct) transmission with influence from society at large. If we keep in mind the example of firms and workers, our assumptions on the dynamics that governs each population seems a good approximation to actual societies.

We are especially interested in the influence of punishment institutions on the long-run distribution of preferences and behavior, particularly in the punishment coordination problem. In our model, these institutions are on the one hand the capacity to collectively punish the allocator and the cost of coordinated punishment and on the other hand the level of peer pressure. However, it is important to notice that for the punishment to be effective, not only laws and institutions (exogenous to the individual) are needed but

[3] See for example, Ostrom (1990, 2000)

jpet12090 W3G-jpet.cls February 20, 2014 13:8





also the willingness of these individuals to incur the costs of implementing the punishment.

We present an overlapping-generations dynamic model to analyze this team trust game with coordinated punishment. We denote as culture any stable steady state of the dynamics where the same equilibrium of the team trust game is played.

Our main results are the following. Although some other different equilibria can appear in the short-run where punishment is observed and/or unfair returns are offered and are not punished, the only long-run outcomes of our dynamics are either a fully cooperative culture (FCC) with high levels of trust and cooperation and fair returns or a non-cooperative culture (NCC) with no cooperation at all. In the FCC, cooperation is achieved under the credible threat of effective coordinated punishment. Precisely because of that there is no punishment observed in equilibrium. The credibility of punishment is supported by a relatively high proportion of reciprocators in the investor population. By contrast, in the NCC the threat of coordinated punishment is not credible at all because there is a low proportion of punishers in the investor population. As a result there are low levels of cooperation and efficiency.

The FCC is only feasible for high values of the institutional capacity of coordinated punishment, and therefore cooperation evolves only if the law allows for a sufficiently high punishment capacity in society.

But the law and institutions to punish opportunistic allocators are not enough. The main determinant of the basin of attraction of the FCC is a sufficiently high level of peer pressure relative to the individual cost of coordinated punishment targeted at the allocator. High peer pressure and, therefore, institutions that favor it have a strong impact on the feasibility of effective coordinated punishment and consequently on the levels of cooperation and efficiency. This result can explain the importance of belonging to organizations or clubs where peer pressure is more easily exerted, like a union for example. But this is not the only example. For instance, belonging to a community or a gang increases the damage inflicted by the group on the free-rider.

The intuition behind our main results is the following. Strong punishment institutions, related to coordinated punishment and to peer punishment, increase the effectiveness of punishment in the short run. But they also increase the incentives to socialize on preferences that display negative reciprocity. This in turn will increase the effectiveness of punishment of future generations because of the presence of a larger proportion of punishers in the population. This might happen both because it increases the probability of having a punisher as a teammate and also because the individual expected cost of coordinated punishment diminishes. For a sufficiently high proportion of punishers, and provided the level of peer pressure is high enough, even selfish investors are willing to punish unfair return offers by the allocator. Hence, the credibility of the threat of punishment is the highest, and provided the capacity of punishment of the team is high enough, both types of allocator prefer to set fair returns. If society reaches this situation, both types of investor lose their incentives to actively socialize their children. Society has reached a FCC.

However, if society starts in a distribution with a low proportion of punishers, the above logic works exactly in the opposite direction and society will get stuck in a very inefficient outcome.

Our paper is related to two important strands of literature. First, the experimental analysis of so called altruistic punishment that starts with Fehr and Gachter (2000, 2002) and produces an impressive amount of evidence (see for example, Falk, Fehr, and Fischbacher 2005). All this evidence poses an important question for theoretical literature: how altruistic punishment can evolve in a large society where repeated game effects are negligible. This issue has been addressed from an evolutionary dynamics approach (see for instance, Sigmund *et al.* 2010). It would take us too long to review the vast experimental and evolutionary literature on punishment. To the best of our knowledge, the only work on coordinated punishment is by Boyd, Gintis, and Bowles (2010), who analyze a public goods game. Punishment is coordinated in the sense that it is contingent on the number of others predisposed to participate, and it shows increasing returns of scale (the individual cost of punishment decreases at an increasing rate with the number of punishers). The main difference with our work, apart from the fact that we analyze a team trust game, is that those authors assume that punishment is equally effective whatever the number of participants is. Instead, we assume that punishment is effective only if a minimum number of individuals participate.

Finally, our paper is closely related to the literature on cultural transmission and socialization (Bisin and Verdier 2000, 2001), and more particularly, to the work on the endogenous determination of preferences and its interaction with institutions. For instance, Huck and Kosfeld (2007) in an evolutionary model analyze how what they call weak institutions interact with preferences for punishment. As in our approach, institutions and the law are only effective if individuals are willing to engage in an individually costly implementation of these tools. Our paper differs from theirs in three main aspects: they analyze a public goods game, punishment is not coordinated, and they use a replicator-like dynamics.

The paper is organized as follows: In the next section we present the model. In Section 3, we introduce social preferences. In Section 4, we show the punishment and rewarding policy of the players. In Section 5, we compute the equilibria of the team trust game played by each generation. In Section 6, we present the dynamics of the model. In Section 7, we obtain the cultures in the long run. And finally Section 8 concludes.

## 2. The Team Trust Game with Coordinated Punishment

We consider a strategic situation in which a team of investors, composed of two players randomly drawn from a continuum of investors of mass 2, is matched with an allocator, randomly drawn from a continuum of allocators of mass 1, to play the following sequential game.

In the first stage, called the investment phase, both investors have to decide simultaneously and independently whether to invest in a project (action $I$) or not (action $NI$). If both investors choose to participate in the project, a joint surplus of size 2 is produced. Otherwise, if just one or both investors decide not to invest, no surplus is produced. In this latter case, we assume that the game ends and all players obtain a payoff of zero. We suppose that the gross gain per investor is 1 and that investment has a cost $c \in (0, 1/2)$ and that not investing is costless.

In the second stage, the rewarding phase, the allocator (she), after observing a surplus of size 2, has to set a payoff $b$ to each investor (he), where $0 \leq b \leq 1$. As we are interested in symmetric outcomes, we will assume that the allocator will pay the same reward $b$ to both investors.

In the third stage, the coordinated punishment phase, the investors have to decide simultaneously whether to punish the allocator (action $p$) or not to punish (action $np$). Only if both investors choose to punish, a proportion $\lambda$ of the payoff obtained by the allocator is destroyed, where $0 < \lambda \leq 1$. But if just one or neither of them decides to punish, then there is no surplus destruction. Therefore, punishment is only effective if both investors choose to punish.

We assume that choosing to punish is costly, but its cost depends on the number of investors that adhere to this activity. In particular, we suppose that if only one member decides to punish, he has to bear all the cost $z > 0$ of the (ineffective) punishment. But if both members choose to punish, the individual cost of (effective) punishment reduces to $z/2$. This simple model with only two actions captures a crucial feature of coordinated punishment: the existence of a threshold on the number of participants in the punishment activity that affects in a discontinuous way both its effectiveness and its individual cost. Only if the team succeeds in reaching this threshold, the damage caused to the allocator can be large enough and the individual cost of punishing can be lower than the cost of ineffective punishment.

Finally, the fourth stage is the peer punishment phase. If the coordinated punishment of the allocator has not succeeded because of the defection of a teammate, then the nondefecting investor now has the option of punishing the defecting mate at some cost. We assume that this peer punishment creates a loss in the utility of the punished mate of size $\gamma$.

Some comments on the institutional parameters that characterize the coordinated punishment of the allocator and peer punishment are in order. In a labor market context, $\lambda$ would be the punishment that the team of workers could inflict on the firm. This depends on the workers' capacity for money burning (sabotage, strikes, . . . ), which in turn might depend on the workers' degree of unionization, their ability to organize collectively, their legal rights in society, etc … It might also differ across different types of jobs depending on the strategic position of the worker in the production process. Parameter $\lambda$ can also be interpreted as the maximum punitive sanction that the legal system provides an agent with in order to punish the opportunistic behavior of the other party when the punisher is not able to recover the full cost of his investment (see Dufwenberg, Smith, and Van Essen 2011).

Concerning the peer punishment phase, parameter $\gamma$ is our measure of the level of peer pressure. We think that it captures a realistic feature of labor markets in which some workers punish the strike-breaking behavior of other coworkers and this is an important constraint of the behavior of many organizations (unions, communities, clubs, churches).

Suppose now that all players have self-regarding preferences and there is complete information. We can obtain the subgame perfect equilibrium solving the game by backward induction. In the last subgame, selfish investors will never punish either individually or collectively because it is costly and does not increase their payoff. Given that the allocator will not be punished, she will offer the investors a reward $b = 0$ and therefore the optimal action for them will be to choose not to invest. This is a very inefficient outcome in which all players obtain a payoff of zero.

