Coordinated Punishment and the Evolution of Cooperation

Olcina, G. and Calabuig, V. *
Department of Economic Analysis, ERI-CES, University of Valencia

July 2011

Abstract

In this paper we analyze a team trust game with coordinated punishment of the investors to the allocator 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 the cost of the coordinated punishment depends 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 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 fully cooperative culture is larger, the higher is the institutional capacity of coordinated punishment, the higher is the level of peer pressure and the smaller is the individual cost of coordinated punishment.

*The authors acknowledge financial support from the Spanish Ministry of Education and Science project ECO2008-06035/ECON. 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 la Neuve, 2010), Public Economic Theory PET '11 (Bloomington, Indiana, 2011) and Fifth International Conference on Dynamic Games (Saint Petersburg, 2011).

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 performed with people from these societies show significant levels of trusting and cooperative behaviour in this class of games even when cooperative behavior is costly. These observations raise the question of which are the sources of cooperation (and efficiency) in a non-repeated incomplete contract scenario.

Both, the experimental evidence[1] and recent theoretical work[2] from anthropology and evolutionary theory show that individually costly cooperative behaviour can be sustained by costly punishment. However, from our point of view, there are two important weaknesses on most of the existing models on punishment and on the experimental work performed on this issue. On the one hand, current models and experiments usually assume that punishment is made on an individual basis, that is, it is uncoordinated. But this way of modelling punishment is quite unrealistic because most of the punishment exerted in real-life situations is coordinated. In particular, almost all models and experiments ignore the empirically relevant fact that both the capacity of punishing (inflicting a damage) to the punished target and the costs of punishing typically depend on the number of individuals involved in the punishing activity. For instance, it is needed a minimal number of punishers in order to obtain effective punishment in a strike or in a boycott.

On the other hand, the other weakness of most of the experimental literature on the role of punishment is that it has focused on symmetric team or group situations such as the public goods game. But there has been very few works on the impact of punishment on the kind of assymmetric situations that characterize modern markets and privately owned corporations. By this we mean economic games based on specialization and on the division of labour such as, for instance, the principal - agent relationship, the hold-up game or in general any sequential transaction between a seller and a buyer.

[1] See for example, Fehr and Gachter (2000, 2002), Gachter, Renner and Sefton (2008), Herrman, Thoeni and Gachter (2008), Yamagishi (1986), Hauert, Traulsen, Nowak, Brandts and Sigmund (2007), Erta, Page and Putterman (2009).

[2] See for example Henrich et al. (2006), Boyd, Gintis, Bowles and Richerson (2003), Kosfeld, Rields and Okada (2009), Rockenbach and Milinski (2006), Carpenter (2007)

A well-known social dilemma that captures these assymmetric 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). To invest 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 labour market. In many employment relations, a group of employees is hired by a single employer (the firm). The labour contract in these cases is highly incomplete and it assigns usually 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. Baron and Kreps (1999) point out that this threat of hold up of the investors (the employees) by the allocator (the employer) can be mitigated by a balance of power, arising from the credible threat by the employees to retaliate if they are exploited.

In this paper we analyze a team trust game with coordinated punishment of the investors to the allocator and where we also add a final stage of peer punishment. It seems crucial for the effectiveness of punishment, in addition to institutions and law, the ability of different groups to overcome the collective action problem that is at the core of almost any form of coordinated punishment.

The empirical literature on collective action[3] agrees on the importance of a number of factors that affect the likelihood of successful collective action. In our analysis we incorporate the two most important factors: the possibility of peer punishment and the heterogeneity of preferences in the population.

Regarding to the first factor, notice that punishment itself is a public good among the investors and it is also subject to free-riding behaviour. Peer punishment is a "second-order" punishment to those individuals who free-ride in the coordinated punishment phase to the allocator. We will denote this sort of punishment as peer pressure.

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

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 in the coordinated punishment phase provided the cost of punishing is low enough and/or to punish free-riding behavior of their team-mate. There are also two types of allocators: selfish or profit-maximizers and fair-minded allocators who always set a fair return.

To assume heterogeneity of preferences is nowadays quite standard but our main assumption is that preferences are endogenous, that is, the distribution of preferences in both populations evolves along time. Different forces govern the evolution of preferences in both populations. The dynamics of the allocators´ population is driven by market forces: profits. But the dynamics in the investors' population is governed by a cultural transmission process that combines intentioned and costly parental (direct) transmission with oblique transmission from the 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 aproximation to actual societies.

We are specially interested in the influence of the punishment institutions on the long-run distribution of preferences and behavior, particularly in the punishment coordination problem. Why and how do different societies succeed in solving this coordination problem? And what is the relation, if any, with the strength of the punishment institutions? The punishment institutions in our model are on the one hand the capacity of collectivelly punishing the allocator and the cost of coordinated punishing and on the other hand the level of peer pressure. However, it is important to notice that for punishment to be effective it is needed not only laws and institutions (exogenous to the individual) but also the willigness of these individuals to incur costs to implement the punishment. In other words, to implement the law or to make the institution work is individually costly.

We present an overlapping generations dynamic model in order to analyze this team trust game with coordinated punishment. We denote as a culture any stable steady state of the dynamics where the same particular 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 with high levels of trust and cooperation and fair returns or a Non-Cooperative Culture with no cooperation at all. In the fully cooperative culture 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 investors' population. By contrast in the Non-Cooperative Culture the threat of coordinated punishment is not credible at all because there is a low proportion of punishers in the investors´ population. As a result there are low levels of cooperation and efficiency.

The fully cooperative culture 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 the society.

But law and institutions to punish opportunistic allocators are not enough. The main determinant of the basin of attraction of the cooperative culture is a sufficiently high level of peer pressure relative to the individual cost of coordinated punishment targeted to the allocator. High peer pressure and therefore institutions that do favour 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 as for example, a union. But it is not the unique example. For instance, belonging to a community or a gang increases the damage inflicted by the group to the free-rider.

Summing up, the basin of attraction of the fully cooperative culture is larger, the higher is the institutional capacity of coordinated punishment, the higher is the degree of peer pressure and the smaller is the individual cost of coordinated punishment. And also the smaller is the cost of investment.

The intuition behind our main results is the following. Different institutions facilitate different behaviors and what is more important, they influence the long run incentives to socialize on particular preferences. In particular, strong punishment institutions related to coordinated punishment and to peer punishment, increase the effectiveness of punishment in the short-run through its effects on the constraints that individuals face. But they also increase the incentives to socialize on preferences that display negative reciprocity. This will, in turn, increase the effectiveness of punishment of future generations because of the presence of a larger proportion of punishers in the population. Therefore, this new distribution of preferences (with a greater proportion of punishers) will reinforce the effectivenes of a given institution. This might happen both because it increases the probability of having a punisher as a team-mate 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 as compared to the individual cost of coordinated punishment even the 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 (that is, strong institutions) both types of allocator prefer to set fair returns. If the society reaches this situation both types of investor lose their incentives to actively socialize their children. The society has reached the fully cooperative culture.

However, uniqueness is not achieved. Our model shows hysteresis: initial conditions matter, because they can lead the society to a different steady state. If the society starts in a distribution with a low proportion of punishers the above logic works exactly in the opposite direction and the society will get stuck in a very inefficient outcome.

Our work is related to two important strands of the literature. First, the experimental analysis of the so-called altruistic punishment that starts with Fehr and Gachter (2000, 2002) and continues with an impressive amount of evidence (see for example, Falk et al, 2005). All this evidence poses an important question for the 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 lead us too far to review the vast experimental and evolutionary literature on punishment. We have already explained that an important weakness of this literature is the assumption of uncoordinated punishment. To the best of our knowledge, the only work on coordinated punishment is Boyd, Gintis and Bowles (2010). These authors 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 to 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 these authors assume that punishment is equally effective whatever is the number of participants. Instead, we assume that punishment is effective only if a minimal 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 in particular, to the work on the endogenous determination of preferences and its interaction with institutions. For instance, Huck and Kosfeld (2007) analyze in an evolutionary model how what they call weak institutions interact with preferences for punishment. As in our approach, institutions and law are only effective if individuals are willing to engage in an individual costly implementation of these tools. Our paper differs from theirs in three main aspects. They analyze a public good game, punishment is not coordinated and the replicator-like dynamics they use does not display any cultural bias.

