Carrot or Stick?
Group Selection and the Evolution of Reciprocal Preferences*

Florian Herold†

This version: June 28, 2003

Abstract

This paper studies the evolution of both characteristics of reciprocity - the willingness to reward friendly behavior and the willingness to punish hostile behavior. Firstly, preferences for rewarding as well as preferences for punishing can survive evolution provided individuals interact within separated groups. This holds even with randomly formed groups and even when individual preferences are unobservable. Secondly, preferences for rewarding survive only in coexistence with self-interested preferences. But preferences for punishing tend either to vanish or to dominate the population entirely. Finally, the evolution of preferences for rewarding and the evolution of preferences for punishing influence each other decisively. The existence of rewarders enhances the evolutionary success of punishers, but punishers crowd out rewarders.

JEL-Classification: C72, D63, D64, D83

Keywords: Reciprocity, Evolution of Preferences, Group Selection, Co-evolution, Fairness

1 Introduction

This paper addresses three questions concerning the evolution and co-evolution of both characteristics of reciprocity - the willingness to reward friendly behavior and the willingness to punish hostile behavior: 1) How can preferences for rewarding and preferences for punishing survive the evolutionary competition with purely self-interested preferences? 2) What structural differences distinguish the evolution of the willingness to reward from the evolution of

*I am very grateful to Klaus M. Schmidt, Theodore C. Bergstrom, Robert Evans and Christoph Kuzmics for thorough reading of the paper and very valuable suggestions. I would like to thank Florian Englmaier, Sten Nyberg, Pedro Rey Biel and Ferdinand von Siemens for helpful comments and discussions. Funding by the EDGE-Program and a Marie-Curie-Fellowship are gratefully acknowledged.

†Contacts: Department of Economics, University of Munich, Ludwigstr. 28 (Rgb.), D-80799 Munich, Germany, Tel.: +49 89 2180-2107, Fax.: +49 89 2180-3510 email: florian.herold@lrz.uni-muenchen.de

the willingness to punish? 3) How is the evolution of one side of reciprocity influenced by the evolution of the other side?

Self-interested preferences are a standard assumption in economic theory. But several experimental studies offer substantial evidence that at least some people are not exclusively driven by self-interest. A significant number of people are willing to reward friendly and/or to punish hostile behavior of an opponent even if this is costly and does not maximize their own material payoffs[1]. From an evolutionary standpoint[2] the observed coexistence of reciprocal and self-interested preferences seems puzzling: A rational self-interested individual can always mimic reciprocal behavior, if thereby, he maximizes his expected material payoff. But in the event of his action being unobservable or punishment being impossible, he behaves selfishly and receives a higher material payoff. Apparently, self-interested individuals should always outperform reciprocal individuals and reciprocal preferences should vanish due to natural selection.

This paper shows that preferences for rewarding as well as preferences for punishing can survive evolutionary competition with purely self-interested preferences if players interact within separated groups and if they can condition their strategy on the distribution of preferences within their own group. This holds even if individual preferences are unobservable, groups are formed randomly and players interact anonymously in random pairings.

However, there are crucial structural differences between the evolution of preferences for rewarding and the evolution of preferences for punishing. Rewarders can successfully invade a population of self-interested players. But they cannot drive them out completely. Preferences for rewarding survive only in coexistence with self-interested preferences. But, preferences for punishing either drive out self-interested preferences or they die out themselves. The option to punish hostile behavior results either in a "culture of punishment" - where all players are willing to punish hostile behavior - or in a "culture of laissez faire" - where nobody is willing to incur the costs of punishing.

The co-evolution of both aspects of reciprocity influences our results decisively if there is an option to reward friendly behavior as well as an option to punish hostile actions. Rewarders enhance the evolutionary success of preferences for punishing, but punishers tends to crowd out preferences for rewarding. In fact, rewarders may serve as a catalyst for the evolution of punishers. Rewarders can invade a population of self-interested types. Their existence can enable punishers to invade successfully and finally to crowd out self-interested as well as rewarding types.

Our results are driven by the marginal effect a player has on the distribution of preferences within his group. This marginal effect is advantageous for a reciprocator and can outweigh

[1]For a survey of the experimental literature see Fehr and Gächter [9] or Fehr and Schmidt[10].

[2]The process of evolution may be interpreted in terms of biological- as well as cultural evolution or even as a process of learning. Under the weak assumptions described below, our results hold independent from the interpretation. Therefore, we delay the discussion of the relevance of each interpretation to Section 3.

the costs of rewarding or punishing. To see this consider pairwise interactions of the following structure: A first moving player *(player 1)* may either cooperate or defect. Cooperation is costly for player 1 but profitable for a second moving player *(player 2)*. Player 2 observes this action and can then reward and/or punish player 1 (both is costly) or remain inactive. Player 1 cooperates only if he expects player 2 to be reciprocal with a sufficiently high probability. Since individual preferences are unobservable player 2 estimates the probability of meeting a certain type by the fraction of this type in his group. Hence, players 1 cooperate if the number of reciprocal players in their group is above a certain threshold, otherwise they defect. Having reciprocal preferences leads to a material advantage for player 2 when he is pivotal in his group, i.e. his type is decisive in whether the number of reciprocal preferences in his group is just above or just below the threshold for cooperation.

Under which circumstances outweighs this material advantage the losses incurred for rewarding or punishing? The intuition for our main result is derived from the following observation: The hope for reward as well as the fear of punishment can induce players 1 to cooperate. But, when most players 1 cooperate it is relatively expensive for player 2 to reward cooperation whereas the willingness to punish is almost for free. On the other hand when most players 1 defect, the willingness to reward is almost for free, whereas it is expensive to punish defection. A higher fraction of rewarders or punishers leads to a higher fraction of groups in which cooperation occurs. Therefore, rewarders are relatively successful when most players 2 are self-interested, whereas punishers become more successful the more players 2 have reciprocal preferences.

The question as to how reciprocity or social preferences can survive evolution has been tackled by different authors from biology, psychology, economics and other social sciences. Sethi and Somanathan [30] have reviewed this literature recently, in which they classify existing explanations for the survival of reciprocity in sporadic interactions into three basic mechanisms[3] [4]: *commitment, assortation,* and *parochialism*.

Commitment: If preferences are observable reciprocal preferences may serve as an advantageous commitment device. A reciprocal player is credibly committed to reward friendly or punish unfriendly behavior. Therefore, he may induce friendly behavior of a first-moving player. This may enhance his evolutionary success. The results of Güth and Yaari [15], Güth [14], Bester and Güth [4] and partly Sethi [27] and Höffler [17] are based on this argument.

Assortation: Efficiency enhancing behavior becomes evolutionary more successful, if play-

[3]They also survey "repetition" as a fourth explanation in non-sporadic interactions.

[4]Huck and Oechssler [19] find a further mechanism to explain how preferences for punishing unfair behavior might survive evolution. They look at ultimatum games in which costs of punishing unfair behavior are very small compared to the punishment and the inverse group-size. In the role of a proposer punishers have the relative advantage over materialists in their group, that they are slightly less likely to be matched with a punisher. Therefore, unfair offers of materialists are more likely to be rejected. In the role of responders punishers have the disadvantage of incurring the costs of punishing. But if these costs are sufficiently small this disadvantage is more than compensated by the relative advantage when being in the role of a proposer.

ers are not matched randomly, but interact with higher probability with players of their own type. In particular, most of the literature on group selection focuses on this idea to explain the evolutionary survival of social preferences. Price [25] first offered a mathematical description. Bergstrom [2] investigated the relation between assortative matching and the evolution of cooperation. Notice that initially random groups may become assortative over time by the evolution of preferences inside groups if no reshuffling of groups occurs[5] (compare e.g. Cooper and Wallace [8]). For surveys of this literature on assortative group selection see Sober and Wilson [31] and more recently Bergstrom [3].

Parochialism: Types that act to enhance efficiency if they are in a group of mainly their own type and act to reduce efficiency if they are in a group of mainly self-interested players may survive evolution even if the matching is non-assortative. Sethi and Somanathan [29] showed that conditional altruists who behave friendly towards other altruists but spiteful towards materialists may be more successful than pure materialists. Similarly Gintis [13] looked at conditional punishers, who punish defectors only if there are enough other punishers in their group. They also survive evolution[6].

This paper contributes to the literature by weakening the degree of observability required to explain the survival of social preferences. The explanation of this paper for the survival of reciprocal preferences is related to the idea of *commitment*. But individual preferences need not be observable[7]. Players observe only the overall distribution of preferences within their own group. The marginal effect of a player on the distribution of preferences in his group drives our results[8]. Even more importantly, our setting allows the analysis of both sides of reciprocity in a unified framework. We find crucial structural differences between the evolution of preferences for rewarding and the evolution of preferences for punishing. Finally, our framework enables us to demonstrate that the co-evolution of both sides of reciprocity influences the results decisively.

The remainder of the paper is organized as follows: Section 2 presents the model and analyzes the three cases which arise naturally: 1) Player 2 might have only the costly option to reward cooperation. 2) Player 2 might have only the costly option to punish defection. 3) Player 2 might have the options to reward cooperation as well as the option to punish

[5] A different endogenous justification of assortative matching arises if preferences are partly observable, see e.g. Frank [11].

[6] Notice that punishers may enhance efficiency, because they can induce cooperation of a first-moving player.

[7] Bowles and Gintis [6] and Friedman and Singh [12] consider also the case when individual preferences are unobservable for the evolution of types who punish non-cooperative behavior . But, in contrast to our paper, the results of Bowles and Gintis depend critically on the assumption that there does not exist a type of player who cooperates but refuses to punish non-cooperative behavior. Friedman and Singh need implicitly the assumption that second-order punishment (i.e. the punishment of non-punishers) is costless. See Sethi and Somanathan [30] for a more detailed discussion of these papers.

[8] A similar effect plays a role in the model by Höffler [17]. He considers a learning process of bounded rational workers in a stylized principal agent model. In equilibrium some agents play fair and other don't - similar to the coexistence result in case 1 of our model.

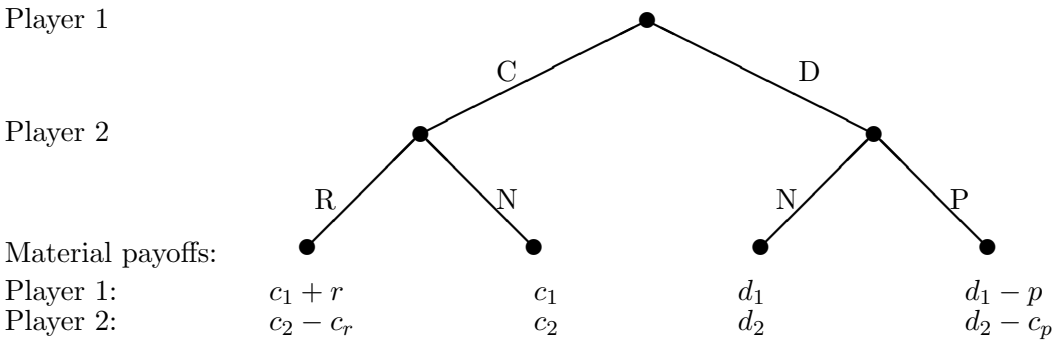
defection. Section 3 discusses our results and finally Section 4 concludes.

2 The Model

We use the indirect evolutionary approach[9] to describe the evolution of preferences: Individuals may have different preferences. We only impose the restriction that preferences can be described by subjective utilities for each possible outcome. The subjective utility an individual assigns to an outcome may not coincide with the material payoff he receives. Individuals choose their strategies according to their own preferences and their knowledge about preferences of their opponents, i.e., they play perfect Bayesian equilibria. They receive material payoffs according to played strategies. A type who receives higher material payoffs has more offsprings (or imitators) and his fraction grows. Subjective utilities of an individual are only important to determine his actions. The evolutionary success is only influenced by the resulting material payoffs.

We consider pairwise sequential interactions of the following structure: *Player 1* moves first. He can either cooperate (C) or defect (D). Cooperation leads to a material gain for *player 2* ($c_2 > d_2$) but is costly for player 1 ($d_1 > c_1$). Player 2 observes the action of player 1 and then chooses his reaction. Three cases are analyzed. In case 1, player 2 can reward cooperation of player 1 (by the amount of r) but this is costly (costs c_r) . In case 2, player 2 can punish defection of player 1 (by the amount p) which is also costly (costs c_p). In case 3, player 2 can do both, either reward or punish player 1. The interaction of case 3 is illustrated in figure 1. Case 1 and case 2 are obtained from this figure by removing the option to punish or the option to reward respectively.