In this paper, we will assume that there is a heterogeneity of preferences and that in addition to self-regarding people, there is also a significant fraction of the population that exhibits social preferences. In the next section, we introduce this type of preference.

## 3. Social Preferences: Reciprocal Altruism

Overwhelming evidence generated by laboratory experiments and also everyday experience, suggests that motives of fairness and reciprocity affect the behavior of many people. By reciprocity we mean the willingness to reward friendly behavior and the willingness to punish hostile behavior.

In each period $t$, there is a certain proportion $q_t$ of investors with reciprocal preferences in the population of investors and a remaining proportion $(1 - q_t)$ of individuals with selfish preferences. We suppose that reciprocal investors (reciprocators) are willing to punish "unfair" offers provided that the cost $z$ is low enough. This kind of player will also punish a teammate who has failed to punish the allocator for an unfair reward.

Reciprocal investors are punishers because they are concerned not only with their monetary payoff but they also aspire to get a fair return compared to the allocator's payoff, and hence any return smaller than what is considered a fair reward will generate disutility for them. We assume that the fair return[4] is $b = 1/2$ and that the disutility derived from getting a reward

[4] Here, we assume that a fair reward is $b = 1/2$, but the results do not change qualitatively if we allow for another "fair" or aspiration reward smaller or greater than $1/2$.

smaller than 1/2 is proportional (captured by the parameter $\alpha \geq 1$) to the distance between this fair reward and a smaller actual reward offered by the allocator. This is similar to the inequity aversion preferences of Fehr and Schmidt (1999), when players face a disadvantageous inequality. For example, if both investors decide to invest in the project, and they also decide to coordinate in punishing the allocator, the utility of a reciprocal investor is given by: $(b - z/2 - c) - \alpha[(1-b)(1-\lambda) - b]$ for any $b < 1/2$.

On the other hand, in each period $t$ there is a proportion of "fair-minded" agents ($p_t$) in the population of allocators and a remaining proportion $(1 - p_t)$ of profit maximizers. The fair-minded allocators are very generous in compensating the team of investors. Namely, setting a reward of $b = 1/2$ to each investor is a dominant action for them.

Note that players do not know the true type of player with whom they are matched in period $t$. In particular, the allocator does not know the true composition of the team and the members of the team do not know either the type of his teammate or the type of allocator. However, we will assume that the preferences distribution $q_t$ or $p_t$ in both groups are common knowledge.

## 4. The Punishment and Rewarding Policy

In this section, we will begin to solve the sequential game by backward induction. First of all, we will make three assumptions about the relationship among the parameters that characterize the punishing institutions. These assumptions guarantee that both types of punishment—coordinated and peer punishment—are chosen at least under some circumstances.

ASSUMPTION 1: $z \leq \alpha\lambda$.

This is a very straightforward assumption which simply states that reciprocators are conditional punishers. The reason is that by successfully punishing the allocator, a reciprocator reduces inequality with respect to the latter and this positive effect on her utility, $\alpha\lambda$, more than compensates for the reduction in his material payoff $z$. If this assumption does not hold, then there will be no difference between the behavior of a selfish and a reciprocal investor and the analysis will be of no interest.

ASSUMPTION 2: $z/2 < \gamma$.

Assumption 2 states that peer pressure is effective, at least under some conditions. The individual cost of a successful coordinated punishment of the allocator is strictly smaller than the damage inflicted by peer punishment. Otherwise, a selfish investor will never participate in the coordinated punishment of the allocator, even if he knows for sure that his teammate is a reciprocal investor who is going to punish him.

ASSUMPTION 3: $\lambda \geq 0.5$.

Assumption 3 states that coordinated punishment has a sufficient impact on the behavior of the selfish allocator. If $\lambda < 0.5$, the allocator would prefer to offer a reward of zero rather than a "fair" reward $(b = 1/2)$, even if she knew for sure that the team was going to punish her. In Section 8 we will comment on how the results change when some of these assumptions do not hold.

### 4.1. Peer Punishment in the Team

In the last stage of the game, each team member has to decide whether to use peer punishment against the other teammate, at a positive cost. A selfish investor will never exert costly peer punishment because it does not increase his payoff, and there is no additional stage to punish non-peer-punishers.

On the other hand, a reciprocal investor dislikes disadvantageous inequality with his teammate and will punish a teammate who has free-ridden in the previous coordinated punishment stage. This will hold whenever the damage inflicted on the defector $\gamma$ is sufficiently high compared with the cost of doing so, because the action of peer punishment reduces inequality with the defecting teammate.[5]

Moreover, a reciprocal investor that has not exerted coordinated punishment will not use peer punishment either. The reason is that if the teammate has not punished the allocator, then there is no inequality between the members of the team. And if his teammate has chosen the action of punishing the allocator, as the coordinated punishment has not been effective, then the defecting investor would get advantageous inequality compared with his teammate.

In order to reduce the complexity of the analysis we do not incorporate the term concerning inequality with the teammate into the utility function of the reciprocal investor and we normalize the cost of peer punishment to zero.[6] Instead, in the rest of the paper we assume the result derived from the previous discussion: only a reciprocal investor who has chosen the action of punishing the allocator and finds out that his teammate has been a defector will choose peer punishment.

### 4.2. The Coordinated Punishment Subgame

In this section, we derive the Bayesian equilibria of the coordinated punishment phase of the sequential game played in each period. We will characterize team behavior in this subgame, anticipating the behavior of players in the peer punishment phase described in Section 4.1.

[5] Let $\theta$ be the cost of peer punishment then, the condition is $\theta \leq (\alpha/(1+\alpha))\gamma$.
[6] This normalization does not affect qualitatively any of the results.

We denote $\mu$ as the updated probability of facing a reciprocal type of investor after a history in which both investors have chosen to invest, that is, $\mu = Prob(r/(I, I))$, where $r$ stands for reciprocal type.

Any coordinated punishment subgame is characterized by a belief $\mu$ and a reward $b$ set by the allocator. Therefore, we will denote this subgame by $CP(\mu, b)$. We represent the (symmetric) Bayesian Nash equilibria (BNE) of this subgame by profiles $(x, y)$, where the first term represents the action of the reciprocator type and the second the action of the selfish type.

Notice that if $b \geq 1/2$, the unique BNE of $CP(\mu, b)$ for any $\mu$ is $(np, np)$, since no type of investor uses any sort of punishment.

The following proposition shows the solution of the $CP(\mu, b)$ for "unfair" rewards.

PROPOSITION 1: *If assumptions 1 and 2 hold and $b < 1/2$, the solution of any subgame $CP(\mu, b)$ is:*

*(i)* *(np,np) for any* $\mu < \mu^*(b) = \frac{z}{z/2+\alpha\lambda(1-b)+\gamma}$,

*(ii)* *(p,np) for any* $\mu \in [\mu^*(b), \overline{\mu} = \frac{z}{2\gamma})$,

*(iii)* *(p,p) for any* $\mu \in [\overline{\mu} = \frac{z}{2\gamma}, 1]$.

*Proof*: See Appendix. ■

Notice that, for $\mu < \mu^*(b)$, no type of investor is willing to punish. In fact the profile $(np, np)$ is the unique BNE.

If $\mu \in [\mu^*(b), \overline{\mu})$, that is, if the proportion of reciprocators is high enough but not very high, only the reciprocators punish while the selfish members of the team will not punish. This bound $\mu^*(b; \alpha, \lambda, z, \gamma)$ is increasing in $z$ and $b$ and decreasing in $\lambda$, $\alpha$, and $\gamma$.

A selfish investor will participate in the punishment of the allocator when $(b - z/2) \geq b - \mu\gamma$, that is, if $\mu \geq \frac{z}{2\gamma}$. Hence, the selfish investor will also punish the allocator as the probability of having a reciprocal mate is high enough. Therefore, for $\mu \geq \overline{\mu}$, both types of investors will punish the allocator.

Notice that there is a multiplicity of equilibria in this subgame and we have made an equilibrium selection, choosing for each $\mu$ the equilibrium with the highest probability of punishment. In particular, the profile $(np, np)$ is a BNE for all $\mu$. If we select this equilibrium for all $\mu$, we will obtain the inefficient outcome of no cooperation as if the game were played without punishment. Clearly, the analysis would be uninteresting. Note also that the profile $(p, np)$ is a BNE for $\mu \in [\mu^*(b), \widetilde{\mu} = \frac{z}{z/2+\gamma})$ and the profile $(p, p)$ is a BNE for $\mu \in [\overline{\mu}, 1]$. Therefore, as $\overline{\mu} < \widetilde{\mu}$, for $\mu \in [\overline{\mu}, \widetilde{\mu}]$, there are two BNE, but we assume that the BNE *(p,p)* is selected. We will discuss the effects of selecting the equilibrium $(p, np)$ in the next section.