The paper is organized as follows. In the next section we present the model. In section 3 we introduce the social preferences. In section 4 we show the punishment and the 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 team trust 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 investment project a joint surplus of size $2H$ is produced. Otherwise, i.e. 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 H is the gross gain per investor and that investment has a cost $c > 0$. We also assume that $c < H/2$ and that not investing is costless. In the rest of the paper, for simplicity, we will normalize H to 1.

In the second stage, the rewarding phase, the allocator after observing a surplus of size 2, has to assign a percentage b of $H = 1$ to each investor, 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. So, the interim monetary payoffs are $b - c$ for each investor and $2(1 - b)$ for the allocator.

In the third stage, the coordinated punishment phase, the investors, after being paid, can engage in a costly punishment coordination game. They have to decide simultaneously whether to punish the allocator (action p) or not to punish (action np). Again, only if both investors choose to punish, a proportion λ of the payoff obtained by the allocator is destroyed, where $0 < \lambda \leq 1$. But if just one or none of both 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 adheres 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 $z/2$. So, for instance, the interim material payoffs with effective punishment are: $(b - z/2 - c)$ for each investor and $(2(1 - b)(1 - \lambda)$ for the allocator.

Finally, the fourth stage is the peer punishment phase. If the cooordinated punishment to the allocator has not succeeded because of a defection of a team mate, then the non-defecting investor has now the option of punishing the defector mate at some cost. We assume that this peer punishment creates a loss in the utility of the punished mate of size γ.

Some comments on the institutional parameters that characterize the coordinated punishment to the allocator (λ and z) and the peer punishment (γ) are in order. The parameters λ and z capture the two relevant features of coordinated punishment: how much damage can the team inflict on their allocator and how much it costs to each investor.

In a labour market context, λ would be the punishment that the team of workers can inflict on the firm. For example, if $\lambda = 1$, the workers can destroy all the surplus of the firm. It depends on the workers' ability 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 the society, etc... It might also differ across different types of jobs depending on the strategic position of the worker in the production process. The parameter λ can also be interpreted as the maximal punitive sanction that the legal system provides to an agent in order to punish the opportunistic behavior of the other party, when the punisher is not able to recover all the cost of his investment. (see Dufwenberg *et al*, 2011)

Concerning the peer punishment phase, the parameter γ is our measure of the level of peer pressure. Note that peer punishment is an individual second order punishment, because it intends to punish the defectors in the previous coordinated stage. We think that it captures a realistic feature in labor markets in which some workers punish the strike-breaker behavior of other co-workers and this is an important constraint in 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 neither individually neither collectively because it is costly and does not increase their payoff. Given that the allocator will not be punished, she will offer a proportion $b = 0$ to the investors 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 team trust game, both the promise of rewarding by the allocator and the threat of punishment by the team of investors are not credible.

In this paper we will assume that there is heterogeneity of preferences and 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 preferences.

3 Social Preferences: Reciprocal Altruism.

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

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 players will also punish a team mate who has failed to punish an unfair reward of the allocator.

The reciprocal investors are punishers because they are concerned not only by their monetary payoff but they also aspire to get a fair return compared to the payoff of the allocator and, hence, any return smaller than what it 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 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 on 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 the 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 neither the type of his/her team mate nor the type of the 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. In particular, we will analyse the punishing behavior of the different

[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 $1/2$.

types of investors, both in the coordinated and in the peer punishment phase, and the optimal reward policy of the selfish allocator.

First of all, we will make three assumptions on 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 assumption states that for a reciprocal investor to punish an unfair reward of the allocator is the best response to the choice of punishment by the other member of the team whatever is the latter´s type. The reason is that by successfully punishing the allocator, a reciprocator reduces inequality with respect to the latter and this positive effect in her utility, $\alpha\lambda$, more than compensates the reduction in his material payoff z. Consequently, this is a very straightforward assumption that simply states that reciprocators are conditional punishers. 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 lack any interest.

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

If this assumption holds, the individual cost of a succesful coordinated punishment to the allocator is strictly smaller than the damage inflicted by the peer punishment. Otherwise, a selfish investor will never participate in the coordinated punishment to the allocator, even if he knows for sure that his team mate is a reciprocal investor who is going to punish him. That is, peer pressure will never be effective. We want obviously to analize the more interesting case where the assumption holds and peer pressure is effective, at least, under some conditions.

Assumption 3 : $\lambda \geq 0.5$.

This assumption sates that the 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. That is, the threat of damage inflicted by the coordinated punishment is not enough to induce her to make a generous offer because she obtains a higher expected payoff setting a very low reward.

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 member of the team has to decide whether to use peer punishment against the other team mate, at a positive cost, or not to do it. 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 who dislikes disadvantageous inequality with his team mate will punish a team mate who has free-rided in the previous coordinated punishment stage. This will hold whenever the damage inflicted to the defector γ is sufficiently high as compared to the cost of doing so, because the action of peer punishment reduces the inequality with the defector team mate.

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

In order to reduce the complexity of the analysis we do not incorporate in the utility function of the reciprocal investor the term concerning the inequality with the team mate and we normalize the cost of peer punishment to zero. Instead we assume in the rest of the paper 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 team mate has been a defector will choose the peer punishment.

4.2 The Coordinated Punishment Subgame.

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

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

Any coordinated punishment subgame is characterized by a belief μ 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 μ 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)) The BNE (np,np) for any $\mu < \mu^*(b) = \frac{z}{z/2+\alpha\lambda(1-b)+\gamma}$.
ii) The BNE (p,np) for any $\mu \in [\mu^*(b), \overline{\mu} = \frac{z}{2\gamma})$.
iii) The BNE (p,p) for any $\mu \in [\overline{\mu} = \frac{z}{2\gamma}, 1]$

Proof: See Appendix.

Notice that 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 if the probability of having a reciprocal mate is high enough. Therefore, for $\mu \geq \overline{\mu}$, as the proportion of reciprocators is so high, both types of investors will punish the allocator.

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 do not punish. This bound $\mu^*(b; \alpha, \lambda, z, \gamma)$ is increasing in z and b and decreasing in λ, α and γ.

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

Notice that there is multiplicity of equilibria in this subgame and we have made an equilibrium selection, choosing for each μ the equilibrium with the highest probability of punishment. In particular, the profile (np, np) is a BNE for all μ. And 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 it is selected the BNE *(p,p)*.

We turn now 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 the punishment, that is, the minimal reward at which no type of investor will punish. The expected cost of the punishment will depend on the proportion of reciprocators in the investor population μ 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 checked that in that case, the best reward 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: See Appendix.

Note that $\mu^*(0)$ is the maximal proportion of reciprocators in the investors population such that both types of investors do not punish at $b = 0$. 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 allocator prefers to offer the lowest return $b_s = 0$.

Lemma 2 *For any* μ, *such that* $\mu^*(0) < \mu < \mu^*(0.5) = \frac{z}{(\alpha/2)\lambda+\gamma+z/2}$, *there exists a* $\widehat{b}(\mu)$, *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 μ 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 μ[5], only the reciprocators choose to punish any offer $b < 1/2$.

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

Proof: See Appendix.

For this range of values of μ 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 μ' that comes from comparing the previous both expressions. Notice that only when $\lambda \leq 0.5$, then $\mu' \geq 1$, which means that the optimal reward policy is to set $b_s = 0$ for all μ.