Figure 1: Interaction



with $c_1 + r > d_1 > c_1 > d_1 - p$ and $c_2 > c_2 - c_r > d_2 > d_2 - c_p$.

[9] Compare Güth and Yaari [15].

If all players maximize only their own material payoff (and are known to do so) all three games are solved easily by backward induction. Player 2 never incurs any costs in the last stage. This is anticipated by player 1. Therefore, player 1 defects in the first stage. This outcome tends to result also in an evolutionary setting, if individual preferences are unobservable and if all players are "playing the field" (i.e, no group structure is imposed)[10]. The reason is simple: Someone who chooses a strategy which maximizes his material payoffs earns more than someone who doesn't. However, results change when the total population is divided up into separated groups and when players interact only within their own group.

Whether the fraction of players of a certain preferences type grows or shrinks depends on their individual material payoffs. With the law of large numbers in mind we concentrate our analysis on deterministic approximations to the evolutionary dynamics[11]. The results of this paper hold for any payoff-monotonic dynamics. By payoff monotonicity we mean that the fraction of a preference-type with higher (equal) average material payoff grows faster (equally fastly) than the fraction of a preference-type with lower (equal) average material payoff. Furthermore, it is convenient to assume a continuous dynamics.

Assumption 1 *The evolutionary dynamics can be described by regular payoff monotonic growth rates*[12].

Furthermore, we say that a population state forms a stable equilibrium if it is an Asymptotically Stable State - a standard concept in evolutionary game theory[13].

What preference types are relevant for our analysis? In general each player assigns a subjective von Neumann-Morgenstern-utility to each outcome. These subjective utilities depend on the actual position of the player and may differ completely from his material payoffs[14]. We are mainly interested in the evolution of preferences for the position of player 2 - reciprocal behavior is only possible in that position. However, the evolutionary success of preferences for position 2 depends on the behavior of players 1. The following proposition helps to analyze the behavior of players 1:

[10]Nöldeke and Samuelson [23] give an example to illustrate that in general the subgame-perfect Nash equilibrium is not the only evolutionary stable equilibrium. Similarly, the results by Sethi and Somanthan [28] rely on the fact that non-credible threats can survive in certain evolutionary settings. However, Hart [16] and Kuzmics [21] show that the subgame-perfect Nash equilibrium results if certain limits are taken in a suitable way. See also Ok and Vega-Redondo [24] for a justification of the evolution of self-interested preferences when preferences are unobservable.

[11]This is very common in evolutionary game theory even if not entirely innocuous. For some caveats with this approach see Boylan [7]. For a thorough discussion of a deterministic dynamics as limit of a stochastic dynamics see Benaim and Weibull [1].

[12]The formal definition of a regular payoff monotonic growth rate can be found in Weibull [32] or in the appendix of our working paper version.

[13]See Weibull [32] or in the appendix of our working paper version for the precise definition.

[14]In the most general case 3 there exist 4 possible outcomes. A preference type is therefore characterized by a tuple of 8 subjective utilities (modulo a linear transformation). The first 4 subjective utilities describe a players preferences if he happens to play in position of player 1, the remaining 4 subjective utilities describe his preferences in position of player 2.

Proposition 1 *In any stable distribution of preference-types the fraction of players 1 who behave inconsistent with an expected material payoff maximizer must be zero.*

The proof is given in the appendix[15]. This result justifies to simplify the analysis by

Assumption 2 *In position 1 all players maximize their expected material payoffs.*

But when players happen to play in position 2 four classes of preferences may be relevant: We call someone a *"rewarder"* if he is willing to incur costs to reward a friendly action, a *"punisher"* if he is willing to incur costs to punish a hostile action[16], a *"reciprocator"* if he is willing to do both and *"self-interested"* if he is neither willing to incur costs to reward friendly nor to punish hostile behavior. We say someone has *"social preferences"* if he is either a rewarder, a punisher or a reciprocator. In case 1 and 2 only two of these types matter, respectively.

The large (or infinite) population is divided up randomly into separated groups of $(2N)$ players. N players are drawn randomly to play in position 2, the remaining N players play in position 1[17]. By "randomly" we mean that a players type does not influence the probabilities of the types of his group-members, i.e.:

Assumption 3 *The probability that k_+ of the N players 2 in a group are rewarders, k_- are punishers, k_{rc} are reciprocal and k_s are self-interested (with $k_+ + k_- + k_{rc} + k_s = N$) is multinomial distributed*[18]*:*

$$M_{N,\gamma_-,\gamma_+,\gamma_{rc},\gamma_s}(k_+, k_-, k_{rc}, k_s) = \frac{N!}{k_+!k_-!k_{rc}!k_s!}\gamma_+^{k_+}\gamma_-^{k_-}\gamma_{rc}^{k_{rc}}\gamma_s^{k_s} \tag{1}$$

where γ_i is the fraction of the i-th type in the total population (hence $\gamma_+ + \gamma_- + \gamma_{rc} + \gamma_s = 1$).

In case 1 and case 2 only two types of preferences are relevant. Then, the multinomial distribution reduces to the binomial distribution:

$$B_{N,\gamma}(k) = \frac{N!}{k!(N-k)!}\gamma^k(1-\gamma)^{N-k}. \tag{2}$$

[15]The intuitive reason for why Prop.1 holds for position 1 but not for position 2 is easier to understand once the entire model is presented. The crucial point is that player 2 observes the action of player 1 and therefore does not condition his action on the distribution of types within his group. However, player 1 needs to estimate the probabilities of meeting a certain type by the type distribution within his group and is therefore influenced by this distribution.

[16]The literature calls preferences for rewarding also "positively-reciprocal" and preferences for punishing "negatively-reciprocal".

[17]We might also reshuffle the positions of all players for each interaction. The main results would not change.

[18]An even more natural choice would be the multi-hyper-geometrical distribution (drawing without replacement). For simplicity we approximate it by the multinomial distribution. The qualitative results are not affected and the approximation is good for a large total population.

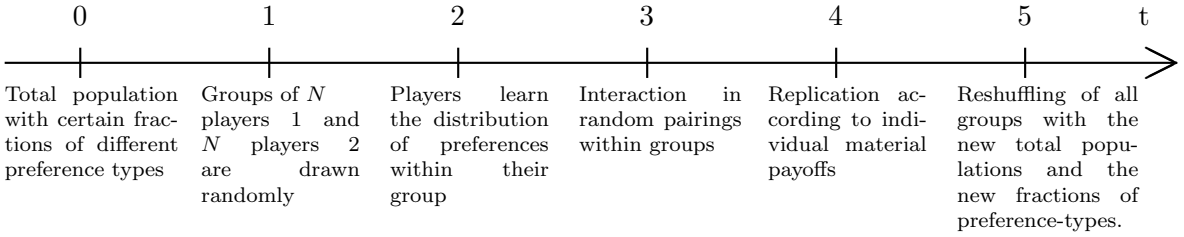
Individuals interact in random pairings within their group. Individuals do not know the type of their respective counterpart, but we assume them to know the frequency of each preference-type in their group:

Assumption 4 *Individuals know the fractions of the different types within their own group (but they don't know the type of their randomly matched opponent).*

This assumption is not entirely innocuous, but we relegate the detailed discussion to section 3. In order to abstract from repeated games effects, we assume individuals to play anonymously and finitely often. Hence, player 2 need not fear any consequences in a later stage whatever action he takes[19].

After a finite number of interactions preferences are replicated according to received material payoffs and all groups are completely reshuffled. A new cycle starts with the new fractions γ_i of the different preference types in the total population. Timing of events in our model is illustrated graphically in Figure 2.

Figure 2: Timing of events



Case 1: Costly Rewarding

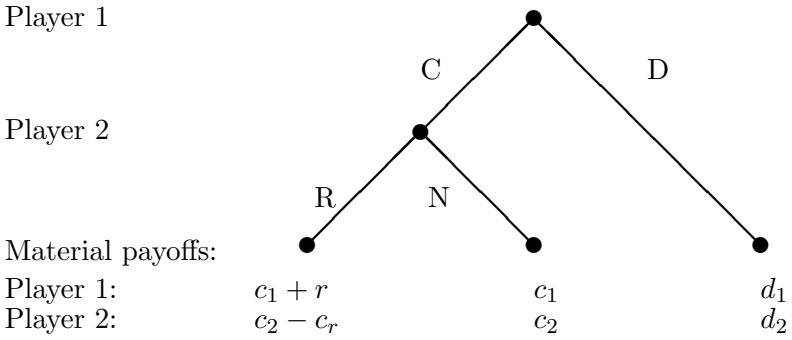
Case 1 concentrates on the possibility for player 2 to reward friendly behavior of player 1. First, player 1 decides whether to cooperate or defect. Player 2 observes this action. In case player 1 cooperates, player 2 may either incur the costs to reward player 1 or refuse to do so[20]. This game is also known as "trust game" and is illustrated in figure 3.

Player 1 maximizes his expected material payoff. Player 2 has either preferences for rewarding and rewards cooperation or he has self-interested preferences and does not reward

[19]This argument would become more involved if players guessed the fraction of e.g. rewarding players in their group from observed behavior instead of simply knowing it as assumed in Assumption 4. Then a self-interested player might tradeoff between costs of rewarding cooperation and gains from other players believing to live in a group with a higher fraction of a certain (e.g. reciprocal) type. See also the discussion in section 3.

[20]We could give player 2 an additional option to reward player 1 after defection. But it is straightforward to show that preferences for rewarding defection cannot be part of any stable equilibrium. Therefore, we ignore this possibility.

Figure 3: Interaction in case 1



with $c_1 + r > d_1 > c_1$ and $c_2 > c_2 - c_r > d_2$.

cooperation of player 1. The evolutionary process determines the fractions of each type in equilibrium.

We consider a group where k of the N players 2 have preferences for rewarding cooperation. Player 1 will base his decision whether to cooperate or whether to defect on his expected material payoff. Player 1 does not know the type of his opponent, but he knows the fraction $\frac{k}{N}$ of players 2 in his group, who would reward cooperation. Hence, player 1 expects an average material payoff of $(c_1 + \frac{k}{N}r)$ for cooperation. If player 1 defects he receives surely a payoff of d_1. Therefore, player 1 will cooperate if $c_1 + \frac{k}{N}r > d_1$ or equivalently if[21] $k > N\frac{d_1 - c_1}{r}$. Hence, cooperation occurs in a group only if the number of rewarding players 2 is above this threshold. We denote this threshold by k^*.

Definition 1 *k^* is the highest number of rewarding players 2 in a group which is still not sufficient to induce player 1 to cooperate. In other words*
$k \leq k^* \Rightarrow$ player 1 defects
$k > k^* \Rightarrow$ player 1 cooperates.

Calculation of k^* is straightforward:

$$k^* = \left\lfloor N\frac{d_1 - c_1}{r} \right\rfloor, \tag{3}$$

where $\lfloor x \rfloor$ denotes the largest natural number smaller or equal to the real number x. k^* is an integer with $0 \leq k^* \leq N - 1$.

In groups with k^* or less rewarding players 2 no cooperation occurs. Players 1 defect and players 2 receive a material payoff of d_2 - independently of their types. In groups with more

[21] For notational simplicity we define the tie breaking rule that player 1 defects if his expected payoff for defecting equals that for cooperating.

than k^* rewarding players 2 players 1 cooperate. A rewarding player 2 receives a material payoff of $(c_2 - c_r)$. A self-interested player 2 exploits cooperation of player 1 and receives a material payoff of c_2. These payoffs are summarized in table 1.