We now turn to the optimal rewarding policy of selfish allocators.

### 4.3. The Rewarding Policy of a Selfish Allocator

Notice first that the return policy of a fair-minded allocator does not change when there is incomplete information. However, the rewarding policy of a selfish allocator is indeed affected by the proportion of reciprocators in the investor population. From now on we will denote the offer of the selfish allocator by $b_s$ and the offer of the fair-minded allocator by $b_f$.

It is obvious that when the allocator sets $b = 1/2$, she will not be punished by any type of investor and her payoff will be 1.

The optimal return policy of the selfish allocator will depend, basically, on the comparison between the expected cost of being punished and the cost of avoiding punishment, that is, the minimum reward at which no type of investor will punish. The expected cost of punishment will depend on the proportion of reciprocators in the investor population $\mu$ and the equilibrium played in the coordinated punishment game. In other words, the selfish allocator has two options: (i) to offer a low return $b_s$ such that there is punishment, and it is easily proven that in that case, the best return is to offer $b_s = 0$, or (ii) to offer a sufficiently generous reward $b_s > 0$ to avoid the coordinated punishment of the team.

We describe the optimal reward policy of the selfish allocator with the following lemmata:

LEMMA 1: *For any* $\mu < \mu^*(0) = \frac{z}{\alpha\lambda+\gamma+z/2}$, *the selfish allocator will set* $b_s = 0$.

*Proof*: Note that for any given $\widehat{\mu} < \mu^*(0)$, then $\widehat{\mu} < \mu^*(b)$, $\forall b < 1/2$, which implies that the BNE of $CP(\mu, b)$ is $(np, np)$. Thus, if the selfish allocator offers $b_s = 0$, she will not be punished and her payoff will be 2. ■

Note that $\mu^*(0)$ is the maximum proportion of reciprocators in the investor population such that both types of investors do not punish at $b_s = 0$. When the proportion of reciprocators is so low that no type of investor punishes, the expected cost of being punished is zero and thus the selfish allocator prefers to offer the lowest return $b_s = 0$.

LEMMA 2: *For any* $\mu$ *such that* $\mu^*(0) < \mu < \mu^*(0.5) = \frac{z}{(\alpha/2)\lambda+\gamma+z/2}$, *there is a reward* $\widehat{b}(\mu)$, *such that* $\mu^*(\widehat{b}) = \mu$, *where* $0 < \widehat{b}(\mu) = \frac{q(\alpha\lambda+\gamma+z/2)-z}{q\alpha\lambda} < 1/2$. *The optimal reward policy of the selfish allocator is unique and it will be one of the following:* $b_s = 0$ *or* $b_s = \widehat{b}(\mu)$.

*Proof*: See Appendix. ■

Notice that $b_s = \widehat{b}(\mu)$ is the minimal reward for a given $\mu$ such that reciprocators do not punish. In this case, the allocator has to choose between setting $b_s = \widehat{b}(\mu)$ and avoiding punishment or setting $b_s = 0$ and being punished only by the reciprocators with an expected cost of $\mu^2\lambda$. Recall that for this range of values of $\mu$[7] only the reciprocators choose to punish any offer $b < 1/2$.

LEMMA 3: *For any $\mu$ such that $\overline{\mu} = \frac{z}{2\gamma} > \mu > \mu^*(0.5)$, the selfish allocator sets $b_s = 0$ if $\mu < \mu' = \frac{1}{\sqrt{2\lambda}}$ and sets $b_s = 1/2$ if $\mu \geq \mu'$.*

For this range of values of $\mu$ the reciprocators choose to punish any offer $b < 1/2$ and the only way to avoid punishment is to offer $b = 1/2$. Then for the selfish allocator the cost of avoiding punishment is 1 (offering $b_s = 1/2$) and the cost of being punished is $2\mu^2\lambda$ (offering $b_s = 0$). Therefore, the optimal reward policy depends on a critical value $\mu'$, which comes from comparing both of the previous expressions. Notice that $\mu' > 1$ only when $\lambda < 0.5$. Therefore, in this case, the optimal reward policy is to set $b_s = 0$ for all $\mu$.

LEMMA 4: *For all $\mu \geq \overline{\mu} = \frac{z}{2\gamma}$ and $\lambda \geq 0.5$, the optimal return for the selfish allocator is to set $b_s = 1/2$.*

This result is due to the fact that in this range of values of $\mu$, the BNE of the coordinated punishment phase is $(p, p)$, by which both types of investors do punish if the allocator offers $b_s < 1/2$. Therefore, the allocator has a cost of $\lambda$ of being punished if she offers $b_s < 1/2$, while the cost of avoiding punishment is $1/2$ by setting $b_s = 1/2$.

Recall that under Assumption 3, $\lambda \geq 0.5$. However, if $\lambda < 0.5$, the optimal reward policy would be $b_s = 0$.[8]

## 5. Equilibria within a Generation

We are now ready to obtain the equilibria of the whole team trust game. To begin with, we characterize the efficient or cooperative equilibria. All types of investors choose to invest in the project, there is no punishment in equilibrium and, therefore, there is no surplus destruction. In these pooling equilibria, $q = \mu = Prob(r/(I, I))$. The difference among the various equilibria that might exist is the reward chosen by the selfish allocator.

PROPOSITION 2: ***The fully cooperative equilibrium (FCE) with $b_s = 1/2$.***

*If $\lambda \geq 0.5$ and for $q_t \geq \min\{\overline{q} = \frac{z}{2\gamma}, q' = \frac{1}{\sqrt{2\lambda}}\}$ and any $p_t$, there is a pooling equilibrium in which both types of investors choose to invest, both types of allocators set $b = 1/2$ and no punishment is observed in equilibrium.*

---

[7] The particular values of $\mu$ for which the first policy or the second one is optimal, depends on the particular location of the roots of the cubic equation $z = \mu(\alpha\lambda + \gamma + z/2) - \alpha\lambda^2\mu^3$, as it is explained in the Appendix.

[8] If we select in the interval $[\overline{\mu}, \widetilde{\mu}]$ the BNE $(p, np)$ instead of the BNE $(p, p)$, lemmata 3 and 4 still hold simply replacing $\overline{\mu}$ with $\widetilde{\mu}$.

*Proof*: Suppose $\lambda \geq 0.5$, by Lemma 4 the selfish allocator will set $b_s = 1/2$ if $q_t \geq \overline{q}$ and by Lemma 3, if $q_t \geq q'$. Therefore, this pooling equilibrium exists whenever $q_t \geq \min(\overline{q} = \frac{z}{2\gamma}, q' = \frac{1}{\sqrt{2\lambda}})$. The payoff for any team member is $1/2 - c$ and any type of allocator gets a payoff of 1. By Proposition 1, there is no punishment in equilibrium. ■

This equilibrium is supported by the credible threat of coordinated punishment if the allocator sets a return of $b < 1/2$. This can only happen with a relatively high fraction of reciprocal investors. Given this high proportion, even the selfish investor will punish unfair offers, fearing peer punishment. As a consequence, the selfish allocator will also set $b_s = 1/2$, there is no punishment and both types of investors invest. This result is driven by the fact that the high number of reciprocators in this equilibrium increases the probability of having a reciprocal teammate. Therefore, the cost of coordinated punishment is reduced to $z/2$, and the possibility of suffering peer punishment in the case of a defection increases.

Notice that the FCE does not exist for $\lambda < 1/2$. This is because selfish allocators would choose $b_s = 0$ and there would be punishment.

PROPOSITION 3: ***The cooperative equilibrium with*** $b_s = \widehat{b} > 0$. *If* $q^*(0.5) \geq q_t \geq q^*(0)$ *and* $p_t \geq p'(q, \lambda, \alpha, c, z)$, *there is a pooling equilibrium in which both types of investors choose to invest, where* $q^*(0.5) = \frac{z}{(\alpha/2)\lambda + \gamma + z/2}$, $q^*(0) = \frac{z}{\alpha\lambda + \gamma + z/2}$, $p' = \frac{c - (\widehat{b} - \alpha(1-2\widehat{b}))}{0.5 - (\widehat{b} - \alpha(1-2\widehat{b}))}$, *and* $\widehat{b} = \frac{q(\alpha\lambda + \gamma + z/2) - z}{q\alpha\lambda}$. *Profit maximizing allocators set* $b_s = \widehat{b} < 1/2$ *and fair-minded allocators set* $b_f = 1/2$. *No punishment is observed in equilibrium.*

*Proof*: If $q_t \in [q^*(0), q^*(0.5)]$ the selfish allocators set $b_s = \widehat{b} < 1/2$, by Lemma 2. A selfish investor will invest when the following condition holds: $p_t(0.5) + (1 - p_t)\widehat{b} - c \geq 0$. That is, if $p_t \geq \frac{c - \widehat{b}}{0.5 - \widehat{b}}$.