Lemma 4 *For any μ such that $\mu \geq \overline{\mu} = \frac{z}{2\gamma}$ and if $\lambda \geq 0.5$, the optimal return for the allocator is to set $b_s = 1/2$.*

Proof: See Appendix

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

Recall that by assumption 3, $\lambda \geq 0.5$. However if $\lambda < 0.5$ the optimal reward policy would be $b_s = 0$.

Now we are ready to obtain the equilibria of the whole team trust game in the next section.

[5] The particular values of μ for which is optimal the first policy or the second one, 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.

5 Equilibria within a Generation.

To start with we characterize the efficient or cooperative equilibria of the team trust game. 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 with*** $b_s = 1/2$***.*** *If* $\lambda > 0.5$ *and for an investors´ preference distribution* $q_t \geq \min\left\{\overline{q} = \frac{z}{2\gamma}, q^{'} = \frac{1}{\sqrt{2\lambda}}\right\}$ *and for any allocators´ preference distribution* p_t*, there exists 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.*

Proof: See Appendix.

The equilibrium payoff for any member of the team is $(1/2 - c)$ and any type of allocator gets a payoff of 1. 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. In particular, this critical fraction depends on the parameters that represent the punishment institutions (namely, coordinated punishment: λ and z and peer punishment γ).

Given this high proportion of reciprocal investors, even the selfish investor will punish unfair offers fearing the peer punishment. As a consequence, the selfish allocators will also set $b_s = 1/2$, there is no punishment and both types of investors will invest. This result is driven by the fact that the high number of reciprocators in this equilibrium increases the probability of having a reciprocal team mate. Therefore, it reduces the cost of a potential "coordinated punishment" (to $z/2$) and also increases the possibility of suffering peer punishment in case of a defection.

Notice that the fully cooperative equilibrium does not exist for λ smaller or equal than 1/2. This is because for $\lambda \leq 1/2$, it holds that $q^{'} \geq 1$. Therefore this equilibrium exists if q is greater than $\overline{q}$. However, for $\lambda \leq 1/2$ the selfish allocators choose $b_s = 0$ and there would be punishment.

Proposition 3 ***A Cooperative Equilibrium with*** $b_s = \widehat{b} > 0$.
For an investors´ preference distribution q_t *such that* $q^*(0.5) > q_t > q^*(0)$

and an allocators´ preference distribution p_t *such that* $p_t \geq p'(q, \lambda, \alpha, c, z)$, *there exists 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 maximizer allocators set* $b_s = \widehat{b} < 1/2$ *and fair-minded allocators set* $b_f = 1/2$. *No punishment is observed in equilibrium.*

Proof: See Appendix.

In this equilibrium, the payoff for the fair minded allocator is 1 and the payoff for the selfish type of allocator is $2(1-\widehat{b})$. The selfish allocator offers $b = \widehat{b} < 1/2$, the minimal reward such that reciprocators do not punish. This reward depends on the proportion of reciprocal investors in the population. It is worth to choose 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 ***A Cooperative Equilibrium with*** $b_s = 0$.
For an investors´ preference distribution such that $q_t < q^*(0)$ *and an allocators´ preference distribution such that* $p_t \geq p''(c, \alpha) = \frac{\alpha+c}{\alpha+0.5}$, *there exists a pooling equilibrium in which both types of investors choose to invest. Profit maximizer allocators set* $b_s = 0$ *and fair-minded allocators set* $b_f = 1/2$. *No punishment is observed in equilibrium.*

Proof: See Appendix.

In this equilibrium, the proportion of reciprocators is so low that it is not worthwile 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 in this equilibrium.

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 ***Non-Cooperative Equilibrium (NCE).***
For every investors´ preference distribution q_t *and every allocators´ preference distribution* p_t, *there exists 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.

In the NCE both types of investors do not choose to make investment. Next, we characterize separating equilibria in which just one type of investor chooses the efficient action of investing.

Proposition 6 ***The Inefficient Separating Equilibrium.***
For any pair of preference distributions $(q_t,\, p_t)$ *such that* $\frac{c+\alpha(1-q_t)}{1-q_t} \geq p_t \geq \frac{2c}{1-q_t}$*, there exists an Inefficient Separating Equilibrium in which the selfish investor chooses to invest in the project whereas the reciprocator chooses not invest. Profit maximizer allocators set* $b_s = 0$ *and fair-minded allocators set* $b_f = 1/2$.

Proof. See Appendix.

The beliefs in the equilibrium path are $\mu(r/(I, I) = 0$. That is, if there is investment the allocator believes that the team is composed entirely of selfish members. The expected payoff for each selfish member of a team is $0.5 \cdot p \cdot (1-q) - c$, whereas the payoff for a reciprocator is zero. The expected equilibrium payoff for the selfish allocator is $2(1-q)^2$ and the expected payoff of the fair-minded allocator is $(1-q)^2$.

In this equilibrium the surplus is generated with a probability $(1-q)^2$ which is the probability that the team is composed by two selfish investors. Note that, paradoxically, selfish members of the team choose the efficient action while the reciprocators do not. Selfish investors invest because of the presence in the 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 that this equilibrium only exists for an intermediate range of values of p.

There exists another separating equilibrium in which, by 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 that we will introduce in the next section.

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

Proposition 7 ***A Quasi-Cooperative Equilibrium with punishment.*** *For an investors´ preference distribution* q_t *such that* $q_t \in [\overleftrightarrow{q}_1, \min\left\{\overline{q}, q'\right\}]$, *and an allocators´ preference distribution* p_t *such that* $p_t \geq p'''(q, \lambda, \alpha, c, z)$, *there exists 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* $\overleftrightarrow{q}_1$ *is the smallest positive real root of the cubic equation* $z = q(\alpha\lambda+\gamma+z/2)-\alpha\lambda^2 q^3$. *Profit maximizer 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 that punish low rewards from the selfish allocator. However, there has to be a high proportion of fair minded allocators that make profitable for both types of investors to choose to invest, despite that the reciprocal investors will have to punish unfair rewards. Note that the critical value on the proportion of fair-minded allocators comes from the incentive constraint of the reciprocal investor.

Notice that for some regions of (p, q) there is 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 is going to be introduced 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 mantain equilibrium play. Therefore, in each period individuals coordinate in a 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 investors 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 Allocators' Population.

We assume that at the end of each period, allocators which 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 maximizer type of allocator, respectively. Notice that φ is a positive constant low enough in order 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 Investors' Population.

Preferences in the investor population are culturally transmitted according to an intergenerational transmission process. Children acquire preferences through 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 of Bisin and Verdier (2001) which is the economic version of the anthropological model of Cavalli-Sforza and Feldman (1981).

We consider overlapping generations of investors who only live for two periods (as a young 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[6] and has to make a (costly) decision regarding his/her child education, trying to transmit his/her 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 this preferences with probability τ^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 in the investors´ population (oblique transmission).

The transition probabilities[7] P^{ik}, determined by this socialization mechanism, can be easily computed and then obtained the dynamic evolution of the distribution of preferences which is given by the following equation on differences[8]:

$\Delta q_t = q_t(1 - q_t)[\tau_t^r - \tau_t^s]$

Direct transmission is also costly. Let $C(\tau^i)$ denote the cost of the education effort τ^i and $i \in \{r, s\}$. 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 education effort $\tau^i \in [0,1]$ at time t, which maximizes

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

Where V^{ik} is the utility to a parent with preferences i if his child is of type k. Notice that the utility V^{ik} depends on q_{t+1}^E , which denotes the expectation

[6] As it is customary in this class of models we will assume that reproduction is asexual, with a parent per child and thus the population remains constant.