Table 1: Material payoffs of player 2

	Payoffs in groups with $k \leq k^*$	Payoffs in groups with $k > k^*$
Rewarder	d_2	$c_2 - c_r$
Self-interested	d_2	c_2

For a player 2 the probability that exactly k of the other $(N-1)$ players 2 in his group have preferences for rewarding is $B_{N-1,\gamma}(k) = \frac{(N-1)!}{k!(N-1-k)!}\gamma^k(1-\gamma)^{N-1-k}$, where γ is the fraction of rewarders in the total population. If this player has preferences for rewarding the total number of rewarders in his group is $(k+1)$, otherwise it remains k. Hence, a self-interested player 2 receives an expected material payoff of[22]

$$\bar{u}_s(\gamma) = d_2 \sum_{k=0}^{k^*} B_{N-1,\gamma}(k) + c_2 \sum_{k=k^*+1}^{N-1} B_{N-1,\gamma}(k) \tag{4}$$

and a rewarding player 2 receives an expected material payoff of

$$\bar{u}_+(\gamma) = d_2 \sum_{k=0}^{k^*-1} B_{N-1,\gamma}(k) + (c_2 - c_r) \sum_{k=k^*}^{N-1} B_{N-1,\gamma}(k). \tag{5}$$

Due to the assumption 1 of payoff monotonicity the fraction of rewarding players grows (falls) if they receive a higher (lower) average payoff than the self-interested type. Hence, we can see from the sign of the difference

$$\begin{aligned} \bar{u}_+(\gamma) - \bar{u}_s(\gamma) &= (c_2 - c_r - d_2) B_{N-1,\gamma}(k^*) - c_r \sum_{k=k^*+1}^{N-1} B_{N-1,\gamma}(k) \qquad (6) \\ &= (c_2 - c_r - d_2)\binom{N-1}{k^*}\gamma^{k^*}(1-\gamma)^{N-1-k^*} - c_r \sum_{k=k^*+1}^{N-1}\binom{N-1}{k}\gamma^k(1-\gamma)^{N-1-k} \end{aligned}$$

when the fraction of rewarding players increases, decreases or remains stable. First, we consider the case $c_2 - d_2 - c_r \leq 0$, i.e. gains of cooperation for player 2 are smaller than costs of rewarding. Then, all terms on the right hand side of equation 6 are negative (or zero) and a self-interested player 2 earns always more than a rewarding player 2.

[22] A different way to calculate this is to multiply the payoff of a rewarding (self-interested) player 2 in a group of k rewarding players 2, multiply it by k $(N-k)$ and weight it by the probability that a group has k rewarding players 2 (i.e. the binomial coefficient). If we sum this up over all $0 \leq k \leq N$ and divide it by the total number of players 2 of that type we get the average payoff. Of course, the results remain unchanged.

Proposition 2 *If $c_2 - d_2 - c_r \leq 0$, i.e. the cost for rewarding exceed player 2's gains from player 1's cooperation, then only an entirely self-interested population is stable*[23].

But mainly we are interested in the case $c_2 - d_2 - c_r > 0$, i.e. gains from cooperation for player 2 exceed his costs of rewarding. Then, there is a chance for the survival of preferences for rewarding . In fact, for $k^* < N-1$ preferences for rewarding and self-interested preferences coexist in any stable equilibrium:

Proposition 3 (Coexistence) *Let $c_2 - d_2 - c_r > 0$ and $k^* < N-1$. Then, the monomorphic population states (i.e. states with a fraction $\gamma = 0$ or $\gamma = 1$ of rewarders) are unstable for any payoff monotonic dynamics. Preferences for rewarding can invade a self-interested population and self-interested preferences can invade a population of rewarders.*

Remark 1 *If $c_2 - d_2 - c_r > 0$ and $k^* = N - 1$ then only a monomorphic population of preferences for rewarding forms a stable equilibrium.*

All proofs are relegated to the appendix.

For an intuitive understanding of Proposition 3 first consider a population consisting almost entirely of rewarders. Then, almost all groups consist almost entirely of rewarding players 2. Therefore, players 1 cooperate in almost all groups. A rewarding player 2 receives a payoff of $(c_2 - c_r)$ only, whereas a self-interested player 2 saves the costs of rewarding and earns the higher payoff of c_2. Therefore, the fraction of self-interested players grows.

The intuition why preferences for rewarding can invade a self-interested population is slightly more involved. Consider a population consisting almost entirely of self-interested players. Then, the vast majority of groups contain too few rewarding players 2 to induce cooperation of players 1. In these groups self-interested and rewarding players 2 receive the same payoff d_2. But in a small number of groups the fraction of rewarding players 2 is above the threshold k^* and players 1 are willing to make the advanced concession of cooperation. Every player in these groups receives a higher payoff than most players in groups without cooperation. But the fraction of rewearding players 2 in this groups is at least $\frac{k^*}{N}$ and therefore far above the fraction of rewarders in the total population (which is close to zero). Therefore, rewarding players profit relatively more from this successful groups and can invade a self-interested population.

If $k^* = N-1$ the result changes for the following reason: Then, self-interested preferences cannot invade a population of rewarders. Even if a self-interested player 2 is the only invader in his group he destroys cooperative behavior of players 1. Hence, if $k^* = N - 1$ rewarding players 2 do always at least as good as self-interested players 2.

[23] A population is called stable if it is an asymptotical stable state of the dynamics. For details see e.g. Weibull [32] or the appendix of the working paper version.

According to proposition 3 only mixed populations are candidates for stable preference distributions. In fact, there exists a unique stable equilibrium.

Theorem 1 (Unique mixed equilibrium) *Let* $c_2 - d_2 - c_r > 0$ *and* $k^* < N-1$. *Then there exists a unique stable equilibrium. Self-interested preferences and preferences for rewarding coexist in this equilibrium.*

Figure 4 illustrates the dynamics of the evolutionary process for an example. The difference in average material payoffs between rewarders and self-centered individuals is plotted as a function of γ for $N = 20, d_1 = 1, c_1 = 0, r = 2, d_2 = 5, c_2 = 0, c_r = 1$. The fraction of rewarders in the stable equilibrium of this example is $\gamma^{eq} \approx 0.5876$. If the fraction γ of rewarding individuals is below γ^{eq} then they earn a higher average material payoff and their fraction γ increases. If $\gamma > \gamma^{eq}$ rewarding players earn less and γ decreases. Due to the assumed continuity of the evolutionary dynamics γ converges to γ^{eq}.

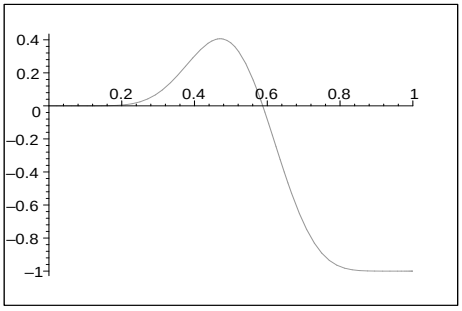


Figure 4: $(\bar{u}_+ - \bar{u}_s)$ as function of γ for $N = 20, d_1 = 1, c_1 = 0, r = 2, d_2 = 5, c_2 = 0, c_r = 1$.

Efficiency: Player 1 cooperates only if his expected material payoff under cooperation is higher than under defection. On the other hand preferences for rewarding can only survive if player 2 receives a higher material payoff after cooperation and rewarding than after defection. Hence, the existence of preferences for rewarding can only lead to a Pareto-improvement (in material payoffs) in comparison to a purely self-interested population. But for $k^* < N-1$ non-rewarding self-interested players survive, too. Hence, inefficient defection occurs in some groups and the outcome is still inefficient.

Comparative Statics for Case 1

The fraction γ^{eq} of rewarders in the unique stable equilibrium is characterized by the equation $\bar{u}_+(\gamma^{eq}) - \bar{u}_s(\gamma^{eq}) = 0$. Inserting equation 6 and rearranging leads to

$$c_2 - c_r - d_2 = c_r \sum_{k=1}^{N-1-k^*} \frac{(N-1-k^*)!k^*!}{(N-1-k^*-k)!(k^*+k)!} \left(\frac{\gamma^{eq}}{1-\gamma^{eq}}\right)^k, \tag{7}$$

for $0 < \gamma^{eq} < 1$. The comparative statics is easily derived from this condition.

First, we consider the dependence of the equilibrium fraction γ^{eq} of rewarding players on the group-size N. N enters into equation 7 not only directly but also via $k^* = \left[N\frac{c_2-d_2}{r}\right]$. Due to the truncation k^* is only almost proportional to N. In general, a higher group size N tends to decrease γ^{eq}. But for some values the truncation can invert this effect slightly. To avoid such problems we concentrate in the following proposition on sequences of N for which $\frac{k^*}{N} \equiv c$ is kept constant.

Proposition 4 *An increase in the group size N, keeping $\frac{k^*}{N}$ constant, lowers the fraction γ^{eq} of preferences for rewarding in equilibrium.*

Intuitively, larger groups reduce the probability of being pivotal. Therefore, the advantage of being a rewarder is reduced. Hence, the fraction of rewarding players decreases in equilibrium[24]. This result is consistent with the common feeling that in large anonymous groups the level of cooperation is lower. The influence of a single player on the reputation of a large group is small. In larger groups a lower number of rewarding players survive in equilibrium.

Now we consider the dependence of γ^{eq} on the parameters of the game. We start with the influence of the costs c_r player 2 has to incur if he rewards cooperation.

Proposition 5 *Higher costs c_r of rewarding lead to a lower fraction γ^{eq} of the preferences for rewarding in equilibrium. Furthermore, $\lim_{c_r \to 0} \gamma^{eq} = 1$ and $\lim_{c_r \to (c_2-d_2)} \gamma^{eq} = 0$.*

Intuitively, higher costs of rewarding do not influence the incentives of player 1, but reduce the fitness of rewarding players 2. Therefore, their fraction is reduced in equilibrium.

Proposition 6 *Higher gains of cooperation (c_2-d_2) lead to a higher fraction γ^{eq} of rewarding players 2 in equilibrium. Furthermore, $\lim_{(c_2-d_2) \to c_r} \gamma^{eq} = 0$ and $\lim_{(c_2-d_2) \to \infty} \gamma^{eq} = 1$.*

The intuition is straightforward, again. If gains of cooperation increase then gains from being pivotal increase for a rewarding player 2. The costs are not effected. Therefore, the fraction of rewarding players 2 increases.

Lemma 1 *If the threshold k^* of rewarding players 2 in a group (above which players 1 in that group start to cooperate) increases, then the fraction of rewarding players 2 in equilibrium increases.*

The intuition for this lemma is slightly involved. An increase in k^* means that there have to be more rewarding players 2 in a group in order to induce cooperation of player 1. Hence, a

[24]The last argument is not entirely complete. The probability of having to bear the costs for rewarding may also decrease and therefore a counterbalancing effect may arise. We can show, that the equilibrium fraction of rewarders decreases with the group size, but so far we have not be able to show whether the equilibrium fraction does or does not converge to zero if the group size goes to infinity. Numerical results suggests that γ^{eq} decreases only slowly and may not converge to zero.

lower number of self-interested players 2 can free-ride without putting cooperation in danger. Therefore, the total number of self-interested players 2 decreases.

From lemma 1 we can easily derive two further results. The costs of cooperation for player 1 ($d_1 - c_1$) and the amount of the possible reward r enter in equation 7 only through k^*. Hence,we obtain

Corollary 1 *The equilibrium fraction γ^{eq} of rewarding players increases (weakly) if the costs $(d_1 - c_1)$ of cooperation for player 1 increase.*

Corollary 2 *The equilibrium fraction γ^{eq} of rewarding players decreases (weakly) if the amount r by which a player 1 can be rewarded cooperation increases.*

Both corollaries might seem counterintuitive at first glance. But the intuition is similar to that of lemma 1. Increasing costs of cooperation or a decreasing reward make it more difficult to induce player 1 to cooperate. Therefore, free-riding by a self-interested player 2 becomes more likely to destroy cooperation. Hence, the fraction of self-interested players has to decrease in equilibrium.

Case 2: Costly Punishment

In case 1, player 2 had only the possibility to reciprocate positively, i.e. to reward a friendly action. In case 2 we analyze the evolution of preferences if there exists the possibility for player 2 to punish hostile behavior (i.e. defection) of player 1. This punishment is costly[25]. The interaction is illustrated in Figure 5.