The condition for a reciprocal investor is: $p_t(0.5) + (1 - p_t)[\widehat{b} - \alpha(1 - 2\widehat{b}) - c \geq 0$. That is, if $p_t \geq \frac{c - (\widehat{b} - \alpha(1-2\widehat{b}))}{0.5 - (\widehat{b} - \alpha(1-2\widehat{b}))}$. This latter is the binding condition. The payoff of the fair-minded allocator is 1 and the expected payoff of the selfish allocator is $2(1 - \widehat{b})$. By Proposition 1, there is no punishment in equilibrium. ■

In this equilibrium, the selfish allocator offers $b_s = \widehat{b} < 1/2$, the minimal reward such that reciprocators do not punish. This reward depends on the proportion of reciprocal investors in the population. It is worth choosing investment for both types of investors if the proportion of fair-minded allocators is high enough, in particular, if it is greater than a critical value $p'(q, \lambda, \alpha, c, z)$.

PROPOSITION 4: ***The cooperative equilibrium with*** $b_s = 0$. *If* $q_t \leq q^*(0)$ *and* $p_t \geq p''(c, \alpha) = \frac{\alpha + c}{\alpha + 0.5}$, *there is a pooling equilibrium in which both types of investors choose to invest. Profit maximizing allocators set* $b_s = 0$ *and fair-minded allocators set* $b_f = 1/2$. *No punishment is observed in equilibrium.*

*Proof*: If $q_t < q^*(0)$ the selfish allocators set $b_s = 0$, by Lemma 1. A selfish investor will invest when the following condition holds: $p_t(0.5) - c \geq 0$. That is, if $p_t \geq 2c$. The condition for a reciprocal investor is: $p_t(0.5) + (1 - p_t)(-\alpha) - c \geq 0$. That is, if $p_t \geq \frac{\alpha + c}{\alpha + 0.5}$. The latter is the binding condition.

The payoff of the fair-minded allocator is 1 and the payoff of the selfish allocator is 2. ■

In this equilibrium, the proportion of reciprocators is so low that it is not worthwhile for them to punish for any reward and the selfish allocator anticipating this behavior sets the lowest possible return, $b_s = 0$. However, both types of investors decide to invest due to the very high proportion of fair-minded allocators.

Next, we switch to the inefficient or non-cooperative equilibria of the team trust game. In these equilibria either one or both types of investors do not invest or, even if both choose to invest, there is punishment and thus surplus destruction, with positive probability.

PROPOSITION 5: ***The non-cooperative equilibrium (NCE).*** *For every* $q_t$ *and* $p_t$, *there is an inefficient pooling equilibrium in which both types of investors choose not to invest.*

The proof is straightforward and is left to the reader. Just notice that the equilibrium payoff for all types of players is 0 and there is no profitable deviation of investors.

Next, we characterize separating equilibria in which just one type of investor chooses the efficient action of investing.

PROPOSITION 6: ***The inefficient separating equilibrium (ISE).*** *For any* ($q_t$, $p_t$) *such that* $\frac{c + \alpha(1 - q_t)}{(1 - q_t)(\alpha + 0{,}5)} \geq p_t \geq \frac{2c}{1 - q_t}$, *there is an ISE in which the selfish investor chooses to invest in the project whereas the reciprocator chooses not to invest. Profit maximizing allocators set* $b_s = 0$ *and fair-minded allocators set* $b_f = 1/2$.

*Proof*: The incentive compatibility constraint for the selfish investor is $(1 - q_t)p_t(0.5) - c \geq 0$. That is, $p_t \geq \frac{2c}{1 - q_t}$. The constraint for a reciprocal investor is $0 \geq (1 - q_t)[p_t(0.5) + (1 - p_t)(-\alpha)] - c$. That is, $p_t \leq \frac{c + \alpha(1 - q_t)}{(1 - q_t)(\alpha + 0{,}5)}$. The set of pairs $(q_t, p_t)$ that satisfies both incentive compatibility constraints is not empty. The expected payoff of the selfish allocator is $2(1 - q_t)^2$, while the expected payoff of the fair-minded allocator is $(1 - q_t)^2$. ■

In this equilibrium, surplus is generated with probability $(1-q)^2$, which is the probability of the team being composed of two selfish investors. Note that, paradoxically, selfish team members choose the efficient action while reciprocators do not. Selfish investors invest because of the presence in society of a significant fraction of fair-minded allocators who pay high returns, and reciprocal investors choose not to participate in the project because of the presence of a significant fraction of selfish allocators who pay low returns. This explains why this equilibrium only exists for an intermediate range of values of $p$.

There is another separating equilibrium in which, in contrast to the previous case, the reciprocators invest and the selfish types do not. However, this equilibrium only exists for a degenerate distribution of preferences, $q = 2c$. We will not take into account this equilibrium in the main text because it is easily shown in the Appendix that it never constitutes a stable steady state of the dynamics, which we will introduce in Section 6.

Finally, there is another inefficient equilibrium in which although both types of investors choose to invest, there is punishment with positive probability.

PROPOSITION 7: ***The quasi-cooperative equilibrium with punishment.*** *If* $q_t \in [q_1, \min\{\overline{q}, q'\}]$, *and* $p_t \geq p'''(q, \lambda, \alpha, c, z)$, *there is a pooling equilibrium in which both types of investors choose to invest, where* $p''' = \frac{z+\alpha+c-q(z/2+\alpha\lambda)}{z+\alpha+0.5-q(z/2+\alpha\lambda)}$ *and* $q_1$ *is the smallest positive real root of the cubic equation* $z = q(\alpha\lambda + \gamma + z/2) - \alpha\lambda^2 q^3$. *Profit maximizing allocators set* $b_s = 0$ *and fair-minded allocators set* $b_f = 1/2$. *Only reciprocators punish the selfish allocators in equilibrium.*

*Proof*: See Appendix. ■

This equilibrium exists for a relatively high proportion of reciprocal investors who punish low rewards. However, there has to be a high proportion of fair-minded allocators that makes it profitable for both types of investors to choose to invest, despite the fact that the reciprocal investors will have to punish unfair rewards.

Notice that for some regions of $(p, q)$ there is a multiplicity of equilibria. We will assume that the NCE is only played in the region where it constitutes the unique equilibrium. The reason is that it is at least (weakly) Pareto dominated by any other equilibrium. As we will prove with the dynamic analysis which we introduce in the next section, in all the remaining cases our results do not depend on the particular equilibrium which is played in each period.

## 6. Dynamics of the Model

Our setting is a two-speed dynamic model. Changes in preferences are gradual over time, while changes in behavior are instantaneous to maintain equilibrium play. Therefore, in each period individuals coordinate in a perfect Bayesian equilibrium (PBE) of the team trust game and, assuming adaptive expectations, they believe that this equilibrium will be played by the next generation.

The dynamics in each population is governed by different forces. The evolution of the proportion of the different types of allocators is driven by market forces: the level of profits. However, the dynamic evolution of the distribution of preferences in the investor population is influenced by cultural motives, more precisely by an intergenerational transmission of preferences that, in turn, is affected by an intentional process of socialization, not exclusively driven by material payoffs.

### 6.1. The Dynamics of the Allocator Population

We assume that at the end of each period, allocators who follow the less profitable reward policy have a positive probability of being replaced by allocators with a reward policy that provides more profits. The probability of change is assumed to be an increasing function of the profit differences. Then the dynamic behavior of $p_t$ is given by the following difference equation: $\Delta p_t = p_t(1-p_t)\varphi[\Pi_t^f(p_t, q_t) - \Pi_t^s(p_t, q_t)]$, where $\Pi_t^f(p_t, q_t)$ and $\Pi_t^s(p_t, q_t)$ are the profits, in period t, for the fair-minded and the profit maximizing type of allocator, respectively. Notice that $\varphi$ is a positive constant which is low enough to have $p_t \in [0, 1]$. This is analogous to the replicator-dynamics and hence it is payoff-monotonic. This dynamics is not influenced by any kind of intergenerational transmission of preferences in the allocator population. The reason is that because of the usual motive of competition among firms, independently of the cultural traits of the managers, firms with lower rates of profits would be more likely to leave the market.