[7] P^{ik} denote the probability that a child of a parent with preferences i is socialized to preferences k

[8] We relegate to the appendix the particular details of this process.

about the proportion of reciprocal investors in period $t+1$ in the population. In this work 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 that in the current period, that is, $q_{t+1}^{E} = q_t$.

Direct transmission is justified because parents are altruistic towards their children. But their socialization decisions are not based on the purely material payoff expected for their children but on the payoff as perceived by their 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, ΔV^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.

Maximizing the above expression with respect to τ^i, $i \in \{r, s\}$, 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 (ΔV^i) and (negatively) on the proportion of their own type in the current preferences distribution in the population.

Substituting the optimal education efforts in 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. As it is shown by Montgomery (2010) this cultural dynamics is analogous to the replicator dynamics if we substitute the material payoffs by the levels of cultural intolerance. Instead of material payoffs, levels of cultural intolerance are the main determinants that govern the dynamic evolution of the preferences distribution in the investors' 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 non-linear differences equation system:

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

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 Fully Cooperative Culture (FCC), any stable steady state of the dynamics where a Fully Cooperative Equilibrium is played. And a similar definition applies for the other equilibria of the game.

Our model yields different long run outcomes, cultures, depending on the particular values of the institutional parameters: λ, γ and z, of the trust-punishment relationship, and also depending on the initial conditions of the dynamics. Some of these cultures 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 where the PBE exists.

Let us start checking the "robustness" of the efficient (Cooperative) Perfect Bayesian Equilibria.

Proposition 8 *If $\lambda > 1/2$, the only efficient culture is the Fully Cooperative Culture and it exists for any pair of distributions (q_t, p_t) such that $q_t \geq \min\left\{\overline{q} = \frac{z}{2\gamma}, q' = \frac{1}{\sqrt{2\lambda}}\right\}$.*

Proof. See Appendix.

We give here a sketch of the proof and defer the rest of the details to the Appendix. Firstly, we show that the Fully Cooperative Equilibrium constitutes a culture and secondly that the other two cooperative equilibria can not be stable steady states of the dynamics, and hence, they never become cultures.

The Fully Cooperative Culture is based on the FCE. In the region in which this equilibrium exists, for any value of p and for high values of q, both types of investors invest and both types of allocators choose to offer $b = 1/2$. Therefore, $V^{ik} = 1/2 - c$ for both types of parents of investors. Hence, the net gains for any type of parent of investor obtained from transmitting their own preferences ΔV^{ik}, that is, their levels of cultural intolerance are zero. Thus, the optimal education effort functions for both types are also zero and there are not incentives at all for socialization. Consequently, the distribution of preferences in the investors population will remain unchanged, that is, $q_{t+1} = q_t$.

Concerning the dynamic evolution in the allocators population, note that the levels of profits of both types of allocators are the same $\Pi_t^f = \Pi_t^s = 1$, since they have the same reward policy, $b = 1/2$, and thus, the preferences 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 is a local attractor of the dynamics and, thus, a culture.

There exist 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 it 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. In the equilibrium in which $b_s = \widehat{b} < 1/2$, $\Pi^s > \Pi^f$, since $\Pi^s = 2(1 - \widehat{b})$ and $\Pi^f = 1$. And in the equilibrium in which $b_s = 0$, also $\Pi^s > \Pi^f$, since $\Pi^s = 2$ and $\Pi^f = 1$. 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 can not constitute long run cultures because the proportion of selfish allocators will increase over time (reducing p) until the society ends up playing the inefficient separating equilibrium. The reason is that the dynamics has reached the region with low values of q and intermediate values of p, and for these lower proportion of fair-minded allocators, the reciprocal investors prefer not to invest.

Next, we turn to check the robustness of the inefficient equilibria of the team trust game.

Proposition 9 *The only inefficient culture is the Non Cooperative Culture and exists for any pair of distributions* (q_t, p_t) *such that* $q_t < \min\left\{\overline{q} = \frac{z}{2\gamma}, q^{'} = \frac{1}{\sqrt{2\lambda}}\right\}$ *and* $p_t < \frac{2c}{1-q_t}$.

Proof. See Appendix.

We also give here a sketch of the proof and defer the details to the Appendix.

Recall that in the Non-cooperative Equilibrium both types of investors do not invest in the project and there is no surplus creation. In this equilibrium the payoff of every player is zero. Thus, the optimal education efforts levels are zero and there are no incentives to socialize. Hence if the society coordinates in this equilibrium, the investors population will remain locked in in 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. That is, if the dynamics reaches this region remains in it. 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. In the inefficient separating equilibrium only the selfish investors choose to invest. And in the quasi cooperative equilibrium, although both types of investors choose to invest, there is punishment because reciprocal investors punish the unfair rewards of selfish allocators. We are going to check that these equilibria can not result in a culture because the dynamics leaves the region in which these equilibria exist.

First, regarding the dynamics of the investors population in the Inefficient Separating Equilibrium region, we obtain that the levels of cultural intolerance of both types are non-negative. 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. As it can be easily computed the levels of cultural intolerance are given by:

$\Delta V^r(q,p) = (1-q)(\alpha - p(0.5+\alpha)) + c \geq 0,$

$\Delta V^s(q,p) = (1-q)p(0.5) - c \geq 0,$

where we have dropped the subindex t for clarity of exposition.

We substitute these levels of cultural intolerance in the socialización effort functions of both types of parents and we equate these functions to obtain the demarcation curve, that is, the locus of pairs (q,p) such that the distribution of preferences in the investors population remains constant over time. This curve is given by the expression $q(p) = (\alpha p - \alpha)q^2 + q(2\alpha - p(2\alpha + 0.5) + (p(\alpha + 0.5) - \alpha) - c = 0$.

Note that for a given p, if $q > q(p)$,then $\tau^{r*}((q,p)) < \tau^{s*}((q,p))$ and q decreases and if $q < q(p)$, then $\tau^{r*}((q,p)) > \tau^{s*}((q,p))$ and q increases. This demarcation curve belongs to the region in which the equilibrium is defined.

In the inefficient separating 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 allocators population in this region is given by the expression $\Delta p_t = p_t(1-p_t)\varphi[-(1-q)^2]$, which is negative for all q. Thus, in all this region, the proportion of fair-minded allocators (p) falls.

Summing up, in the inefficient separating equilibrium region p always decreases and q changes depending on its location, above or below of the demarcation curve. But, in any case 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, the society will remain there because both of them constitute a culture.

A formal analysis showing this result is relegated to the Appendix.

The other inefficient equilibrium is the Quasi Cooperative Equilibrium with punishment, in which both types of investors choose to invest, but only reciprocators punish the unfair rewards set by the selfish allocators. This equilibrium exists for an intermediate proportion of reciprocal investors and a high proportion of fair minded allocators.

In this situation the profits of the selfish allocators are higher than those of the fair minded allocators. The dynamics in this region leads to a decrease in p, because $q^2 2(1-\lambda) = \Pi^s > \Pi^f = 1$.

However, concerning the dynamics in the investors 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 a unfair reward while a selfish investor parent dislikes the behavior of his reciprocal child of punishing an unfair reward and losing material payoffs. The levels of cultural intolerance are given by:

$\Delta V^r(q,p) = (1-p)[q(z/2+\gamma+\alpha\lambda) - z] > 0$

$\Delta V^s(q,p) = (z - q(\gamma + z/2)(1-p) > 0.$

Using the same procedure that in the previous case we obtain the demarcation curve. This curve is given by the expression $-q^2\alpha\lambda + q(z/2 + \gamma + \alpha\lambda) - z = 0$ and it turns out that it is independent of p. In particular it is given by the solution (q'') to the previous quadratic function. If $q > q''$, then $\tau^{r*}((q,p)) > \tau^{s*}((q,p))$ and q increases and if $q < q''$, then $\tau^{r*}((q,p)) < \tau^{s*}((q,p))$, and q decreases.