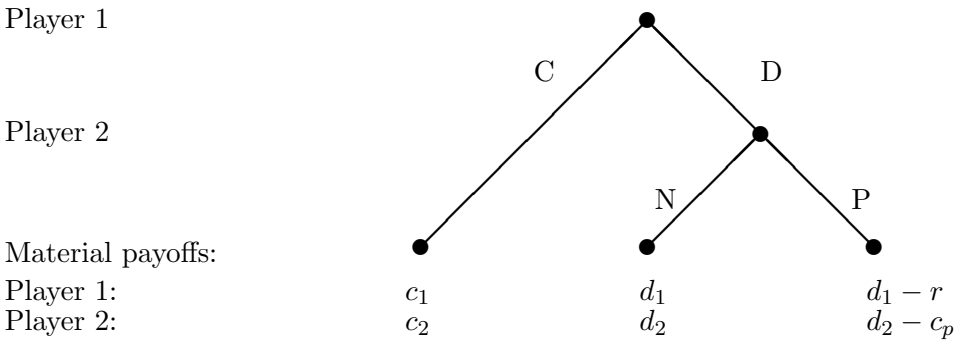
Player 2 has either preferences for punishing or self-interested preferences. A punishing player 2 is willing to incur the costs to punish player 1 in case of defection. But if player 2 is self-interested he avoids these costs and does not punish defection of player 1. Player 1 maximizes his expected material payoff. In a group where k of the N players 2 have preferences for punishing player 1 expects an material payoff of $(d_1 - \frac{k}{N}p)$ after defection. After cooperation he receives a material payoff of c_1. Therefore, player 1 cooperates if and only if[26] $c_1 > d_1 - \frac{k}{N}p$ or equivalently if $k > N\frac{d_1 - c_1}{p}$. Analogously to case 1 we denote the threshold by k^{**}.

Definition 2 *k^{**} is the highest number of punishing players 2 in a group which is still insufficient to induce a self-interested player 1 to cooperate. In other words*
$k \leq k^{**} \Rightarrow$ player 1 defects
$k > k^{**} \Rightarrow$ player 1 cooperates.

[25] We might allow for this punishment after cooperation as well as after defection of the first player. But - similar to case 1 - preferences which lead to punishment after cooperation (e.g. spiteful preferences) vanish in our model due to natural selection. Again, we simplify the analysis by looking at the possibility of punishment only if player 1 defected.

[26] Again, we assume the tie breaking rule that player 1 defects if he is indifferent.

Figure 5: Interaction in case 2



with $c_1 + p > d_1 > c_1; c_2 > d_2; c_p > 0$.

The calculation of k^{**} is straightforward:

$$k^{**} = \left[N \frac{d_1 - c_1}{p} \right]. \tag{8}$$

k^{**} is an integer with $0 \leq k^{**} \leq N - 1$.

In groups with k^{**} or less punishing players 2 no cooperation occurs. Players 1 defect. In response punishing players 2 receive material payoffs of $(d_2 - c_p)$. Self-interested players 2 avoid costs of punishing and receive higher material payoffs of d_2. In groups with more than k^{**} punishing players 2 players 1 cooperate. Therefore, players 2 receive - independently of their types - material payoffs of c_2. The payoff structure is summarized in table 2.

Table 2: Material payoffs of player 2

	Payoffs in groups with $k \leq k^{**}$	Payoffs in groups with $k > k^{**}$
punisher	$d_2 - c_p$	c_2
Self-interested	d_2	c_2

Now let γ be the fraction of punishers in the total population. Analogously to case 1 self-interested players 2 receive an expected material payoff of

$$\bar{u}_s(\gamma) = d_2 \sum_{k=0}^{k^{**}} B_{N-1,\gamma}(k) + c_2 \sum_{k=k^{**}+1}^{N-1} B_{N-1,\gamma}(k) \tag{9}$$

and punishing players 2 the expected material payoff of

$$\bar{u}_-(\gamma) = (d_2 - c_p) \sum_{k=0}^{k^{**}-1} B_{N-1,\gamma}(k) + c_2 \sum_{k=k^{**}}^{N-1} B_{N-1,\gamma}(k). \tag{10}$$

Due to the assumption of payoff monotonicity the fraction of punishers grows (falls) if punishing players 2 receive a higher (lower) average payoff than self-interested players 2. Hence, we are interested in the sign of the difference

$$\bar{u}_-(\gamma) - \bar{u}_s(\gamma) = (c_2 - d_2) B_{N-1,\gamma}(k^{**}) - c_p \sum_{k=0}^{k^{**}-1} B_{N-1,\gamma}(k) \tag{11}$$

$$= \quad (c_2 - d_2) \frac{(N-1)!}{(k^{**})!(N-1-k^{**})!} \gamma^{k^{**}} (1-\gamma)^{N-1-k^{**}} - c_p \sum_{k=0}^{k^{**}-1} \frac{(N-1)!}{k!(N-1-k)!} \gamma^k (1-\gamma)^{N-1-k}.$$

For $0 < \gamma < 1$ follows

$$\bar{u}_-(\gamma) - \bar{u}_s(\gamma) \quad \gtreqless 0 \tag{12}$$

$$\Leftrightarrow \qquad c_2 - d_2 \qquad \gtreqless c_p \sum_{k=0}^{k^{**}-1} \frac{(k^{**})!(N-1-k^{**})!}{k!(N-1-k)!} \left(\frac{1-\gamma}{\gamma} \right)^{k^{**}-k}. \tag{13}$$

The right hand side of equation 13 is strictly decreasing and continuous in γ, tends to 0 if γ tends to 1 and to infinity if γ tends to zero. The left hand side of equation 13 has a fixed positive value. Hence, there exists only one equilibrium of mixed types which is unstable. We denote the fraction of punishers in this unstable equilibrium by γ^{cut}. The only stable equilibria are the corner solutions.

Theorem 2 *Let $k^{**} > 0$. Then, the two monomorphic equilibria - in which either all players have preferences for punishing or all players have self-interested preferences - are stable. The unique mixed equilibrium is not stable.*

In contrast to case 1 the option to punish defection drives the population to a monomorphic state. Either a "culture of punishment" develops, where all players are willing to punish, or a "culture of laissez faire", where nobody bothers to punish defectors. The evolutionary dynamics is illustrated by Figure 6.

Theorem 2 is very intuitive. If virtually no player 2 is willing to punish defection a single punisher is very unfit. In almost any group he is the only punisher and unable to enforce cooperation of player 1. Player 1 defects and the punishing player 2 has to pay the costs c_p of punishing. Therefore, he is less fit than a self-interested player 2 who does not punish. On the other hand if virtually all players 2 are willing to punish they seldom have to prove this. Players 1 in almost all groups cooperate in order to avoid punishment. Only in few groups in which the number of punishing players 2 is below the threshold k^{**}, self-interested and

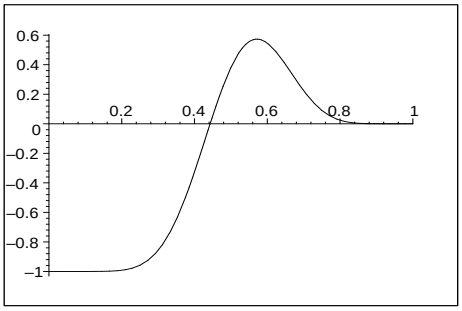


Figure 6: $(\bar{u}_{-} - \bar{u}_s)$ as function of γ for $N = 20, d_1 = 1, c_1 = 0, p = 2, d_2 = 5, c_2 = 0, c_p = 1$.

punishing players 2 receive different payoffs. But most of these groups are just one punisher below the threshold. In these groups a punishing player 2 is pivotal in inducing cooperation. Therefore, he benefits from his preferences.

In the equilibrium of a population of punishers players 1 always cooperate and no player 2 has to prove his willingness to punish. How would the results change if players 1 make mistakes and fail to cooperate sometimes? Appendix D demonstrates that results change only slightly if probabilities of mistakes are sufficiently small. There remain two stable equilibria. The equilibrium consisting only of self-interested preferences remains stable. However, a population consisting only of punishers is no longer stable. A small fraction of self-interested players can invade. But the fraction of self-interested invaders remains arbitrary small if probabilities of mistakes are sufficiently small[27]. Hence, there might still develop a culture of punishment with a high fraction of punishers and a small fraction of self-interested players.

Results of case 2 differ from case 1 also in terms of efficiency (in material payoffs). In order to be able to rank the outcomes we take the point of view of a player who does not know yet whether he plays in player-position 1 or 2 and might play in each position with equal probability. Then, cooperation is efficient if $d_1 - c_1 < c_2 - d_2$, i.e. if player 2 profits more from cooperation than player 1 loses. However, defection (and no punishment) is efficient if $d_1 - c_1 > c_2 - d_2$. The option to punish defection can enforce complete cooperation (in a world without mistakes and in the right equilibrium). If $d_1 - c_1 < c_2 - d_2$ this is efficient. But cooperation can be enforced by the threat of punishment also in cases where cooperation is inefficient. Hence, the possibility to punish defection can be efficiency enhancing as well as efficiency reducing.

The unstable mixed equilibrium separates the basins of attraction of both stable equilibria. If the initial fraction of punishers is below γ^{cut} then only self-interested players survive, otherwise only punishers. The lower the value of γ^{cut} the more initial population states evolve to a population of punishers. We relegate the comparative statics of γ^{cut} to appendix C.

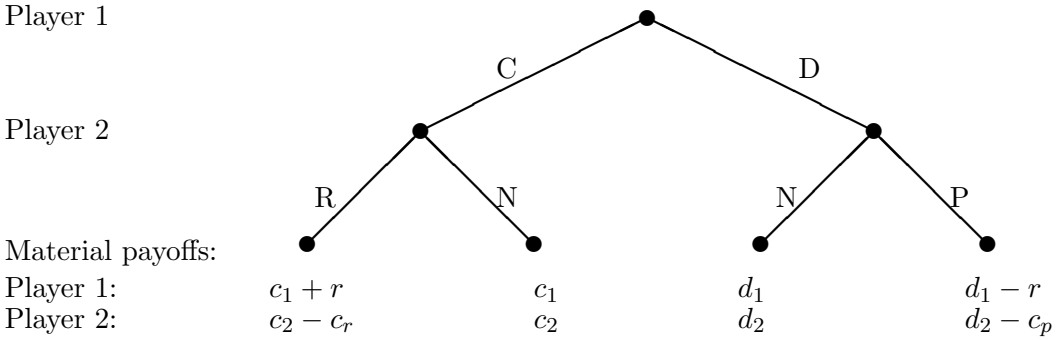
[27]However, a moderate probability of mistakes may result in a significant shift of the punisher equilibrium. See Appendix D for details.

In case 2 there are two equilibria and we don't know wether a "culture of punishment" or a "culture of laissez faire" develops. However, case 3 suggests that the survival of preferences for punishing becomes more likely if player 2 has both options - punishing and rewarding. In fact, under suitable conditions only an entirely population forms an evolutionary stable equilibrium in case 3.

Case 3: Costly Rewarding or Costly Punishment

In case 3 player 2 has both options - costly punishing after defection and costly rewarding after cooperation. This allows us to analyze the co-evolution of preferences for rewarding and preferences for punishing, i.e. how the evolution of one side of reciprocity influences the evolution of the other side. The interaction is illustrated in Figure 7.

Figure 7: Interaction in case 3



with $c_1 + r > d_1 > c_1$ and $c_2 > c_2 - c_r > d_2 > d_2 - c_p$.

All players 1 maximize their expected material payoffs. There are four different types of players 2 [28]: **Self-interested** players neither reward cooperation nor punish defection. **Punishers** do not reward cooperation, but punish defection. **Rewarders** reward cooperation, but do not punish defection. **Reciprocal** players reward cooperation as well as they punish defection. In order to reduce technical problems, we make the following[29]

Assumption 5 *The material loss p for player 1 after being punished equals his material gain r after being rewarded, i.e.* $p = r$.

Due to Assumption 5 punishers and rewarders have exactly the same influence on the behavior of players 1 in their group. Hence, material payoffs of all other players 2 are not effected if

[28] Again, we neglect generic cases of preferences which associate the same subjective utility to different outcomes.

[29] The general intuition for the results of this section holds without this assumption, but assumption 5 simplifies the analysis considerably.

we replace a punisher by a rewarder or vice versa. We know from the analysis of case 2 that preferences for punishing are evolutionary more successful if their own fraction grows. Hence, punishers profit also from a growing fraction of rewarders. Any kind of reciprocity helps to induce cooperation of players 1 and reduces costs of being a punisher.