### 6.2. The Cultural Dynamics of the Investor Population

Preferences in the investor population are culturally transmitted according to an intergenerational transmission process. Children acquire preferences through the observation, imitation, and learning of cultural models prevailing in their social and cultural environment, that is, in their family and in their social group. The transmission of preferences which is the result of social interaction between generations is called cultural transmission. We will draw from the model of cultural transmission by Bisin and Verdier (2001), which is the economic version of the anthropological model by Cavalli-Sforza and Feldman (1981).

We consider overlapping-generations of investors who only live for two periods (as a young person and as an adult). In the first period, the investor is a child and is socialized to certain preferences. In the second period, the investor (as an adult with well-defined preferences) is randomly matched with an adult investor to form a team and play the team trust game with a randomly matched allocator. Also in this second period, the adult investor has one offspring[9] and has to make a (costly) decision regarding his child's education, trying to transmit his own preferences.

Therefore, the investor population will evolve according to a purposeful and costly socialization process that we describe next. Let $\tau^i \in [0, 1]$ be the educational effort made by an investor parent of type $i$ where $i \in \{s, r\}$ and $s$ denotes selfish and $r$ denotes reciprocator.

The socialization mechanism works as follows: Consider a parent with $i$ preferences. His child is first directly exposed to the parent's preferences and is socialized to these preferences with probability $\tau^i$ chosen by the parent (vertical transmission); if this direct socialization is not successful, with probability $1 - \tau^i$, he is socialized to the preferences of a role model picked at random from the investor population (oblique transmission).

The transition probabilities[10] $P^{ij}$, determined by this socialization mechanism, can be easily computed and then the dynamic evolution of the distribution of preferences can be obtained, which is given by the following equation on differences:[11] $\Delta q_t = q_t(1 - q_t)[\tau_t^r - \tau_t^s]$.

Direct transmission is also costly. Let $C(\tau^i)$ denote the cost of the educational effort $\tau^i$. While it is possible to obtain similar results with any increasing and convex cost function, we will assume, for simplicity, the following quadratic form $C(\tau^i) = (\tau^i)^2/2k$, with $k > 0$. Therefore, a parent of type $i$ chooses the educational effort $\tau^i \in [0, 1]$ at time $t$, which maximizes

$P_t^{ii}(\tau^i, q_t)V^{ii}(q_{t+1}^E) + P_t^{ij}(\tau^i, q_t)V^{ij}(q_{t+1}^E) - (\tau^i)^2/2k,$

where $V^{ij}$ is the utility to a parent with preferences $i$ if his child is of type $j$. Notice that utility $V^{ij}$ depends on $q_{t+1}^E$, which denotes the expectation about the proportion of reciprocal investors in the population in period $t + 1$. In this paper, we will assume that parents have adaptive or backward-looking expectations, believing that the proportion of reciprocal investors will be the same in the next period as in the current period, that is, $q_{t+1}^E = q_t$.

Direct transmission is justified because parents are altruistic toward their children. However, their socialization decisions are not based on the purely material payoff expected for their children, but on the payoff as perceived by the parents according to their own preferences. This is the notion of imperfect empathy. According to this notion, parents obtain a higher utility if their children share their preferences. Let us define $\Delta V^r = V^{rr} - V^{rs}$ and $\Delta V^s = V^{ss} - V^{sr}$. That is, $\Delta V^i$ is the net gain from socializing your child to your own preferences, or the cultural intolerance of parents with respect to cultural deviation from their own preferences.

---

[9] As is customary in this class of models, we will assume that reproduction is asexual, with one% per child, and thus the population remains constant.

[10] $P^{ij}$ denotes the probability of a child of a parent with preferences $i$ being socialized to preferences $j$.

[11] We relegate the particular details of this process to the Appendix.

Maximizing the above expression with respect to $\tau^i$, we get the following optimal education effort functions:

$$\tau^{r*}(q_t) = k\Delta V^r(q_t)(1-q_t).$$

$$\tau^{s*}(q_t) = k\Delta V^s(q_t)q_t.$$

Note that the optimal education effort functions of both types of parent depend (positively) on their level of cultural intolerance ($\Delta V^i$) and (negatively) on the proportion of their own type in the current preferences distribution in the population.

Substituting the optimal educational efforts into the differences equation that characterizes the dynamic behavior of $q_t$, we obtain:

$\Delta q_t = q_t(1-q_t)k[\Delta V^r(1-q_t) - \Delta V^s q_t]$.

This is the Bisin-Verdier cultural dynamics. Instead of material payoffs, levels of cultural intolerance are the main determinants that govern the dynamic evolution of the preferences distribution in the investor population.

Summing up, the joint dynamics of the preference distribution in both populations (allocators and investors) is determined by the dynamical system defined by the following two nonlinear differences equation system:

$$\begin{aligned} \Delta p_t &= p_t(1-p_t)\varphi\left[\Pi_t^f(q_t,p_t) - \Pi_t^s(q_t,p_t)\right], \\ \Delta q_t &= q_t(1-q_t)k[\Delta V^r(1-q_t) - \Delta V^s q_t]. \end{aligned} \tag{1}$$

## 7. Cultures in the Long Run

We adhere to the notion of culture, used by Rob and Zemsky (2002), as a stable or self-reproducing pattern of behavior and beliefs in a group or a society. Therefore, we identify it as a stable steady state of the preference dynamics.

DEFINITION 1: *A culture is any stable steady state of the dynamical system (1) where the same perfect Bayesian equilibrium of the team trust game is played.*

We will denote, for example, as a FCC any stable steady state of the dynamics where a FCE is played. A similar definition applies for the other equilibria of the game. Our model yields different long-run outcomes, cultures; some of them are efficient, that is, both types of investors invest and there is no punishment, and other cultures are inefficient because some of the previous conditions do not hold.

Our strategy will consist of analyzing whether the different PBE of the team trust game are "robust" under our dynamical system. By robust, we mean that the dynamics does not take the distribution of preferences out of the region of the space $(q, p)$ where the PBE exists.

Let us start by analysing the efficient (cooperative) PBE.

PROPOSITION 8: *If* $\lambda \geq 0.5$, *the only efficient culture is the FCC and it exists for any pair of distributions* $(q_t, p_t)$ *such that* $q_t \geq \min\{\overline{q} = \frac{z}{2\gamma}, q' = \frac{1}{\sqrt{2\lambda}}\}$.

*Proof*: See Appendix. ■

We give here a sketch of the proof and leave the rest of the details in the Appendix. First, we show that the FCE constitutes a culture, and second that the other two cooperative equilibria cannot be stable steady states of the dynamics, and hence they never become cultures.

The FCC is based on the FCE in which both types of investors invest and both types of allocators set $b = 1/2$. Therefore, $V^{ik} = 1/2 - c$ for both types of parents of investors. Hence, their levels of cultural intolerance, and consequently their optimal educational efforts, are zero. There are no incentives at all for socialization. Thus, the distribution of preferences in the investor population will remain unchanged, that is, $q_{t+1} = q_t$.

Concerning the dynamic evolution in the allocator population, note that the levels of profits of both types of allocators are the same $\Pi_t^f = \Pi_t^s = 1$, and thus the preference distribution in the population of allocators will also remain constant. Concluding, any pair of preference distributions $(q, p)$ of a FCE is a rest point of the dynamical system and a local attractor of the dynamics and thus a culture.

There are two other cooperative equilibria of the team trust game for high values of $p$ and low values of $q$. In both equilibria the investors invest and there is no punishment. Hence, the levels of cultural intolerance are zero and therefore there is no movement in $q$. However, these equilibria differ in the rewarding policy of the selfish allocator, as was stated in propositions 3 and 4. Notice that the levels of profits of a selfish allocator are strictly greater than the levels of profits of the fair-minded allocator. Therefore, the proportion of fair-minded allocators $p$ diminishes over time, until eventually the dynamics leaves the region for which any of these two equilibria exist.

In other words, although for an initial high $p$ and a low $q$, the first generations coordinate in cooperative equilibria where selfish allocators set unfair rewards, these cannot constitute long-run cultures because the proportion of selfish allocators will increase over time until society ends up playing the ISE. The reason is that for this lower proportion of fair-minded allocators, the reciprocal investors prefer not to invest.

Next, we check the robustness of the inefficient equilibria of the game.

PROPOSITION 9: *The only inefficient culture is the NCC and it exists for any pair of distributions* $(q_t, p_t)$ *such that* $q_t < \min\{\overline{q} = \frac{z}{2\gamma}, q' = \frac{1}{\sqrt{2\lambda}}\}$ *and* $p_t < \frac{2c}{1-q_t}$.