Eventually, depending on the initial conditions, the dynamics will leave this region either to the FCC region with a very high q or to the Inefficient separating equilibrium 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.

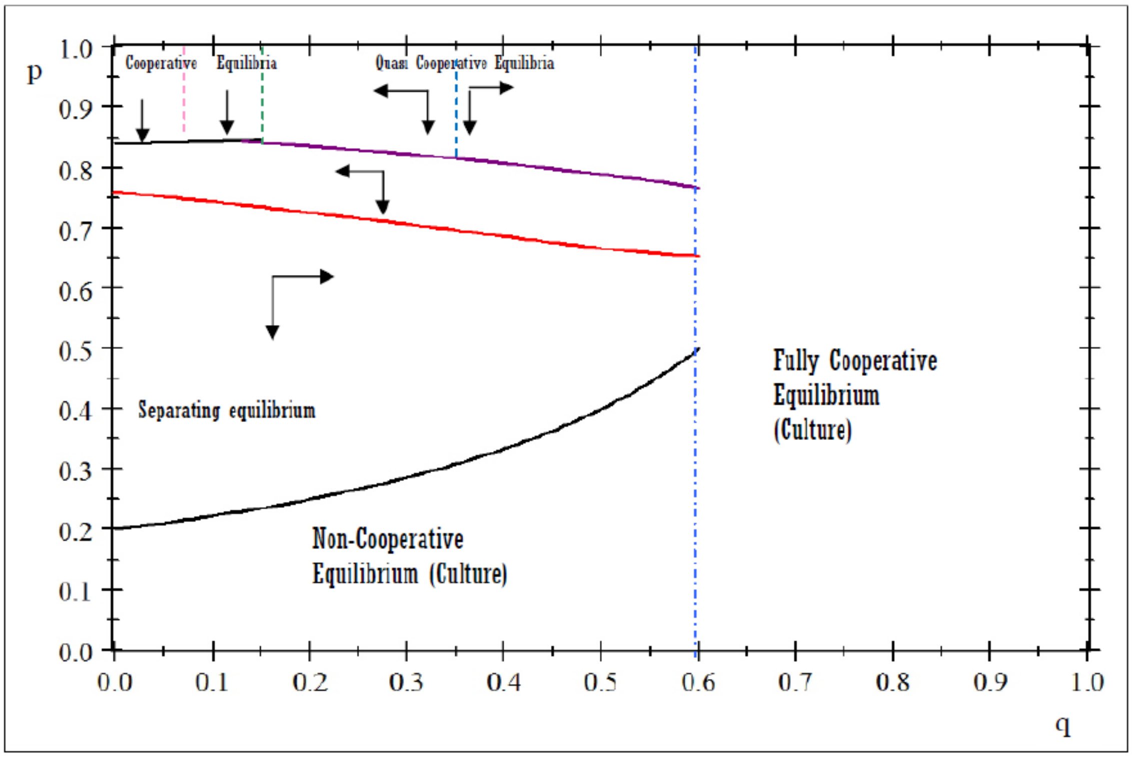
We sum up the results obtained in this section in the following corollary.

Corollary 1 *The only long-run outcomes of the dynamical system 1 are the Fully Cooperative Culture or the Non-Cooperative Culture.*

The next figure 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



p
q
Cooperative Equilibria
Quasi Cooperative Equilibria
Separating equilibrium
Fully Cooperative Equilibrium (Culture)
Non-Cooperative Equilibrium (Culture)

reciprocal investors and by any preference distribution of allocators. On the other hand, the inefficient NCE, that exists for 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. Surprinsingly, there is not observed punishment in any of both cultures. The ultimate reason for this result relies in the credibility of punishment. In the FCC, the threat of punishment is so high and credible that modifies the behavior of selfish allocators, leading them to set fair rewards in order to avoid punishment. This is the unique situation in which the selfish allocators do not have any competitive advantage over the fair minded allocators in term of profits.

However, in the NCC the small number of reciprocators in the investors population generates a situation in which the punishment is not credible at all and therefore the 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 to obtain a cooperative culture with fair returns in the long-run.

Some of the assumptions concerning the punishment institutions play an important role to obtain 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 $\lambda > 0.5$, that is, the damage caused by the coordinated punishment is big enough to make the threat of punishment effective. Otherwise, the punishment can not lead to an increase in the cooperation. The reason is that as $\lambda \leq 0.5$, that is, the inflicted damage is low, the selfish allocator will prefer to set $b_s = 0$ and to be punished with probability one than to offer $b_s = 1/2$ and to avoid the punishment. Therefore the FCE does not exist for $\lambda \leq 0.5$. Nevertheless for a sufficiently high value of p there will exist a quasicooperative equilibrium with punishment. But this will never constitute a culture because p decreases over time, since the profits of the selfish allocators are greater than those of the fair-minded allocators.

Summarizing, if $\lambda \leq 0.5$, the NCC is the unique global attractor of the dynamics.

Second, we want to know the effects if $z/2 > \gamma$, that is, if the level of peer pressure is not enough to compel the selfish investors to punish the allocator for high values of q. The first consequence is that only the reciprocators punish in equilibria. Therefore, the basin of attraction of the

FCC decreases. In particular, if $z/2 > \gamma$, then $\overline{q}$ is greater or equal 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 the reciprocal investors can credibly threat with punishment.

The previous remarks reflect the importance of the punishment institutions for the maintenance of the cooperative culture. The influence of these institutions comes from two sources. The first one comes through the incentive constraint these institutions impose in the short run and the second source comes through an indirect way by altering the incentives to socialize in preferences that enhance negative reciprocity. This, in turn, will reinforce the effectiveness of the punishment institutions in the long run. Therefore, a high proportion of reciprocators in future generations will provide the necessary complement to the punishment institutions in order to obtain cooperative behavior.

Some comments on the influence of the degree of aversion to disadvantageous inequality α and the cost of the investment c are necessary. The basin of attraction of the FCC will be greater the greater is α and the smaller is c. But for any $c > 0$, both the fully cooperative and the non-cooperative cultures can only be local attractors. That is, the long run culture that prevails in the society depends on the initial condition of the dynamics. Therefore, uniqueness is not achieved, unless c is zero.

8 Concluding Remarks.

The main result of this model is that cooperation only evolves and maintains if there is enough punishment capacity in the society and there are enough individuals willing to implement both the coordinated and the peer punishment. The Fully Cooperative Culture 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 investors population. This fact illustrates the main difficulty of obtaining cooperation: uniqueness is not achieved. Our model shows hysteresis, initial conditions matter because they can lead the society to a different state in the long run. If the society is able to build strong punishment institutions and can accomplish, through socializacion mechanisms, a preference distribution in the population willing to implement these institutions, then the society can settle in a cooperative and efficient culture. However, if it is not able to reach a "sufficient" proportion of reciprocators in the society, a non cooperative and inefficient culture will be established.

Changes in the punishment institutions as, for instance, in the damage caused by the coordinated punishment or the level of peer pressure, might cause large changes in the long-run distribution of preferences and behaviour (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 the incentives to socialize future generations of investors in a kind of preferences more prone to use any sort of punishment.

Finally, our simple model provides some clear predictions that can be tested with the data (either in the laboratory or with the information contained in, for instance, the World Values Survey). For example, we provide a rationale for the existence of a positive relationship between the investors´ institutional power (measured by their capacity for and their cost of punishing hold up) and their level of cooperation/trust (measured by their aggregated amount of investment when they risk being held up). This relationship is well documented in Aghion, Algan and Cahuc (2011).

9 Appendix

Proof of Proposition 1.

Suppose that the allocator offers $b < 1/2$.