Remark 2 *Higher fractions of rewarders and higher fractions punishers enhance the evolutionary success of preferences for punishing.*

Conversely, we know from case 1 that the evolutionary success of preferences for rewarding relative to self-interested preferences decreases if their own fraction becomes too large. Hence, the same must hold for too large a fraction of punishers. Furthermore, relative to preferences for punishing the success of preferences for rewarding is reduced by an increase of the fraction of any type of reciprocity. The higher the fraction of rewarders or punishers, the more groups are above the threshold for cooperation. Therefore, costs of rewarding grow whereas the costs of punishing fall.

Remark 3 *Higher fractions of rewarders and higher fractions of punishers reduce the evolutionary success of preferences for rewarding relative to the success of preferences for punishing.*

This interdependence between the evolution of both types of reciprocity has interesting consequences. Consider an entirely self-interested population. Preferences for punishing cannot invade such a population directly, as shown in Case 2. But preferences for rewarding can invade (see Case 1). If enough rewarders invade, they may serve as a "catalyst" and enable the invasion of punishers. The more punishers invade, the more successful they become and finally they drive out self-interested players as well as rewarders.

Remark 4 *Preferences for rewarding may serve as a catalyst for the evolution of preferences for punishing. Rewarders can invade an entirely self-interested population. Their existence enables punishers to invade, too. Finally, preferences for punishing become more and more successful and drive out self-interested preferences as well as preferences for rewarding.*

Now we look for stable equilibria in case 3. First, we check for stable monomorphic populations, i.e stable populations of only one preference-type.

Proposition 7 *The only monomorphic stable equilibrium consists entirely of punishers.*

Are there other stable equilibria consisting of several preference types? The answer depends on the parameters of the model. For certain parameters this is the only stable equilibrium. For others there exist further stable equilibria. It is easier to capture the basic intuition if reciprocal preferences are neglected. Hence, for the moment we restrict ourselves to the possibilities of self-interested preferences, preferences for rewarding and preferences for punishing. Consider a population consisting only of rewarders and self-interested and players.

According to Case 1 this population evolves towards a unique equilibrium containing both preference types. Can a small fraction of punishers invade in this equilibrium? The answer depends on the fraction γ^{eq} of rewarders in equilibrium. Since we assumed $p = r$ the effect of a rewarding player 2 on any other player 2 in his group is precisely the same as the effect of a punishing player 2 at the same place. Hence, preferences for punishing can invade this equilibrium if and only if the fraction γ^{eq} of rewarders in this equilibrium (determined by Equation 7) is higher than the threshold γ^{cut} (determined by Equation 13) above which preferences for punishing become more successful than self-interested ones. Preferences for punishing become relatively more successful, the higher their own fraction of the population. Therefore, once preferences for punishing can invade, they drive out all other preferences and the dynamics leads to the monomorphic equilibrium of preferences for punishing.

Proposition 8 *Let γ^{eq} be defined by equation 7 and γ^{cut} by equation 13.*
If reciprocal preferences are neglected, i.e. only the subspace of self-interested preferences, preferences for rewarding and preferences for punishing is considered, then

a) *if $\gamma^{eq} > \gamma^{cut}$, then the only stable equilibrium is a monomorphic population, where all players have preferences for punishing. The population converges to this equilibrium from any interior state.*

b) *If $\gamma^{eq} < \gamma^{cut}$, then there are two stable equilibria. One stable equilibrium is the monomorphic population of preferences for punishing. In the other stable equilibrium preferences for rewarding and self-interested preferences coexist*[30]*. In this equilibrium the fraction of preferences for rewarding is γ^{eq}.*

Figure 8 illustrates the dynamics of the evolutionary process in case 3 with the parameters of our previous examples. Here we have $\gamma^{eq} > \gamma^{cut}$ and the equilibrium with a fraction γ_{eq} of rewarders and a fraction γ^{cut} of self-interested players is not stable: Punisher earn a higher average payoff, invade successfully and drive out all other preferences.

Including reciprocal preferences does not change the basic intuition. Preferences for punishing form still a stable equilibrium and a mixture of a fraction of γ^{eq} with preferences for rewarding and $1-\gamma^{eq}$ self-interested preferences remains a stable equilibrium under the slightly

[30]Notice that even in the case of Prop.8b) where we have still two stable equilibria it is in Case 3 more likely to end up in the monomorphic equilibrium (compared to Case 2). This is meant in the spirit of the model by Kandori et al. [20]: Imagine that each member of the entire (large but finite) population mutates with small probability to any other preference-type. Then, the minimum number of mutations necessary to move from the monomorphic equilibrium to the basin of attraction of the other equilibrium is exactly the same as in case 2. But the other way around less mutations are sufficient to move the population from the bi-morphic equilibrium to the basin of attraction of the monomorphic equilibrium. That is because the rewarders are advantageous for the invasion of the punishing type.

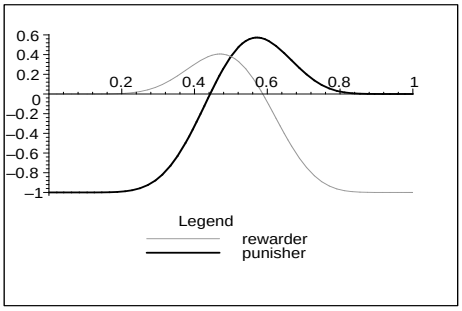


Figure 8: Case 3 with $\gamma^{eq} > \gamma^{cut}$: $(\bar{u}_+ - \bar{u}_s)$ and $(\bar{u}_- - \bar{u}_s)$ as functions of $\gamma \equiv \gamma_+ + \gamma_-$ for $N = 20, d_1 = 1, c_1 = 0, r = 2, d_2 = 5, c_2 = 0, c_r = 1$.

more restrictive condition $\gamma^{eq} < \min\{\gamma^{cut}; \gamma^h\}$, where γ^h is defined by the equation

$$c_2 - d_2 - c_r = c_p \sum_{k=0}^{k^{**}-2} \frac{(k^{**}-1)!}{k!} \frac{(N-k^{**})!}{(N-1-k)!} \left(\frac{1-\gamma^h}{\gamma^h}\right)^{k^{**}-1-k}. \tag{14}$$

The tightening of the condition is necessary to ensure that reciprocal preferences cannot invade the mixed equilibrium either. Furthermore, reciprocal preferences can be part of an equilibrium only under very special conditions. If most groups induce players 1 to cooperate, preferences for punishing tend to outperform reciprocal ones, since they do not have to bear the costs of rewarding. If on the other hand most groups are not able to induce cooperation, then preferences for rewarding tend to outperform reciprocal ones since they do not bear the costs of punishing in the frequent cases of defection. But for certain parameters there exist equilibria with a positive fraction of reciprocal preferences. These additional equilibria are not robust to small changes in parameters of the model and are not very plausible. Therefore, we relegate the discussion of these equilibria to the appendix E and focus attention on the discussed equilibria summarized in the following

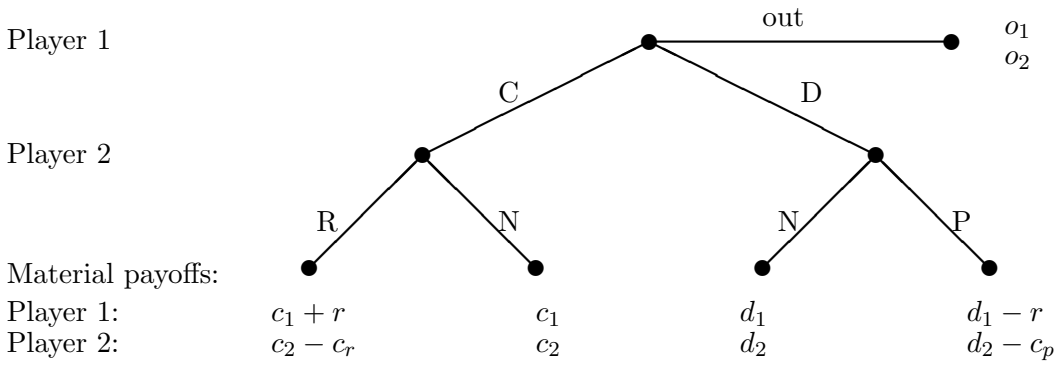
Proposition 9 *Let γ^{eq} be defined by equation 7,γ^{cut} by equation 13 and γ^h by equation 14. Then, for any payoff-monotonic selection dynamics holds*

a) *a monomorphic population of punishers forms a stable equilibrium.*

b) *If $\gamma^{eq} < \min\{\gamma^{cut}, \gamma^h\}$, then also a population with a fraction γ^{eq} of rewarders and a fraction $(1-\gamma^{eq})$ of self-interested players forms a stable equilibrium.*

So far, preferences for punishing will - once they invade - drive out preferences for rewarding completely. In equilibrium either preferences for rewarding or preferences for punishing survive, but not both. However, both sides of reciprocity can survive in one equilibrium, if player 1 cannot be forced to participate in the interaction, i.e. player 1 has an additional outside option as illustrated in figure 9.

Figure 9: Interaction in case 3 with outside option



with $c_1 + r > d_1 > o_1 > c_1$; $c_2 > c_2 - c_r > o_2 > d_2 > d_2 - c_p$.

Now, a population consisting only of punishers is not stable any more. The threat of punishment alone can only deter player 1 from defecting. But player 1 opts out as long as he does not expect to be rewarded for cooperation. Analogous to case 1 reciprocal players who reward and punish can invade the population of punishers. In some groups their willingness to reward induces players 1 to cooperate instead of opting out. This makes reciprocal preferences initially more successful until reciprocal preferences and preferences for punishing are in a mixed equilibrium[31]. Hence, there is an equilibrium in which all or most players are willing to punish defection, some of them do reward cooperation and some don't[32].

A different explanation for the survival of both types of reciprocity arises if individuals engage in different types of interactions - sometimes similar to case 1, sometimes similar to case 2 or 3. If players have general preferences and do not have different preferences for different types of interactions, then some rewarders, some punishers and reciprocators can survive[33].

[31]This equilibrium is only Lyapunov stable but not asymptotically stable. That is because payoffs do not change if some reciprocal players are replaced by rewarders. But the set of population states with a fraction of $(1 - \gamma^{eq})$ of punishers, $\gamma \in [0; \gamma^{eq}]$ of reciprocators and $(\gamma^{eq} - \gamma)$ of rewarders forms an asymptotically stable set of equilibria.

[32]For certain parameters there exist further equilibria, but the detailed analysis is beyond the scope of this paper.

[33]The question how far preferences may depend on the respective interaction is a subtle one. On the one hand preferences should not be expected to evolve independently for any type of interaction. On the other hand people may well classify interactions by broad categories and their preferences may well depend on whether they assign a certain interaction to one category or another. Empirical evidence as well as theoretical approaches in the direction of Samuelson [26] could offer interesting insights to this question.

3 Discussion

Our results hold for any payoff-monotonic evolutionary dynamics. Hence, our model can be interpreted in terms of biological- or cultural evolution as well as processes of learning by success and failure. The assumption 3 of regularly reshuffling may seem to strong for real world applications. But, precisely this assumption allows us to abstract from assortative group selection effects and to isolate the effects we are interested in. From a theoretical standpoint this strengthens our results: Preferences for punishing or rewarding can survive evolution even without effects of assortative matching.

The findings of our evolutionary analysis seem to fit observed human behavior reasonably well. Some people do reward friendly behavior, while some others don't. The willingness to punish hostile behavior seems more robust than the willingness to reward friendly behavior. However, in most experimental studies a significant number of individuals does not punish defection. This may seem to contrast the statement of Theorem 2, that either only punishers survive or no punisher at all. But, if moderate mistakes are included in our analysis, as we do in Appendix D, then a significant number of self-interested players may survive in the "punisher equilibrium".

Furthermore, in our analysis we assumed implicitly that preferences are formed independently for each type of interaction. This seems plausible if different types of interaction are clearly distinguishable and if each of these interactions played an important role during the process of evolution. However, in unfamiliar situations individuals might act as they are used to in situations which they consider similar. For instance, in experimental studies of the *Ultimatum Game* some players might compare the game to a common project of equal partners whereas other compare it to a bargaining with their superior. If in the "real world" a common culture of punishing has developed in the first case and a culture of laissez-faire in the second, then only some players punish unfair offers in the experiment[34].

Further applications arise if we consider the evolutionary dynamics as a learning process and preferences as rules of behavior[35]. All over the world some street traders try to cheat tourists and other sell them fair quality - in accordance with case 1. In contrast, different countries seem to have different norms for how to wait for a bus: In England everybody queues and trying to push your way to the front can provoke harsh reactions. In Italy however nobody queues and there is no way for an individual to enforce queueing. In accordance with case 2 Italy and England seem to have evolved in to different equilibria.