*Proof*: See Appendix. ■

We also give here a sketch of the proof. Recall that in the NCE the payoff of each player is zero. Thus, the optimal education effort levels are zero and there are no incentives to socialize. Hence, if society coordinates in this equilibrium, the investor population will remain locked into this distribution of preferences. Also, as the profits of both types of allocators are zero, there is no movement in $p$. For the regions in the space $(q, p)$ where this equilibrium is unique, it will constitute a local attractor of the dynamics. A society with a very high proportion of selfish individuals (both investors and allocators) will get stuck in this inefficient trap.

There are two other inefficient equilibria of the team trust game. We show that these equilibria cannot result in a culture.

First, regarding the dynamics of the investor population in the ISE region, we observe that the levels of cultural intolerance of both types are nonnegative. This happens because a reciprocal investor parent dislikes the behavior of his selfish child of not punishing an unfair reward, while a selfish investor parent dislikes the behavior of his reciprocal child of not investing.

We equate the socialization effort functions of both types of parents to obtain the demarcation curve $q(p)$, that is, the locus of pairs $(q, p)$ such that the distribution of preferences in the investor population remains constant over time. Note that for a given $p$, if $q > q(p)$, $q$ decreases, and if $q < q(p)$, $q$ rises.

In this equilibrium, the profits of a selfish allocator are strictly higher than the profits of a fair-minded allocator. It can be easily calculated that the dynamics of the allocator population in this region is given by $\Delta p_t = p_t(1 - p_t)\varphi[-(1 - q)^2]$, which is negative for all $q$. Thus, throughout this region, the proportion of fair-minded allocators $(p)$ falls.

Summing up, in the ISE region $p$ always decreases and $q$ changes depending on its location, above or below the demarcation curve. But in both cases the dynamics will eventually leave this region and, depending on the initial condition, it will reach the FCC region for high values of $q$ or the NCC region for low values of $q$ and $p$. Once the dynamics has reached one of these two regions, society will remain there because both of them constitute a culture. A formal analysis showing this result is contained in the Appendix.

The other inefficient equilibrium is the quasi-cooperative equilibrium with punishment. Concerning the dynamics in the investor population, the levels of cultural intolerance of both types are positive. This happens because a reciprocal investor parent dislikes the behavior of his selfish child of not punishing an unfair reward, while a selfish investor parent dislikes the behavior of his reciprocal child of punishing an unfair reward and losing material payoffs.

Using the same procedure as in the previous case, we obtain the demarcation curve. This curve turns out to be independent of $p$. In particular, it is given by the solution $(q'')$ of a quadratic function. If $q > q''$, $q$ increases, and if $q < q''$, then $q$ decreases.

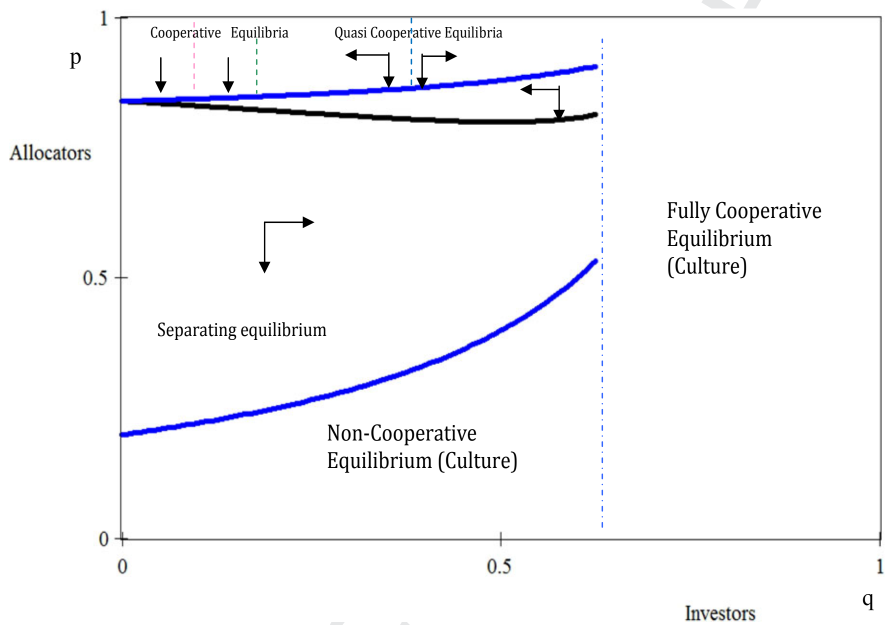


Figure 1:

Q2





On the other hand, the profits of the selfish allocators are higher than those of the fair-minded allocators, and the dynamics in this region leads to a fall in $p$.

Eventually, depending on the initial conditions, the dynamics will leave this region either for the FCC region with a very high $q$, or for the ISE region with a smaller $q$. But we already know that the process will go on and it will end up in the FCC region or in the NCC region, depending on the initial conditions.

We sum up the results obtained in this section in the following corollary.

COROLLARY 1: *The only long-run outcomes of dynamical system 1 are the FCC or the NCC.*

Figure 1 depicts graphically the results we have obtained, for some particular values of the parameters.

### 7.1. Discussion

In the previous section, we have seen that the only cooperative equilibrium that survives as a long-run culture is the FCE. This efficient culture provides a fair retribution to all players and is characterized by a high proportion of reciprocal investors and by any preference distribution of allocators. On the other hand, the inefficient NCE, which exists for a low proportion of reciprocal investors and fair-minded allocators, is the only equilibrium that can survive as a long-run culture among the inefficient equilibria. Surprisingly, there is no observed punishment in either culture. The ultimate reason for this result relies on the credibility of punishment. In the FCC, the threat of punishment is so high and credible that it modifies the behavior of selfish allocators, leading them to set fair rewards in order to avoid punishment. This is the only situation in which selfish allocators do not have any competitive advantage over fair-minded allocators in terms of profits.

However, in the NCC the small number of reciprocators in the investor population generates a situation in which punishment is not credible at all and therefore selfish allocators will set low returns. And as their proportion is so high, the incentives of the team to invest are destroyed. In this sense, the presence of a credible threat of punishment is crucial for obtaining a cooperative culture with fair returns in the long run.

Some of the assumptions concerning the punishment institutions play an important role in obtaining the previous results. Let us discuss the influence of relaxing some of these assumptions in turn. First, note that for FCC to exist, it is crucial that $\lambda \geq 0.5$, that is the damage caused by the coordinated punishment is big enough to make the threat of punishment effective. Otherwise, the punishment cannot lead to an increase in cooperation. The reason is that as $\lambda < 0.5$, that is, the inflicted damage is low, the selfish allocator will prefer to set $b_s = 0$ and to be punished with probability one rather than to offer $b_s = 1/2$ and avoid punishment. Therefore, the FCE does not exist for $\lambda < 0.5$. Nevertheless, for a sufficiently high value of $p$ there will be a quasi-cooperative equilibrium with punishment. But this will never constitute a culture because $p$ decreases over time, since the profits of selfish allocators are greater than those of fair-minded allocators. Summarizing, if $\lambda < 0.5$, the NCC is the unique global attractor of the dynamics.

Second, we want to know the results if $z/2 > \gamma$, that is, if the level of peer pressure is not enough to compel selfish investors to punish the allocator for high values of $q$. The first consequence is that only reciprocators punish in equilibrium. Therefore, the basin of attraction of the FCC decreases. In particular, if $z/2 > \gamma$, then $\overline{q}$ is greater than 1 and then, if $\lambda > 0.5$, the FCC only exists for the interval $[q', 1]$. Notice also that the cooperative behavior of all types of players in the new interval is the same as before, but now only reciprocal investors can credibly threaten punishment.

Third, some comments on the influence of the degree of aversion to disadvantageous inequality $\alpha$ and the cost of the investment $c$ are necessary. The basin of attraction of the FCC will be greater the larger $\alpha$ is, and the smaller $c$ is.

Finally, it is interesting to understand the role of cultural transmission in enhancing cooperation. If the evolution of the preference distribution in the investor population is also governed by material payoffs or profits as in the allocator population, then the basin of attraction of the NCC increases and that of the FCC diminishes. Namely, the basin of attraction of the FCC coincides exactly with the region where the FCE exists. Only if society starts in this region there will be a FCC. This happens because for the rest of equilibria the material payoffs of the selfish investors are greater or equal than those of the reciprocators. Summarizing, cultural transmission in the investor population enhances cooperation and efficiency in the sense that enlarges the basin of attraction of the FCC.

## 8. Concluding Remarks

The main result of this model is that cooperation only evolves and is maintained if there is enough punishment capacity in society and if there are enough individuals willing to implement both coordinated and peer punishment. The FCC is achieved under the threat of effective coordinated punishment, but this threat is, in turn, supported by the presence of a high proportion of reciprocators in the investor population. This fact illustrates the main difficulty of obtaining cooperation: uniqueness is not achieved. Our model shows that initial conditions matter because they can lead society to a different state in the long run. If society is able to build strong punishment institutions and can accomplish, through socialization mechanisms, a preference distribution in the population willing to implement these institutions, then society can settle into a cooperative and efficient culture. However, if it is not able to reach a "sufficient" proportion of reciprocators in society, a non-cooperative and inefficient culture will be established.