To show that the profile (p, np) is a BNE for $\mu \in [\mu^*(b),\ \widetilde{\mu} = \frac{z}{z/2+\gamma}]$ note that for reciprocal investors, choosing punishment is a best response to (p, np) when:

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

$\mu(b - \gamma - \alpha[(1-b)(1-\lambda) - b)] + (1-\mu)[b - \alpha((1-b) - b)).$

That is, if

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

Therefore, if $\mu \geq \mu^*(b)$, the reciprocal types choose to punish an unfair offer of the allocator.

For selfish investors, not punishining the allocator is a best response against (p, np) 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 selfish investors choose punishment as a best response to (p,p).That is, when:

$(b - z/2) \geq \mu(b-\gamma) + (1-\mu)b$.

That is, if

$$\mu \geq \overline{\mu} = \frac{z}{2\gamma}.$$

Summarizing, for $\mu \in [\mu^*(b), \widetilde{\mu}]$, (p, np) is a BNE and for $\mu \in [\overline{\mu}, 1]$, (p,p) is a BNE.

But as $\overline{\mu} < \widetilde{\mu}$, for $\mu \in [\overline{\mu}, \widetilde{\mu}]$ there are two BNE: *(p,np)* and *(p,p)*, however we assume that players will select the equilibrium with the highest probability of punishment , that is, (p,p).

It is easy to check that the profile (np, np) is a BNE for every value of μ. Note that for the reciprocal investor: $b-\alpha[(1-b)-b)] \geq b-z-\alpha[(1-b)-b)]$ and for the selfish investor $b \geq b - z$.

■

Proof of Lemma 1.

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, her maximal possible payoff. This is her optimal return policy.

■

Proof of Lemma 2.

For any given $\widehat{\mu} \in [\mu^*(0), \mu^*(0.5)]$, there exists 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) so the reciprocators would punish. However, for $b \in [\widehat{b}, 0.5]$, $\widehat{\mu} \leq \mu^*(b)$ which implies that the BNE is (np, np) in which no type of investor punishes.

Therefore, the alternatives 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 for the first alternative is $\Pi^s(0) = 2(1 - \mu^2\lambda)$ and the expected payoff for the second alternative is $\Pi^s(\widehat{b}) = 2(1 - \widehat{b})$. That is, $\Pi^s(\widehat{b}) = 2(\frac{z-\widehat{\mu}(\gamma+z/2)}{\widehat{\mu}\alpha\lambda})$.

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 exists 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, there exist three real and unequal roots, one of them is negative. We will call the positive roots : $\overleftrightarrow{\mu}_1$ and $\overleftrightarrow{\mu}_2$. Both roots, $\overleftrightarrow{\mu}_1$ and $\overleftrightarrow{\mu}_2$ are greater than $\mu^*(0)$.Then the optimal reward policy of the allocator is:

If $\widehat{\mu} \in [\mu^*(0), \overleftrightarrow{\mu}_1]$ and $\widehat{\mu} \in [\overleftrightarrow{\mu}_2, \mu^*(0.5)]$ the best policy is to offer $b_s = \widehat{b}$ and if $\widehat{\mu} \in [\ \overleftrightarrow{\mu}_1, \overleftrightarrow{\mu}_2]$ the best policy is to offer $b_s = 0$.

■

Proof of Lemma 3.

For any given $\widehat{\mu} \in [\mu^*(0.5, \overline{\mu}]$, then $\mu^*(b) < \overline{\mu}$ for all $b_s \in [0, 0.5)$ and the BNE of $CP(\mu, b)$ is (p, np) in which only the reciprocators punish. Then, if the selfish allocator sets $b_s = 0.5$ her profits are $\Pi^s(0.5) = 1$. The alternative is to set $b_s = 0$, which results in an expected profit of $\Pi^s(0) = 2(1-\mu^2\lambda)$. Therefore to set $b_s = 0.5$ is better than to set $b_s = 0$, if $1 \geq 2(1-\mu^2\lambda)$. That is, when $\mu \geq \mu' = \frac{1}{\sqrt{2\lambda}}$. Note that $\mu' < 1$ only if $\lambda > 0.5$.

Summarizing, whenever $\widehat{\mu} \in [\mu^*(0.5, \overline{\mu}]$,if $\widehat{\mu} \geq \mu' = \frac{1}{\sqrt{2\lambda}}$, the best policy is to offer $b_s = 0.5$, and if $\mu < \mu'$, the best policy is to offer $b_s = 0$.

If $\lambda \leq 0.5$, then $\mu' \geq 1$, and then it always holds that $\mu < \mu'$ and the best policy is to offer $b_s = 0$.

■

Proof of Lemma 4.

If the allocator offers $b_s < 1/2$ there will be punishment by both types of investors, since the BNE played of $CP(\mu, b)$ is (p, p), and her expected payoff will be $\Pi^s(0) = 2(1-\lambda)$. In contrast, if she offers $b_s = 1/2$, there will no punishment and her payoff will be 1. Therefore, if the allocator compares $\Pi^s(0) = 2(1-\lambda)$ with $\Pi(0.5) = 1$, the result depends on the value of λ.

Hence, for any given $\widehat{\mu} \in [\overline{\mu}, 1]$ and $\lambda > 0.5$, the best return policy is to offer $b_s = 0.5$, but if $\lambda \leq 0.5$, the best return policy is to offer $b_s = 0$.

■

Equilibria of the Team Trust Game with Cooordinated Punishment.

Proof of Proposition 2.

The fully cooperative equilibrium.

Suppose $\lambda > 0.5$, the selfish allocator will set $b_s = 1/2$ if, by lemma 4, $q \geq \overline{q}$ and if, by lemma 3, $q \geq q'$. Therefore this pooling equilibrium exists whenever $q \geq \min(\overline{q} = \frac{z}{2\gamma}, q^{'} = \frac{1}{\sqrt{2\lambda}})$. The beliefs *are* $\mu(r/I, I) = q$, since there is no updating in this sort of equilibria. As both types of investors invest, the surplus is generated for sure and there is no punishment since both types of allocators set a fair reward. The payoff for any member of the team is $1/2 - c$ and any type of allocator gets a payoff of 1. No player has a profitable deviation.

■

Proof of Proposition 3.

A cooperative equilibrium with $b_s = \widehat{b}$.

By lemma 2, if $q_t \in [q^*(0), q^*(0.5)]$ the selfish allocators set $b_s = \widehat{b} = \frac{q(\alpha\lambda+\gamma+z/2)-z}{q\alpha\lambda} < 1/2$.

To obtain this equilibrium, note that to invest is better than not to invest, when the following incentive compatibility restriction holds for a selfish type of investor:

$p(0.5) + (1-p)\widehat{b} - c \geq 0$, and therefore, $p \geq \frac{c-\widehat{b}}{0.5-\widehat{b}}$

The incentive compatibility constraint for a reciprocal investor is:

$p(0.5) + (1-p)[\widehat{b} - \alpha(1-2\widehat{b}) - c \geq 0$, and therefore $p_t \geq \frac{c-(\widehat{b}-\alpha(1-2\widehat{b}))}{0.5-(\widehat{b}-\alpha(1-2\widehat{b}))}$. This latter one is the binding condition.

Note also that by Proposition 1 no punishment is chosen in equilibrium. The payoff of the fair-minded allocator is 1 and the expected payoff of the selfish one is $2(1-\widehat{b})$.

■

Proof of Proposition 4.

A cooperative equilibrium with $b_s = 0$.

By lemma 1 if $q_t < q^*(0)$ the selfish allocators set $b_s = 0$.

As in the previous case, to invest is better than not to invest when the following incentive compatibility restriction holds for a selfish type of investor:

$p(0.5) - c \geq 0$, and therefore, $p \geq 2c$.

The incentive compatibility constraint for a reciprocal investor is:

$p(0.5) + (1-p)(-\alpha) - c \geq 0$, and therefore $p_t \geq \frac{\alpha+c}{\alpha+0.5}$. This latter one is the binding condition.