How robust are our results? Appendix D demonstrates that sufficiently small trembles of players change the results only slightly. But the assumption 4 that the distribution of

[34]This interpretation fits well with findings of Henrich et al. [18]. They conduct an experiment with the ultimatum game in 15 small-scale societies and find that peoples strategies depend strongly on culture and norms of their society.

[35]The time scale to reach an equilibrium is then considerably smaller.

preferences in a group is common knowledge remains to be discussed. The assumption is considerably weaker then assuming observability of individual preferences. For instance, consider the newspaper boxes in many towns. Often you can simply take a newspaper and there is no control for whether you pay for the newspaper or whether you don't. At the end of the day you do not know who payed for the newspaper, but by counting the money you have a good estimator which fraction payed. However, there is still a problem with assumption 4 which becomes apparent by considering a group with a number of e.g. rewarding players 2 just below the threshold k^* which would induce players 1 to cooperate. A self-interested player 2 in this group would prefer to be a rewarder, because this increases his material payoff. In case the fraction of rewarders in the group is not common knowledge - as assumed - but has to be learned from other people's behavior in previous periods, then even a self-interested player 2 might have an incentive to incur the costs of rewarding in order to sustain cooperation. But it has to be noticed, that he prefers the case when another self-interested player 2 of his group mimics rewarding behavior instead of him. Hence, without assumption 4 players 2 might try to build a reputation - not individually, but a reputation of their group. One way to avoid such group reputation building effects is to assume that players have a sufficiently low discount factor. Then player 2 is not willing to forgo current profits for future gains from group reputation and the true fractions of each types can easily be learned in a first round[36]. A different way to avoid group reputation building by players 2 is to relax the assumption of (costless) perfect rationality. Group reputation building is rather complex for player 2. Furthermore, in equilibrium all types receive the same average material payoff. Hence, in equilibrium small costs of thinking can make sure, that player 2 plays according to his type in each period.

4 Conclusions

This non-assortative group selection model offers an explanation for the evolutionary survival of both sides of reciprocal preferences. Despite individual behavior and preferences being unobservable, individuals continue to have a marginal effect on the "reputation" of his group, which influences the behavior of the other players in their group. This effect is sufficient to enable preferences for rewarding and preferences for punishing to survive the evolutionary competition with self-interested preferences. Preferences for rewarding as well as preferences for punishing can induce cooperative behavior. But there is an intrinsic difference between both preference types: Preferences for rewarding tend to coexist with self-interested preferences, whereas preferences for punishing tend either to dominate the population completely or to vanish entirely. Furthermore, rewarders enhance the evolution of preferences for punish-

[36]In case of the evolution of "genes for reciprocity" one may object that the discount factor should be derived endogenously from the process of evolution. However, at least in the context of cultural evolution an exogenously given discount rate seems a reasonable assumption.

ing. Preferences for rewarding are able to invade a self-interested population and may then as a "catalyst" enable the invasion of preferences for punishing. Punishers on the other hand crowd out rewarders and may even drive them out completely.

A Some Definitions from Evolutionary Game Theory

Most standard concepts in evolutionary game theory are formulated for the evolution of strategies. Here we look at the evolution of preferences. The main reason is that this term captures better the aim of our paper - we want to understand why people might have reciprocal preferences. The following definitions are analogous to their counterparts in evolutionary game theory.

Let $t_1, t_2, \ldots, t_n$ be a finite number of possible preference types and $\gamma_1, \gamma_2, \ldots, \gamma_n$ their fractions of the total population, i.e. $\gamma_1, \gamma_2, \ldots, \gamma_n \geq 0$ and $\sum_{i=1}^{n} \gamma_i = 1$. We call the vector $\gamma \equiv (\gamma_1, \gamma_2, \ldots, \gamma_n)$ a population state. The set of all possible population states is therefore a $n-1$ dimensional simplex in R^n. We call this set Δ. The following definitions concerning selection dynamics is analogous to the ones commonly used in evolutionary game theory (see e.g. Weibull [32]). We focus on continuous selection dynamics defined on the simplex Δ in terms of growth rates $g_i(\gamma)$ for the population shares associated with each preference type $i \in n$ as follows

$$\dot{\gamma}_i = g_i(\gamma)\gamma_i \tag{15}$$

where g is a function with open domain X containing Δ.

Definition 3 *A* **regular growth rate function** *is a Lipschitz continuous function $g : X \to R^n$ with open domain X containing Δ, such that $g(\gamma) \cdot \gamma = 0$ for all $\gamma \in \Delta$*

For any regular growth rate function there exists a unique solution $\xi(t, \gamma^0)$ to equation 15 through any initial value γ^0. Moreover ξ is continuous in $t \in T$ and $\gamma^0 \in \Delta$ (Picard-Lindelöf theorem).

Definition 4 *A regular growth rate is called* **payoff monotonic** *if for all $\gamma \in \Delta$*

$$u(t_i, \gamma) < u(t_j, \gamma) \Leftrightarrow g_i(\gamma) < g_j(\gamma), \tag{16}$$

where $u(t_i, \gamma)$ stands for the average material payoff of type t_i when the state of the total population is γ.

Hence payoff monotonicity means that the fraction of types receiving higher average material payoffs grow with a higher rate.

To check the stability of a population state, we look at asymptotic stability. We will refer in all proofs to the metric induced by the maximum-norm. The proofs would extend straightforward to other metrics (e.g. the Euclidian-metric).

Definition 5 *A population-state γ is called* **Lyapunov stable** *if every neighborhood B of γ contains a neighborhood B^0 of γ such that $\xi(t, \gamma^0) \in B$ for all $\gamma^0 \in B^0 \cap C$ and $t \geq 0$.*
A state $\gamma \in C$ is called **asymptotically stable** *if it is Lyapunov stable and there exists a neighborhood B^* such that $\lim_{t\to\infty} \xi(t, \gamma^0) = \gamma$ holds for all $\gamma^0 \in B^* \cap C$.*

Definition 6 *A* **closed set** $A \subset C$ *is* **Lyaponnov stable** *if every neighborhood* B *of* A *contains a neighborhood* B^0 *of* A *such that* $\gamma_+(B^0 \cap C) \subset B$. *A closed set* $A \subset C$ *is* **asymptotically stable** *if it is Lyapunov stable and if there exists a neighborhood* B^* *of* A *such that* $\xi(t, x^0)_{t\to\infty} \to A$ *for all* $x^0 \in B^* \cap C$.

B Proofs

B.1 Proof of Proposition 1

Player 2 conditions his choice of action only on the action chosen by player 1. In particular, he does not condition his behavior on the distribution of preference-types in his group, even if he could do so. This fact is due to our assumption that players have preferences only over outcomes and not over other players preferences or their distribution. Since player 2 can observe the action of player 1, he can guess the outcomes resulting to his own action directly. Hence, player 2 chooses his strategy independent of the distribution of preference-types in his group. Now assume that a positive fraction of players 1 acts in a way that earns them an expected material payoff strictly lower than the maximum. Then, the population cannot be stable since expected material payoff maximizers can either invade in this population or their fraction grows, q.e.d.

B.2 Proof of Proposition 3

Equation 6 describes the difference in expected material payoffs as a function of the fraction γ of rewarders in the entire population. N is kept constant. The first term of the right hand side has a positive sign, the remaining terms are negative. First we consider small γ. The positive first term is of the order k^* in γ whereas the remaining negative terms are at least of the order (k^*+1) in γ. Hence, the right hand side of equation 6 is positive for sufficient small γ. Therefore, γ grows if the fraction of the rewarders is sufficiently small. In other words, an entirely self-interested population is not stable.

In a similar way we prove that the fraction of self-interested players increases if most players have preferences for rewarding, i.e. if γ is close to 1. If γ converges to 1 then the first term of equation 6 converges to zero whereas the remaining sum of negative terms converges to $(-c_r)$ (in fact, the last term converges to $-c_r$ and all the remaining terms to zero). Hence, the fraction γ of rewarders decreases if their fraction of the total population is sufficiently large. In other words, a population entirely of rewardes is not stable either, q.e.d.

B.3 Proof of Theorem 1

From proposition 3 we know that the difference in the average material payoffs $\bar{u}_{pos}(\gamma) - \bar{u}_s(\gamma)$ is above zero for small γ and below zero for γ close to 1. Since $\bar{u}_{pos}(\gamma) - \bar{u}_s(\gamma)$ is continuous in

γ there must exist an interior γ^{eq} with $\bar{u}_{pos}(\gamma^{eq}) - \bar{u}_s(\gamma^{eq}) = 0$. We will see in the next step, that γ^{eq} is the unique value strictly between 0 and 1 satisfying this equation. Hence, for all values below γ^{eq} the difference is above 0 and for all values above γ^{eq} the difference is below 0. Hence, this equilibrium is stable.
Uniqueness follows directly from the necessary condition for an interior equilibrium, i.e. equation 7:

$$c_2 - c_r - d_2 = c_r \sum_{k=1}^{N-1-k^*} \frac{(N-1-k^*)!k^*!}{(N-1-k^*-k)!(k^*+k)!} \left(\frac{\gamma}{1-\gamma}\right)^k. \tag{17}$$

The right hand side is strictly increasing in γ. The left hand side is constant. Therefore, equation 17 is satisfied at most for one γ^{eq}, q.e.d.

B.4 Proof of Proposition 4

We assumed $\frac{k^*}{N} \equiv q$ constant, i.e. $k^* = qN$ with $0 < q < 1$. We can rearrange the equilibrium condition 7 into

$$\begin{aligned} c_2 - c_r - d_2 &= c_r \sum_{k=1}^{N-1-k^*} \left(\left(\prod_{l=1}^{k} \frac{N-k^*-l}{k^*+l} \right) \left(\frac{\gamma}{1-\gamma}\right)^k \right) && (18) \\ &= c_r \sum_{k=1}^{N(1-q)-1} \left(\left(\prod_{l=1}^{k} \frac{(1-q)N-l}{qN+l} \right) \left(\frac{\gamma}{1-\gamma}\right)^k \right). && (19) \end{aligned}$$

Now we prove that for constant γ the right hand side is strictly increasing in N. Since the left hand side is constant, γ^{eq} has to fall in order to equilibrate both sides again.
The number of terms increases with N. Since all terms in equation 18 are positive it is sufficient to prove that each term increases in N. By extending N to real numbers we find

$$\frac{\partial}{\partial N} \left(\frac{(1-q)N-l}{qN+l} \right) = \frac{l}{(qN+l)^2} > 0, \tag{20}$$

q.e.d.

B.5 Proof of Proposition 5

The right hand side of equation 7 is strictly increasing in γ. For any value of c_r equation 7 must hold in equilibrium. If we now choose a new $c_r^{new} > c_r$, the left hand side becomes smaller whereas the right hand side would increase if we keep γ fixed. Therefore, γ has to decrease in order to decrease the right side and satisfy equation 7 again. Hence the new equilibrium fraction of rewarders is lower. Furthermore, if c_r tends to 0, then the left hand side tends to the positive value $c_2 - d_2$, whereas the right hand side would tend to zero if $\frac{\gamma}{1-\gamma}$ remained bounded from above. Therefore, γ must tend to 1 if c_r tends to zero. Finally, if c_r

tends to $(c_2 - d_2)$, then the left hand side tends to zero, but the right hand side can only tend towards zero if γ tends to zero, too, q.e.d.

B.6 Proof of Proposition 6

The proof of proposition 6 is completely analogous to the proof of proposition 5.

B.7 Proof of Lemma 1

We consider the equilibrium condition in form of equation 18. The left hand side is not effected by a change in k^*. The right hand side is effected in two ways if k^* increases. First, the number of terms is reduced and second, each of the remaining terms becomes smaller. Both effects diminish the value of the right hand side. Therefore, γ^{eq} has to increase in order to equilibrate both sides again, q.e.d.

B.8 Proof of Corollary 1 and Corollary 2

$k^* = [N\frac{d_1 - c_1}{r}]$ is weakly increasing in $(d_1 - c_1)$ and weakly decreasing in r. Hence, corollary 1 and 2 follow directly from lemma 1, q.e.d.