Changes in the punishment institutions as, for instance, in the damage caused by coordinated punishment or the level of peer pressure, might cause large changes in the long-run distribution of preferences and behavior (the culture). The new punishment institutions produce these changes not only in the short run but also in the long run through the dynamics of both populations, in particular, by means of incentives to socialize future generations of investors into a kind of preference more prone to using any sort of punishment.

## Appendix

*Proof of Proposition 1*: Suppose that the allocator offers $b < 1/2$. Note that punishment is a best response for reciprocal investors to $(p, np)$ if:

$$\mu(b - z/2 - \alpha[(1-b)(1-\lambda) - b)] \quad +(1-\mu)[b - z - \alpha((1-b) - b))$$
$$\geq \mu(b - \gamma - \alpha[(1-b)(1-\lambda) - b)] + (1-\mu)[b - \alpha((1-b) - b)).$$

That is, if $\mu \geq \mu^*(b) = \frac{z}{z/2 + \alpha\lambda(1-b) + \gamma}$.

Not punishing the allocator is a best response against $(p, np)$ for selfish investors, since $\mu(b - \gamma) + (1-\mu)b > \mu(b - z/2) + (1-\mu)(b - z)$ because $\mu < \widetilde{\mu} = \frac{z}{z/2 + \gamma}$.

The profile $(p, p)$ is a BNE, if punishment is a best response for selfish investors to $(p, p)$. That is, if $(b - z/2) \geq \mu(b - \gamma) + (1 - \mu)b$. Therefore, if $\mu \geq \overline{\mu} = \frac{z}{2\gamma}$.

Hence, for $\mu \in [\mu^*(b), \widetilde{\mu}]$, $(p, np)$ is a BNE and for $\mu \in [\overline{\mu}, 1]$, $(p, p)$ is a BNE. It is easy to check that $(np, np)$ is a BNE for each value of $\mu$. ■

*Proof of Lemma 2*: For any given $\widehat{\mu} \in [\mu^*(0), \mu^*(0.5)]$, there is a $\widehat{b} = \frac{\widehat{\mu}(\alpha\lambda+\gamma+z/2)-z}{\widehat{\mu}\alpha\lambda}$, such that $\mu^*(\widehat{b}) = \widehat{\mu}$ because of the continuity and monotonicity of $\mu^*(b)$. Then for $b \in [0, \widehat{b}]$, $\widehat{\mu} > \mu^*(b)$ which implies that the BNE of $CP(\mu, b)$ is $(p, np)$. However, for $b \in [\widehat{b}, 0.5]$, $\widehat{\mu} \leq \mu^*(b)$ which implies that the BNE is $(np, np)$. Therefore, the options for the allocator are either to offer $b_s = 0$ and the reciprocators will punish or to offer $b_s = \widehat{b}$ and nobody punishes. The expected payoff in the first option is $\Pi^s(0) = 2(1 - \mu^2\lambda)$ and the payoff in the second option is $\Pi^s(\widehat{b}) = 2(1 - \widehat{b})$.

Therefore, $b_s = 0$ is preferred to $b_s = \widehat{b}$, when $2(1 - \widehat{\mu}^2\lambda) \geq 2(\frac{z-\widehat{\mu}(\gamma+z/2)}{\widehat{\mu}\alpha\lambda})$. This defines the following cubic equation: $z = \widehat{\mu}(\alpha\lambda + \gamma + z/2) - \alpha\lambda^2\widehat{\mu}^3$. If the discriminant of this equation is positive, then there is a real root, which is negative and two complex roots, therefore setting $b = \widehat{b}$ is better. If the discriminant of this equation is negative, then there are three real and unequal roots, one of them negative. We will denote the positive roots: $\mu_1$ and $\mu_2$. Both $\mu_1$ and $\mu_2$ are greater than $\mu^*(0)$.Then, the optimal reward policy of the allocator is: for $\widehat{\mu} \in [\mu^*(0), \mu_1]$ and $\widehat{\mu} \in [\mu_2, \mu^*(0.5)]$ to set $b_s = \widehat{b}$ and for $\widehat{\mu} \in [\mu_1, \mu_2]$ to set $b_s = 0$. ■

*ISE Where Reciprocal Investors Invest.*

In this equilibrium, both types of allocators offer $b = 0.5$ and the beliefs are $\mu(r/(I, I) = 1$. The incentive compatibility constraint for a selfish investor is: $0 \geq q_t(0.5) - c$. That is, $2c \leq q_t$.The incentive compatibility constraint for a reciprocal investor is: $q_t(0.5) - c \geq 0$. That is, $q_t \geq 2c$.Therefore, for $q_t = 2c$ both constraints are satisfied. Notice that this equilibrium exists for any $\lambda > 0.5$.

For $\lambda \leq 0.5$, even when the beliefs are $\mu(r/(I, I) = 1$, the selfish allocator prefers to offer $b_s = 0$ and suffering punishment rather than offering $b = 0.5$ and avoiding punishment. The fair-minded allocator sets $b_f = 0.5$. However, note that both incentive compatibility constraints do not hold simultaneously. ■

*Proof of Proposition 7*: In this pooling equilibrium $\mu = q$. By Lemma 2, $b_s = 0$ is preferred to $b_s = \widehat{b}$, when the following cubic in equation condition holds: $z \leq q(\alpha\lambda + \gamma + z/2) - \alpha\lambda^2 q^3$.

The simulations we have run show that, in the range of parameters implied by our assumptions, the discriminant of this equation is negative. Then, there are at most two positive unequal roots. We call the positive roots: $q_1$ and $q_2$, where $q_1 < q_2$. Thus, if $\widehat{q} \in [q_1, q_2]$ the best reward policy of the selfish allocator is to set $b_s = 0$. We can show by numerical simulations that $q_1$ is always greater than $q^*(0)$ and that $q_2$ is always greater than 1. On the other hand, by Lemma 3 if $q_t \in (q^*(0.5), \min(\bar{q}, q'))$ the best reward for the selfish allocator is to set $b_s = 0$, even though she will be punished by the reciprocal investors according to Proposition 1. Therefore, for $q_t \in [\, q_1, \min(\bar{q}, q')]$ the selfish allocator sets $b_s = 0$.

A selfish investor will invest when the following condition holds: $p_t(0.5) + (1 - p_t)[-q_t\gamma] - c \geq 0$.That is, if $p_t \geq \frac{q_t\gamma + c}{q_t\gamma + 0.5}$.

The condition for a reciprocal investor is: $p_t(0.5) + (1 - p_t)[(1 - q_t)(-z - \alpha) + q_t(-z/2 - \alpha(1 - \lambda))] - c \geq 0$. That is, if $p_t \geq \frac{z + \alpha + c - q_t(z/2 + \alpha\lambda)}{z + \alpha + 0.5 - q_t(z/2 + \alpha\lambda)}$. This last expression is the binding incentive compatibility constraint. The payoff for the fair-minded allocator is 1 and for the selfish allocator is $2(1 - \lambda q_t^2)$. ■

*Transition Probabilities of the Socialization Process.*

Let $P^{ij}$ denote the probability that a child of a parent with preferences $i$ is socialized to preferences $j$. The socialization mechanism is characterized by the following transition probabilities where $q_t$ is the proportion of reciprocal investors: $P_t^{ss} = \tau_t^s + (1 - \tau_t^s)(1 - q_t)$, $P_t^{sr} = (1 - \tau_t^s)q_t$, $P_t^{rr} = \tau_t^r + (1 - \tau_t^r)q_t$,and $P_t^{rs} = (1 - \tau_t^r)(1 - q_t)$.

Given these transition probabilities it is easy to characterize the dynamic behavior of $q_t$: $q_{t+1} = [q_t P_t^{rr} + (1 - q_t)P_t^{sr}]$.

*Proof of Proposition 8*: In the cooperative equilibria with $b_s = \widehat{b}$, the payoff of the fair-minded allocator is $\Pi^f = 1$ and the payoff of the selfish allocator is $\Pi^s = 2(1 - \widehat{b}) = 2(\frac{z - q(\gamma + z/2)}{q\alpha\lambda})$, where $\widehat{b} < 1/2$. The dynamics of the allocator population is given by $\Delta p_t = p_t(1 - p_t)\varphi(1 - 2(\frac{z - q(\gamma + z/2)}{q\alpha\lambda}))$. This expression is negative if $\widehat{b} < 1/2$, so $p$ decreases, $\forall q$.