Then both types of investors invest and the payoff of the fair allocator is 1 and the payoff of the selfish allocator is 2. Note also that by Proposition 1 no punishment is chosen in equilibrium.

■

Proof of Proposition 6.

The Inefficient Separating Equilibrium where selfish investors choose to invest.

In this equilibrium the surplus is generated with a probability $(1-q)^2$ when the team is formed by two selfish investors. The selfish allocator offers $b_s = 0$ and her expected payoff is $2(1-q)^2$, while the fair-minded allocator offers $b_f = 1/2$ and her expected payoff is $(1-q)^2$.

The beliefs are $\mu(r/(I,I) = 0$, that is, if there is investment by both members of the team then it is believed for sure that the team is composed by two selfish investors.

To obtain this equilibrium the following incentive compatibility restriction has to hold for a selfish type of investor:

$(1-q)p(0.5) - c \geq 0$, and therefore, $p \geq \frac{2c}{1-q}$.

The incentive compatibility constraint for a reciprocal investor is:

$0 \geq (1-q)[p(0.5) + (1-p)(-\alpha)] - c$, and therefore $p \leq \frac{c+\alpha(1-q)}{1-q}$.

The set of pairs (q,p) that satisfies both incentive compatibility constraints is not empty.

■

The Inefficient Separating Equilibrium where reciprocal investors choose to invest.

There exists another separating equilibrium in which only reciprocal investors decide to invest and the selfish investors decide not to invest if $q = 2c$. In this equilibrium both types of allocators offer $b = 0.5$ and the beliefs are $\mu(r/(I,I) = 1$, that is, if there is investment by both members of the team then it is believed for sure that the team is composed by two reciprocal investors.

To obtain this equilibrium the following incentive compatibility restriction has to hold for selfish types of investors:

$0 \geq q(0.5) - c$ and therefore, $2c \leq q$.

The incentive compatibility constraint for a reciprocal investor is:

$q(0.5) - c \geq 0$, and therefore $q \geq 2c$.

Therefore, for $q = 2c$ both incentive compatibility 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 the punishment, than offering $b = 0.5$ and avoiding the punishment. The fair minded allocator sets $b_f = 0.5$. However, note that now both incentive compatibility constraints do not hold simultaneously.

The constraint for the selfish type is $0 \geq q[p(0.5)+(1-p)(0-z/2)]-c$ and the constraint for the reciprocal type is $q[p(0.5)+(1-p)(-z/2-\alpha(1-\lambda)]-c \geq 0$.

■

Proof of Proposition 7

A Quasi Cooperative Equilibrium with Punishment.

In this pooling equilibrium μ is equal to q.We know by lemma 2 that $b_s = 0$ is preferred to $b_s = \widehat{b}$, when the following cubic 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 exist at most two positive unequal roots. We call the positive roots: $\overleftrightarrow{q}_1$ and $\overleftrightarrow{q}_2$, where $\overleftrightarrow{q}_1 < \overleftrightarrow{q}_2$. Thus if $\widehat{q} \in [\overleftrightarrow{q}_1, \overleftrightarrow{q}_2]$ the best reward policy of the selfish allocator is to set $b_s = 0$. We can show by numerical simulations that $\overleftrightarrow{q}_1$ is always greater than $q^*(0)$ and that $\overleftrightarrow{q}_2$ is always greater than 1.

On the other hand, by lemma 3 if $q_t \in (q^*(0.5), \min(\overline{q}, q'))$ the best reward for the selfish allocator is to set $b_s = 0$, even though she will be punished by the reciprocal investors acording to Proposition 1.

Therefore, for $q_t \in [\overleftrightarrow{q}_1, \min(\overline{q}, q')]$ the selfish allocator sets $b_s = 0$.

Choosing to invest is better than choosing not to invest, for selfish investors, when the following incentive compatibility restriction holds:

$p(0.5) + (1-p)[-q\gamma] - c \geq 0$, and therefore, $p \geq \frac{q\gamma + c}{q\gamma + 0.5}$.

The incentive compatibility constraint for a reciprocal investor is:

$p(0.5)+(1-p)[(1-q)(-z-\alpha)+q(-z/2-\alpha(1-\lambda)]-c \geq 0$, and therefore $p_t \geq \frac{z+\alpha+c-q(z/2+\alpha\lambda)}{z+\alpha+0.5-q(z/2+\alpha\lambda)}$.

This last expression is the binding incentive compatibility constraint and determines the existence of the equilibrium.

The payoff of the fair-minded allocator is 1 and the expected payoff of the selfish allocator is $2(1-\lambda q^2)$.

■

Transition Probabilities of the Socialization Process.

Let P^{ik} denote the probability that a child of a parent with preferences i is socialized to preferences k. The socialization mechanism is then characterized by the following transition probabilities where q_t is the proportion of reciprocal investors:

$P_t^{ss} = \tau_t^s + (1-\tau_t^s)(1-q_t)$

$P_t^{sr} = (1-\tau_t^s)q_t$

$P_t^{rr} = \tau_t^r + (1-\tau_t^r)q_t$

$P_t^{rs} = (1-\tau_t^r)(1-q_t)$

Given these transition probabilities it is easy to characterize the dynamic behavior of q_t:

$q_{t+1} = [q_t P_t^{rr} + (1-q_t)P_t^{sr}]$

Substituting we obtain the following equation on differences:

$q_{t+1} = q_t + q_t(1-q_t)[\tau_t^r - \tau_t^s]$

Cultures in the Long Run.

Proof of Proposition 8

The *Dynamics in the Cooperative Equilibrium with $b_s = \widehat{b}$:*

The *allocators population dynamics.*

The payoff of the fair 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 $\widehat{b} = \frac{q(\alpha\lambda+\gamma+z/2)-z}{q\alpha\lambda} < 1/2$.

And the dynamics of the allocators 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 $\widehat{b} < 1/2$ as it is the case, so p decreases, $\forall q$.

The *investors population dynamics.*

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 are zero because $V^{rr} = V^{rs}$ and $V^{ss} = V^{sr}$, and the optimal education efforts are trivially zero: $\tau^{r*}(q_t) = \tau^{s*}(q_t) = 0$, and then q does not change.

Therefore, in this region p always decreases and q does not change.

The dy*namics in the Cooperative Equilibrium with $b_s = \widehat{0}$.*

The allocators population dynamics.

The payoff of the fair allocator is $\Pi^f = 1$ and the payoff of the selfish allocator is $\Pi^s = 2$.

The dynamics of the allocators population is given by $\Delta p_t = p_t(1 - p_t)\varphi(-1) < 0$, so p decreases, $\forall q$.

The *investors population dynamics.*

The payoff of the reciprocal investor is $U_r = p(0.5) + (1-p)(-\alpha) - c \geq 0$.

The payoff of the selfish investor is $U_s = p(0.5) - c \geq 0$.

And again in this equilibrium the levels of cultural intolerance are zero because $V^{rr} = V^{rs}$ and $V^{ss} = V^{sr}$ and the optimal education efforts are trivially zero and then q does not change.

Therefore, in this region p always decreases and q does not change.

It can be checked that the two critical values on p that define the boundary 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 Dynamics in the Inefficient Separating Equilibrium

The allocators population dynamics.

In this equilibrium the payoff of the fair allocator 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: $\Delta p = p(1-p)\varphi(-(1-q)^2) < 0$.Thus, p decreases.