B.9 Proof of Proposition 7

An entirely self-interested population can be invaded by preferences for rewarding and vice versa (case 1). Furthermore, an entirely reciprocal population can be invaded by preferences for punishing (then cooperation occurs still in all groups, but preferences for punishing save the costs of rewarding). Only a monomorphic population of punishers cannot be invaded by other preferences (Self-interested preferences cannot invade as shown in case 2. Reciprocal preferences or preferences for rewarding cannot invade either. Cooperation would still occur in all groups. Hence, rewarding is costly but the threat to punish is for free).

B.10 Proof of Proposition 8

Step 1 proves that the equilibria in proposition 8 are stable. Step 2 shows that no other equilibrium can be stable (in the subspace without reciprocal preferences). Step 3 proves that in case a) the population converges from any interior state to a monomorphic population of punishers.

Step 1: The monomorphic population of punishers forms a stable equilibrium by prop. 7. Theorem 1 in case 1 states that there exists a unique stable equilibrium with a fraction of γ^{eq} rewarders in the subspace of self-interested preferences and preferences for rewarding. It remains to be shown that for $\gamma^{eq} < \gamma^{cut}$ (i.e. in case b)) this equilibrium is stable also in the subspace which includes preferences for punishing. Let γ_+ be the fraction of rewarders,

γ_- the fraction of punishers in the total population and define $\tilde{\gamma} \equiv \gamma_+ + \gamma_-$. Due to $p = r$ (assumption 5) the action of player 1 depends only on the total fraction $\tilde{\gamma}$ of rewarders and punishers. Hence, expected material payoffs of any type of player 2 depend only on $\tilde{\gamma}$, too:

$$\bar{u}_s(\tilde{\gamma}) = d_2 \sum_{k=0}^{k^*-1} B_{N-1,\tilde{\gamma}}(k) + d_2 B_{N-1,\tilde{\gamma}}(k^*) + c_2 \sum_{k=k^*+1}^{N-1} B_{N-1,\tilde{\gamma}}(k) \quad (21)$$

$$\bar{u}_+(\tilde{\gamma}) = d_2 \sum_{k=0}^{k^*-1} B_{N-1,\tilde{\gamma}}(k) + (c_2 - c_r)\, B_{N-1,\tilde{\gamma}}(k^*) + (c_2 - c_r) \sum_{k=k^*+1}^{N-1} B_{N-1,\tilde{\gamma}}(k) \quad (22)$$

$$\bar{u}_-(\tilde{\gamma}) = (d_2 - c_p) \sum_{k=0}^{k^*-1} B_{N-1,\tilde{\gamma}}(k) + c_2 B_{N-1,\tilde{\gamma}}(k^*) + c_2 \sum_{k=k^*+1}^{N-1} B_{N-1,\tilde{\gamma}}(k) \quad (23)$$

Hence, we obtain for the differences in average material payoffs

$$\bar{u}_+(\tilde{\gamma}) - \bar{u}_s(\tilde{\gamma}) = (c_2 - d_2 - c_r)\, B_{N-1,(\tilde{\gamma})}(k^*) + (-c_r) \sum_{k=k^*+1}^{N-1} B_{N-1,(\tilde{\gamma})}(k) \quad (24)$$

$$\bar{u}_-(\tilde{\gamma}) - \bar{u}_s(\tilde{\gamma}) = -c_p \sum_{k=0}^{k^*-1} B_{N-1,(\tilde{\gamma})}(k) + (c_2 - d_2)\, B_{N-1,(\tilde{\gamma})}(k^*) \quad (25)$$

$$\bar{u}_-(\tilde{\gamma}) - \bar{u}_+(\tilde{\gamma}) = -c_p \sum_{k=0}^{k^*-1} B_{N-1,(\tilde{\gamma})}(k) + c_r \sum_{k=k^*}^{N-1} B_{N-1,(\tilde{\gamma})}(k) \quad (26)$$

The first equation corresponds to equation 7 of case 1 - only γ is now replaced by $(\tilde{\gamma})$. In particular, the fraction of rewarders will increase relative to the fraction of self-interested preferences if $\tilde{\gamma} < \gamma^{eq}$ (and decrease if $\tilde{\gamma} > \gamma^{eq}$). Similarly, the second equation corresponds to equation 13 of case 2. In particular, this means that the fraction of punishers decreases relative to the fraction of self-interested players as long as $\tilde{\gamma} < \gamma^{cut}$. By putting these two observations together we see that for $\gamma^{eq} < \tilde{\gamma} < \gamma^{cut}$ both the fraction of preferences for rewarding γ_+ as well as the fraction of preferences for punishing γ_- decrease relative to the fraction $(1-\gamma_+-\gamma_-)$ of self-interested preferences. Hence, $(\tilde{\gamma})$ decreases in absolute terms. The dynamic is continuous in γ_+ and γ_- and therefore also in $(\tilde{\gamma})$. Hence, if initially $\gamma_+^0 + \gamma_-^0 < \gamma^{cut}$ then $(\tilde{\gamma})$ remains below (or equal to) $\min\{\gamma_+^0 + \gamma_-^0, \gamma^{eq}\}$. In particular, γ_- decreases relative to the fraction of self-interested preferences with a rate strictly above a constant strictly positive rate. Hence, also γ_- converges absolutely to zero. Now it is straightforward to prove Lyapunov- and asymptotic-stability of the population-state $(\gamma_+, \gamma_-, \gamma_s) = (\gamma^{eq}, 0, 1 - \gamma^{eq})$. Let $\epsilon < \frac{\gamma^{cut} - \gamma^{eq}}{3}$. For any initial population state in the ϵ-neighborhood of $(\gamma^{eq}, 0, 1 - \gamma^{eq})$ holds $\gamma_+^0 + \gamma_-^0 < \gamma^{cut}$. Therefore, γ_- converges to zero. Due to the continuity of all average payoff functions in γ_+ and γ_- the convergence of γ_- to 0 implies convergence of γ_+ to γ^{eq}.

Step 2: Now we prove that there are no other equilibria in the subspace of preferences for rewarding, preferences for punishing and self-interested preferences. First preferences for punishing cannot coexist with self-interested preferences in a stable equilibrium: Replacing any small fraction of the self-interested player by punishers enhances material payoffs of punishers relative to self-interested players. Hence, an equilibrium containing both types cannot be stable. Furthermore, preferences for rewarding and preferences for punishing cannot form an equilibrium: Cooperation would occur in all groups and rewarders earn less because they have to bear the costs of rewarding. A completely rewarding or a completely self-interested population is not stable as shown in case 1. Hence, the only remaining candidates for stable equilibria are either a population of only punishers (in fact, this equilibrium is stable by prop. 7) or a heterogenous population of self-interested preferences and preferences for rewarding. Case 1 showed that in equilibrium the fraction of preferences for rewarding has to be γ^{eq} and the fraction of self-interested preferences $1-\gamma^{eq}$. In step 1 we showed that this equilibrium is stable for $\gamma^{eq} < \gamma^{cut}$. It remains to be shown that this equilibrium is not stable for $\gamma^{eq} > \gamma^{cut}$. This follows from case 2: If $\gamma^{eq} > \gamma^{cut}$ then preferences for punishing are in this equilibrium more successful than self-interested preferences and can therefore invade successfully.

Step 3 It remains to be shown that for $\gamma^{eq} > \gamma^{cut}$ the population converges from any interior state to the equilibrium of a monomorphic population of punishers. From any interior state and for any regular selection dynamics the population state does not reach the boundaries in finite time[37], i.e. no preference-type vanishes completely in finite time. If initially $\tilde{\gamma} < \gamma^{eq}$ then γ_+ grows with positive rate relative to self-interested preferences as long as $(\tilde{\gamma}) < \gamma^{eq}$ and in particular the point where $(\tilde{\gamma}) = \gamma^{cut} + \frac{\gamma^{eq}-\gamma^{cut}}{2}$ is reached in finite time. In the area where $\gamma^{cut} < (\tilde{\gamma}) < \gamma^{eq}$ preferences for rewarding and preferences for punishing are both more successful than self-interested preferences. Therefore, once $(\tilde{\gamma}) \geq \gamma^{cut} + \frac{\gamma^{eq}-\gamma^{cut}}{2}$ holds, the dynamic process never changes this. Hence, after a finite time γ_- increases with a strictly positive rate compared to γ_s. In particular, this implies that γ_s converges to 0. Hence, $(\tilde{\gamma})$ converges to 1 and in particular $(\tilde{\gamma}) > \gamma^{eq}$ after some finite time. Then, rewarding players 2 become less successful than self-interested players 2 (and therefore less successful than punishing players 2) and converge to 0, too. In the end only preferences for punishing survive and the fractions of other preferences converge to zero.

B.11 Proof of Proposition 9

Average material payoffs of each type are continuous in γ_+, γ_-, γ_{rc} and γ_s. Hence, if a single invader of one type earns strictly less than incumbents, then this holds also for a sufficiently small fraction of invaders.

[37] Compare e.g. Weibull [32] page. 141

Part a): A population of punishers is stable in the subspace without reciprocal players (proposition 8). Furthermore, reciprocal players cannot invade either: In this equilibrium of only preferences for punishing cooperation occurs in all groups. Therefore, an invading reciprocal player has no further pivotal benefit of being a rewarder in addition to being a punisher. But he has to pay costs of rewarding. Hence, he is less successful and vanishes again.

Part b): The proof of proposition 8 shows that for $\gamma^{eq} < \gamma^{cut}$ an invading punisher earns a strictly lower average payoff than any incumbent. Hence, it is sufficient to show that for $\gamma_{eg} < \gamma^h$ a reciprocal invader earns on average strictly less than any incumbent. Then, stability of the equilibrium follows by arguments analogous to the proof of proposition 8. The average material payoffs of a rewarder or a single reciprocal invader in this equilibrium are

$$\bar{u}_+ = d_2 \sum_{k=0}^{k^*-1} B_{N-1,\gamma^{eq}}(k) + (c_2 - c_r) \sum_{k=k^*}^{N-1} B_{N-1,\gamma^{eq}}(k) \tag{27}$$

$$\bar{u}_{rc} = (d_2 - c_p) \sum_{k=0}^{k^*-2} B_{N-1,\gamma^{eq}}(k) + (c_2 - c_r) \sum_{k=k^*-1}^{N-1} B_{N-1,\gamma^{eq}}(k) \tag{28}$$

Hence,

$$\bar{u}_{rc} - \bar{u}_+ = -c_p \sum_{k=0}^{k^*-2} B_{N-1,\gamma^{eq}}(k) + (c_2 - d_2 - c_r) B_{N-1,\gamma^{eq}}(k^* - 1). \tag{29}$$

For $0 < \gamma^{eq} < 1$ we obtain by dividing through $B_{N-1,\gamma^{eq}}(k^* - 1)$ the equivalence

$$\bar{u}_{rc} - \bar{u}_+ \gtreqless 0 \tag{30}$$

$$\Leftrightarrow c_2 - d_2 - c_r \gtreqless c_p \sum_{k=0}^{k^*-2} \frac{(k^*-1)!}{k!} \frac{(N-k^*)!}{(N-1-k)!} \left(\frac{1-\gamma^{eq}}{\gamma^{eq}} \right)^{k^*-1-k}. \tag{31}$$

The right hand side of equation 31 is strictly decreasing in γ^{eq}. Furthermore, the right hand side would be equal to the left hand side if $\gamma^{eq} = \gamma^h$ (this was precisely the definition of γ^h). Hence, for $\gamma^{eq} < \gamma^h$ the right hand side is strictly larger than the left hand side and therefore $\bar{u}_{rc} - \bar{u}_+ < 0$, q.e.d.

C Comparative Statics of γ^{cut} in case 2

Let γ^{cut} be the fraction of punishers in the unstable mixed equilibrium. This fraction separates the basins of attraction of the stable equilibria. If the initial fraction of punishing players is below the cutoff γ^{cut} then this fraction decreases until the entire population has self-interested preferences and nobody punishes defection. If on the other hand the initial fraction of punishing players is above the cutoff γ^{cut} then this fraction increases until the entire population

has preferences for punishing. One might therefore interpret the value of γ^{cut} as an indicator how likely it is to end up in one or the other equilibrium[38]. The comparative statics of γ^{cut} is analogous to case 1 and can be derived directly from equation 13.