The payoff of the reciprocal investor is $U_r = p(0.5) + (1 - p)[\widehat{b} - \alpha(1 - 2\widehat{b}) - c \geq 0$ and the payoff of the selfish investor is $U_s = p(0.5) + (1 - p)\widehat{b} - c \geq 0$.

In this equilibrium the levels of cultural intolerance and the corresponding optimal education efforts are zero because $V^{rr} = V^{rs}$ and $V^{ss} = V^{sr}$. Therefore, $q$ does not change.

A similar argument applies for cooperative equilibria with $b_s = 0$.

It can be checked that the two critical values on $p$ that define the boundaries of these two cooperative equilibria, $p'$ and $p''$, are always smaller than $\frac{c + \alpha(1 - q_t)}{1 - q_t}$ for all $q$. Hence, the dynamics will always reach the ISE region. ■

*Proof of Proposition 9*: The payoff of the fair-minded allocator in the ISE is $\Pi_t^f(q, p) = (1 - q)^2$, while the payoff of the selfish allocator is $\Pi^s(q, p) = 2(1 - q)^2$, then the dynamics of the allocators population is given by: $\Delta p = p(1 - p)\varphi(-(1 - q)^2) < 0$, where we have dropped the subindex $t$ for clarity of exposition. Thus, $p$ decreases.

The payoff of the reciprocal investor is $U_r = 0$ and the payoff of the selfish investor is $U_s = (1-q)\cdot p\cdot(0.5) - c$. Thus, the levels of cultural intolerance are nonnegative: $\Delta V^r = V^{rr} - V^{rs} = 0 - ((1-q)(p(0.5) + (1-p)(-\alpha)) - c \geq 0$ and $\Delta V^s = V^{ss} - V^{sr} = (1-q)p(0.5) - c \geq 0$.

Therefore, the optimal educational efforts are given by: $\tau^{r*}(q,p) = k\Delta V^r(q)(1-q) = k(1-q)(c-(1-q)(p(0.5)+(1-p)(-\alpha))$ and $\tau^{s*}(q,p) = k\Delta V^s(q)q = kq((1-q)p(0.5)-c) \geq 0$.

We obtain the demarcation curve $q(p)$ that makes $\Delta q_t = 0$, equating $\tau^{r*}((q,p)) = \tau^{s*}((q,p))$. Then, the demarcation curve is given by: $q(p) = (\alpha p - \alpha)q^2 + q(2\alpha - p(2\alpha+0.5) + (p(\alpha+0.5)-\alpha) - c = 0$. Note that for a given $p$, if $q > q(p), \tau^{r*}((q,p)) < \tau^{s*}((q,p))$ and $q$ decreases and if $q < q(p), \tau^{r*}((q,p)) > \tau^{s*}((q,p))$, and $q$ increases. This demarcation curve belongs to the region in which the equilibrium exists.

The payoff of the fair-minded allocator in the quasi-cooperative equilibrium is $\Pi^f = 1$ while the payoff of the selfish allocator is $\Pi^s((q,p)) = 2(1-\lambda q^2)$. Then, the dynamics of the allocator population is given by $\Delta p = p(1-p)\varphi(2q^2\lambda - 1)$. This expression is negative when $q \leq q'$, so $p$ decreases, $\forall q$.

The payoff of the reciprocal investor is $U_r = p(0.5) + (1-p)[(1-q)(-z-\alpha) + q(-z/2 - \alpha(1-\lambda)] - c \geq 0$ and the payoff of the selfish investor is $U_s = p(0.5) + (1-p)[-q\gamma] - c \geq 0$.

Thus, the levels of cultural intolerance are:

$\Delta V^r = V^{rr} - V^{rs} = p(0.5) + (1-p)[(1-q)(-z-\alpha) + q(-z/2 - \alpha(1-\lambda)] - c - [p(0.5) + (1-p)(-\alpha - q\gamma) - c] = (1-p)[q(z/2+\gamma+\alpha\lambda) - z] > 0$. This expression is positive if $q > z/(z/2+\gamma+\alpha\lambda)$ as it is the case. And $\Delta V^s = V^{ss} - V^{sr} = p(0.5) + (1-p)[-q\gamma] - c - ((0.5)p + (1-p)((1-q)(-z) + q(-z/2) - c) = (z - q(\gamma + z/2)(1-p) > 0$. This expression is positive because $q < z/(\gamma + z/2)$.

The optimal educational efforts are given by:

$\tau^{r*}(q,p) = k\Delta V^r(q)(1-q_t) = k(1-q)(1-p)[q(z/2+\gamma+\alpha\lambda) - z] > 0$.

$\tau^{s*}(q,p) = k\Delta V^s(q)q = kq(z - q(\gamma+z/2)(1-p) > 0$.

We obtain the demarcation curve by equating $\tau^{r*}(q,p) = \tau^{s*}(q,p)$. This curve is given by the expression $-q^2\alpha\lambda + q(z/2+\gamma+\alpha\lambda) - z = 0$ and it turns out to be independent of $p$. In particular, it is given by the solution $(q'')$ to the previous quadratic function. Hence, if $q > q''$, $\tau^{r*}((q,p)) > \tau^{s*}((q,p))$ and $q$ increases and if $q < q''$, $\tau^{r*}((q,p)) < \tau^{s*}((q,p))$, and $q$ decreases. ■

## References

BISIN, A., and T. VERDIER (2000) Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits, *Quarterly Journal of Economics* **115**, 955–988.

BISIN, A., and T. VERDIER (2001) The economics of cultural transmission and the dynamics of preferences, *Journal of Economic Theory* **97**(2), 298–319.

BOYD, R., H. GINTIS, S. BOWLES, and P. J. RICHERSON (2003) The evolution of altruistic punishment, *Proceedings of the National Academy of Sciences of the United States of America* **100**(6), 3531–3535.

BOYD, R., H. GINTIS, and S. BOWLES (2010) Coordinated punishment of defectors sustains cooperation and can proliferate when rare, *Science* **328**, 617–620.

CAVALLI-SFORZA, L., and M. FELDMAN (1981) *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

DUFWENBERG, M., A. SMITH, and M. VAN ESSEN (2013) Hold-up: With a vengeance, *Economic Inquiry* **51**, 896–908.

FALK, A., E. FEHR, and U. FISCHBACHER (2005) Driving forces behind informal sanctions, *Econometrica* **7**(6), 2017–2030.

FEHR, E., and S. GÄCHTER (2000) Cooperation and punishment in public goods experiments, *American Economic Review* **90**, 980–994.

FEHR, E., and S. GÄCHTER (2002) Altruistic punishment in humans, *Nature* **415**, 137–140.

FEHR, E., and K. SCHMIDT (1999) A theory of fairness, competition and cooperation, *Quarterly Journal of Economics* **114**, 817–868.

GÄCHTER, S., E. RENNER, and M., SEFTON (2009) The long-run benefits of punishment, *Science* **322**, 1510–1512.

HAUERT, C., A. TRAULSEN, M. A. NOWAK, H. H. BRANDT, and K. SIGMUND (2007) Via freedom to coercion: The emergence of costly punishment, *Science* **316** 1905–1907.

HENRICH, J., R. McELREATH, A. BARR, J. ENSMINGER, C. BARRETT, A. BOLYANATZ, J. C. CARDENAS, M. GURVEN, E. GWAKO, N. HENRICH, C. LESOROGOL, F. MARLOWE, D. TRACER, and J. ZIKER (2006) Costly punishment across human societies, *Science* **312**, 1767–1770.

HUCK, S., and M. KOSFELD (2007) The dynamics of neighbourhood watch and norm enforcement, *The Economic Journal* **117**, 270–286.

OSTROM, E. (1990) *Governing the Commons: The Evolution of Institutions for Collective Action*. New York: Cambridge University Press.

OSTROM, E. (2000) Collective action and the evolution of social norms, *The Journal of Economic Perspectives* **14**(3), 137–158.

ROB, R., and P. ZEMSKY (2002) Cooperation, corporate culture and incentive intensity, *RAND Journal* **33**, 243–257.

SIGMUND, K., H. DE SILVA, A. TRAULSEN, and C. HAUERT (2010) Social learning promotes institutions for governing the commons, *Nature* **466**, 861–863.

YAMAGISHI, T. (1986) The provision of a sanctioning system as a public good, *Journal of Personality and Social Psychology* **51**, 110–116.

## Queries

Q1: Author:Please provide city and country name in authors' affiliation.

Q2: Author: Citation of Figure 1 is appearing. However, caption of Figure 1 has not been provided. Please provide the same for this article.