The *investors population dynamics.*

The payoff of the reciprocal investor is $U_r = 0$ and the payoff of the selfish investor is $U_s = (1-q) \cdot p \cdot 0.5 - c$

In this equilibrium the levels of cultural intolerance are non negative:

$\Delta V^r = V^{rr} - V^{rs} = 0 - ((1-q)(p(0.5) + (1-p)(-\alpha)) - c \geq 0.$

$\Delta V^s = V^{ss} - V^{sr} = (1-q)p(0.5) - c \geq 0.$

And the optimal education efforts are given by:

$\tau^{r*}(q,p) = k\Delta V^r(q)(1-q) = k(1-q)(c - (1-q)(p(0.5) + (1-p)(-\alpha)).$
$\tau^{s*}(q,p) = k\Delta V^s(q)q = kq((1-q)p(0.5) - c) \geq 0.$

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 increases. This demarcation curve belongs to the region in which the equilibrium is defined.

The Dynamics in the Quasi Cooperative Equilibrium with punishment: The allocators population dynamics.

In this equilibrium the payoff of the fair minded allocator 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 allocators population is given by $\Delta p = p(1-p)\varphi(2q^2\lambda - 1)$. This expression is negative when $q \leq q'$, as it is the case, so p decreases, $\forall q$.

The investors population dynamics.

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: $U_s = p(0.5) + (1-p)[-q\gamma] - c \geq 0$.

In this equilibrium 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.

$\Delta V^s = V^{ss} - V^{sr} =$
$p(0.5) + (1-p)[-q\gamma] - c - ((0.5)p + (1-p)((1-q)(-z) + q(-z/2) - c) =$
$(z - q(\gamma + z/2)(1-p) > 0$, this expression is positive because $q < z/(\gamma + z/2)$.

The optimal education efforts are given by:

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

Therefore, both types of parents have incentives to socialize. We obtain the demarcation curve 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 that is independent of p. In particular is given by the solution (q'') to the previous quadratic function. In particular, if $q > q''$, $\tau^{r*}((q,p)) > \tau^{s*}((q,p))$ and q increases and if if $q < q''$, $\tau^{r*}((q,p)) < \tau^{s*}((q,p))$, and q decreases.

Eventually, depending on the initial conditions, the dynamics will leave this region either to the FCC region with a very high q or to the Inefficient separating equilibrium region with a smaller q. In any case, the dynamics abandons this region and thus this equilibrium can not constitute a culture in the long run.

■

References

AGHION , ALGAN Y. and CAHUC, P. (2011) Civil Society and the State: The Interplay Between Cooperation and Minimum Wage Journal of the European Economic Association, 9(1), 3-42.

ALGAN Y. and CAHUC, P.(2009) Civic Virtue and Labor Market Institutions, *American Economic Journal*: *Macroeconomics*, **1**(1), 111-145.

ANDREONI, J., HARBAUGH,W., VESTERLUND, L. (2003) The carrot or the stick: Rewards, punishment and cooperation, *American Economic Review* **93** (3), 893–902.

ARROW, K. (1972) Gifts and exchanges, Philosophy and Public Affairs (1).

BARON, J. and KREPS, D. (1999) *Strategic Human Resources*, Wiley Ed., New York.

BISIN, A., and VERDIER, T. (1998), On the Cultural Transmission of Preferences for Social Status, *Journal of Public Economics*, **70**, 75-97.

BISIN, A., and VERDIER, T. (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 VERDIER, T. (2001) The Economics of Cultural Transmission and the Dynamics of Preferences, *Journal of Economic Theory*, **97**, 2, 298-319.

BOLTON G. and OCKENFELS A. (2000) ERC: A Theory of Equity, Reciprocity and Competition, *American Economic Review*, **90**, 166 - 193.

BOYD, R., GINTIS, H., BOWLES, S., RICHERSON, P.J. (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. and RICHERSON, P., 1985, *Culture and the Evolutionary Process,* Chicago: University of Chicago Press.

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

CARPENTER, J. (2007) The demand for punishment, *Journal of Economic Behavior and Organization* 62, 522–542.

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

DUFWENBERG, M. and KIRCHSTEIGER, G., (2004), A Theory of

Sequential Reciprocity, *Games and Economic Behavior* 47 (2004), 268-98.

DUFWENBERG, M, SMITH, A. and VAN ESSEN, M. (2011) Hold-up: with a vengeance, *Economic Inquiry,*

EGAS, M. & RIEDL, A., (2008), The economics of altruistic punishment and the maintenance of cooperation, *Proceeds of the Royal Society*, 275, 871–878.

ERTAN, A., PAGE, T. and PUTTERMAN, L. (2009), Who to punish? Individual decisions and majority rule in mitigating the free rider problem, *European Economic Review*, 53, 495–511.

FALK, A., FEHR, E. and FISCHBACHER,U., (2005), Driving Forces Behind Informal Sanctions, *Econometrica*, 7 (6), 2017-2030.

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

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

FEHR, E. and ROCKENBACH, B., (2003), Detrimental effects of sanctions on human altruism, *Nature* 422, 137–140 (2003)

FEHR, E. and SCHMIDT, K., (1999), A Theory of Fairness, Competition and Cooperation, *Quarterly Journal of Economics*, 114, 817-868.

FOWLER, J. H., (2005), Altruistic punishment and the origin of cooperation, *Proceedings of the National Academy of Sciences of the United States of America* 102, 7047–7049.

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

GARDNER, A. and WEST, S. A., (2004), Cooperation and punishment, especially in humans, *American Naturalist* 164, 753–764.

GÜRERK, O., IRLENBUSH, B. & ROCKENBACH, B.,2006, The competitive advantage of sanctioning institutions, *Science* 312, 108–111

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

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

HERRMANN, B., THOENI, C. & GÄCHTER,S., (2008) Antisocial punishment across societies, *Science* 319, 1362–1367

HOUSER, D., XIAO, E., McCABE, K., SMITH, V., (2005) When punishment fails: Research on sanctions, intentions and non-cooperation, *Econ-WPA, series Experimental*, No. 0502001

HUCK, S. and KOSFELD, M., (2007), The Dynamics of Neighbourhood Watch and Norm Enforcement, *The Economic Journal*, 117, 270-286.

KIYONARI, T., BARCLAY, P., WILSON, M. & DALY, D., (2004), Second order punishment in one-shot prisoner's dilemma, *International Journal of Psychology*, 39, 329–334.

KOSFELD, M., RIEDL, A. & OKADA, A., (2009), Institution formation in public goods games, *American Economic Review*, 99, 1335–1355.

MONTGOMERY, J., (2010), Intergenerational Cultural Transmission as an Evolutionary Game, *American Economic Journal: Microeconomics*, 2(4): 115–36.

NAKAMARU, M. & IWASA, Y., (2006), The evolution of altruism and punishment: role of the selfish punisher, *Journal of Theoretical Biology*, 240, 475–488.

NIKIFORAKIS, N.S.,(2008), Punishment and counter-punishment in public goods games: can we really govern ourselves?, *Journal of Public Economics* 92, 91–112.

OSTROM, E.,WALKER, J., GARDNER, R., (1992), Covenants with and without a sword: self-governance is possible, *American Political Science Review* 86, 404–417.

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

OSTROM, E., (2000), Collective Action and the Evolution of Social Norms", *The Journal of Economic Perspectives*, 14, No. 3. pp. 137-158.

PUTNAM, R., (2000), "Bowling Alone: The Collapse and Revival of American Community", New York: Simon and Schuster.

ROB, R, and ZEMSKY, P., (2002), Cooperation, Corporate Culture and Incentive Intensity, *RAND Journal*, 33, 243-257.

ROCKENBACH, B. & MILINSKI, M., (2006), The efficient interaction of indirect reciprocity and costly punishment, *Nature* 444, 718–723.

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

TABELLINI, G., (2008), The Scope of Cooperation: Values and Incentives, *The Quarterly Journal of Economics*, 123(3), 905-950.

YAMAGISHI, T., 1986, The provision of a sanctioning system as a public good, *Journal of Pers. Soc. Psychol.* 51, 110–116.