Higher costs of punishing diminish the basin of attraction of the punisher equilibrium:

Proposition 10 *If the costs c_p - which a player 2 has to bear in order to punish - increase, then γ^{cut} increases, i.e. there have to be initially more punishers in order to end up in the punishing equilibrium. Furthermore,* $\lim_{c_p \to 0} \gamma^{cut} = 0$ *and* $\lim_{c_p \to \infty} \gamma^{cut} = 1$.

The intuition is straightforward: The higher the number of punishing players the cheaper it is to be a punisher. If the costs of punishing increase, punishers become less fit. Hence, punishing players need a higher fraction of punishers in order to be at least as successful as non-punishers.

Higher gains from cooperation for player 2 are good for punishers. Hence, the basin of attraction for their equilibrium becomes larger:

Proposition 11 *If the gains of cooperation for the players 2 $(c_2 - d_2)$ increase, then γ^{cut} decreases, i.e. a lower initial fraction of punishing players is necessary in order to end up in the punishing equilibrium. Furthermore,* $\lim_{(c_2 - d_2) \to 0} \gamma^{cut} = 1$ *and* $\lim_{(c_2 - d_2) \to \infty} \gamma^{cut} = 0$.

Again, the intuition is straightforward: The higher the gains of cooperation for a player 2, the higher his profit from being pivotal in inducing cooperation of players 1. Therefore, a lower fraction of punishers is necessary in order to make punishing more successful than non-punishing.

Lemma 2 *If the threshold k^{**} of punishing players 2 in a group above which the players 1 start to cooperate increases then γ^{cut} increases, i.e. there are more punishing players necessary in order to end up in the punishing equilibrium.*

Intuitively, a higher threshold k^{**} makes it more probable to be in a group in which the number of punishers is too low to induce cooperation. In these groups being a punisher is costly. Therefore, fitness of punishers is lower and a higher initial fraction of punishers is necessary to make punishing more successful than non-punishing.

Corollary 3 *If player 1s costs for cooperation $(d_1 - c_1)$ increase then γ^{cut} increases weakly, i.e. a higher or equal fraction of punishers is necessary in order to end up in the punishing equilibrium.*

Corollary 4 *If player 2s losses due to a punishment p increase, then γ^{cut} decreases weakly.*

[38] Again, this interpretation is in the spirit of the model by Kandori et. al [20], where the size of the basins of attraction determines the long run equilibrium

D Extension: Small Mistakes

This appendix considers the possibility that some players 1 fail to play optimally. Assume that player 1 trembles with small probability ϵ. In that case he defects even so he should cooperate and vice versa. If ϵ is sufficiently small results of case 1 and case 2 change only slightly:

Proposition 12 *If player 1 trembles with sufficiently small probability ϵ then*

in Case 1 *there exist two stable equilibria: In the first equilibrium the fraction γ_ϵ^{eq} of preferences for rewarding is close to the equilibrium fraction without mistakes γ^{eq}. In the second equilibrium only self-interested preferences survive (i.e. $\gamma = 0$). If ϵ tends to zero then γ_ϵ^{eq} tends to γ^{eq}. Moreover, the basin of attraction of the self-centered equilibrium tends to zero.*

in Case 2 *there remain two stable equilibria. The monomorphic equilibrium where all players have self-interested preferences is still stable. But a monomorphic population of punishers is not stable any more. Instead, there is a second stable equilibrium with a high fraction of punishers and a low fraction of self-interested players. If ϵ tends to zero, the fraction of preferences for punishing in this equilibrium is arbitrarily close to 1.*

We discuss and prove only Case 2. The proof for Case 1 is analogous and written upon request.

The intuition for Case 2 of proposition 12 is straightforward. In a world of no mistakes and in the equilibrium where all players 2 are willing to punish, this threat is costless: Player 1 cooperates and no punishment is necessary. But if players 1 make sometimes mistakes being a punisher is costly. If almost everybody else is a punisher the probability of being pivotal tends to zero. But due to mistakes the costs of punishing do not vanish. A monomorphic population of punishers is therefore no longer stable.

For a more formal proof consider average payoffs of both types. A self-interested player 2 receives an average material payoff of

$$\bar{u}_s(\gamma) = (d_2 + \epsilon\,(c_2 - d_2)) \sum_{k=0}^{k^{**}} B_{N-1,\gamma}(k) + (c_2 - \epsilon\,(c_2 - d_2)) \sum_{k=k^{**}+1}^{N-1} B_{N-1,\gamma}(k) \tag{32}$$

and the punishing type receives

$$\bar{u}_-(\gamma) = (d_2 - c_p + \epsilon\,(c_2 - d_2 + c_p)) \sum_{k=0}^{k^{**}-1} B_{N-1,\gamma}(k) + (c_2 - \epsilon\,(c_2 - d_2 + c_p)) \sum_{k=k^{**}}^{N-1} B_{N-1,\gamma}(k). \tag{33}$$

Hence, the difference in average payoffs between both types is

$$\begin{aligned}
& \bar{u}_-(\gamma) - \bar{u}_s(\gamma) \\
= & -(1-\epsilon)\, c_p \sum_{k=0}^{k^{**}-1} B_{N-1,\gamma}(k) + ((1-2\epsilon)\,(c_2-d_2) - \epsilon c_p)\, B_{N-1,\gamma}(k^{**}) \\
& -\epsilon c_p \sum_{k=k^{**}+1}^{N-1} B_{N-1,\gamma}(k) \\
= & (1-2\epsilon)\left(-c_p \sum_{k=0}^{k^{**}-1} B_{N-1,\gamma}(k) + (c_2-d_2)\, B_{N-1,\gamma}(k^{**})\right) - \epsilon c_p. \qquad (34)
\end{aligned}$$

This difference is continuous in γ and ϵ. For $\gamma = 0$ the difference is negative. Hence, the monomorphic equilibrium of self-interested preferences remains stable. Also for $\gamma = 1$ the difference is negative. Therefore, self-interested preferences can invade a population of punishers. However, for ϵ sufficiently small there still exists a second stable equilibrium in addition to $\gamma = 0$. In this second equilibrium punishers and self-interested types coexist. The fraction of punishing types in this equilibrium converges to 1 if ϵ tend to 0.

Proof: Existence: For $\epsilon = 0$ there exits a γ_0 where the difference is positive. Due to continuity in ϵ the difference at this γ_0 is still positive for sufficiently small ϵ. Since the difference is negative at $\gamma = 1$ there must exists a stable equilibrium between γ_0 and 1 due to continuity in γ.

Exactly one more stable equilibrium: The term in large brackets in equation 34 is a polynomial of finite order. Hence, there are only a finite number of local minima. Let Δ be the minimum value of all local minima above zero. For $\epsilon < \frac{\Delta}{2\Delta + c_p}$ we obtain $(1-2\epsilon)\Delta - \epsilon c_p > 0$ and therefore all local minima with positive value remain positive. Hence, for sufficiently small ϵ there are still only two γ for which $\bar{u}_-(\gamma) - \bar{u}_s(\gamma) = 0$ - one (still unstable) equilibrium close to the old unstable equilibrium and one stable equilibrium close to $\gamma = 1$, q.e.d.

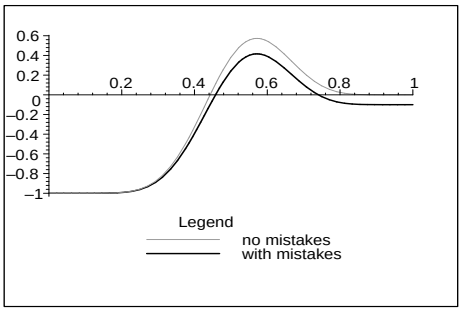


Figure 10: Case 2 with mistakes of probability $\epsilon = 0.1$ and with $\gamma^{eq} > \gamma^{cut}$: $(\bar{u}_+ - \bar{u}_s)$ and $(\bar{u}_- - \bar{u}_s)$ as functions of $\gamma \equiv \gamma_+ + \gamma_-$ for $N = 20, d_1 = 1, c_1 = 0, r = 2, d_2 = 5, c_2 = 0, c_r = 1$.

Equation 34 is also helpful to analyze the case of mistake probabilities which are not arbitrarily small, but of moderate size. We can adjust the payoff difference $(\bar{u}_- - \bar{u}_s)$ by re-scaling it slightly with $(1-2\epsilon)$ and then shifting it downwards by ϵc_p. Figure 10 demonstrates this for our example of case 2 with a mistake probability of $\epsilon = 0.1$. Here, in the stable "punisher equilibrium" with mistakes a fraction of $\gamma_- \approx 0.74$ has preferences for punishing, but a fraction of $(1-\gamma_-) \approx 0.26$ has self-interested preferences.

E Further equilibria in case 3

First we derive a sufficient condition under which there are no other stable equilibria than those of proposition 9. Second we analyze under which conditions there exist stable equilibria consisting only of reciprocal and self-interested preferences and third we give the intuition why no further stable equilibria exist.

The following lemma limits the possible candidates for stable equilibria.

Lemma 3 *An equilibrium with a positive fraction γ_{rc} of reciprocal players can only be stable if the fraction of γ_s self-interested players is also positive.*

Proof: If $\gamma_s = 0$ then players 1 cooperate in all groups. Therefore, preferences for punishing earn c_1 in all groups whereas reciprocal preferences earn only $c_1 - c_r$. Hence, the equilibrium is not stable, q.e.d.

In case 3 players 1 cooperate in their group if and only if $c_1 + \frac{k_+ + k_{rc}}{N} r > d_1 - \frac{k_- + k_{rc}}{N} r$, i.e. $k_+ + k_- + 2k_{rc} > N\frac{d_1 - c_1}{r}$. We define $k_{ef} \equiv k_+ + k_- + 2k_{rc}$, $\tilde{\gamma} \equiv \gamma_+ + \gamma_-$ and $W(k_{ef}) \equiv W_{N-1,\gamma_- + \gamma_+, \gamma_{rc}}(k_{ef})$ as the probability that a group of $N-1$ players has the characteristic k_{ef}, i.e.

$$W(k_{ef}) = \sum_{\substack{\tilde{k}, k_{rc}=0 \\ \tilde{k}+2k_{rc}=k_{ef}}}^{N-1} \binom{N-1}{\tilde{k}, k_{rc}} \tilde{\gamma}^{\tilde{k}} \gamma_{rc}^{k_{rc}} \left(1 - \tilde{\gamma} - \gamma_{rc}\right)^{N-1-\tilde{k}-k_{rc}}. \tag{35}$$

The probabilities of a group having any characteristic k_{ef} must be 1, i.e.

$$\sum_{k_{ef}=0}^{2N-2} W(k_{ef}) = 1. \tag{36}$$

We can write down average material payoffs in new notation

$$\bar{u}_{rc} = (d_2 - c_p) \sum_{k_{ef}=0}^{k^*-2} W(k_{ef}) + (c_2 - c_r) \sum_{k_{ef}=k^*-1}^{2N-2} W(k_{ef}) \tag{37}$$

$$\bar{u}_{+} = d_2 \sum_{k_{ef}=0}^{k^*-1} W(k_{ef}) + (c_2 - c_r) \sum_{k_{ef}=k^*}^{2N-2} W(k_{ef}) \tag{38}$$

$$\bar{u}_{-} = (d_2 - c_p) \sum_{k_{ef}=0}^{k^*-1} W(k_{ef}) + c_2 \sum_{k_{ef}=k^*}^{2N-2} W(k_{ef}) \tag{39}$$

$$\bar{u}_{s} = d_2 \sum_{k_{ef}=0}^{k^*} W(k_{ef}) + c_2 \sum_{k_{ef}=k^*+1}^{2N-2} W(k_{ef}) \tag{40}$$

An equilibrium with $\gamma_{rc} > 0$ can only be stable if $\bar{u}_{rc} = \bar{u}_s$ (by lemma 3) and if $\bar{u}_+ \leq \bar{u}_{rc}$ and $\bar{u}_- \leq \bar{u}_{rc}$:

Lemma 4 *The following conditions are all necessary for a stable equilibrium with a fraction $\gamma_{rc} > 0$:*

1.

$$c_p \sum_{k_{ef}=0}^{k^*-2} W(k_{ef}) + c_r \sum_{k_{ef}=k^*-1}^{2N-2} W(k_{ef}) = (c_2 - d_2) \left(W\left(k^* - 1\right) + W\left(k^*\right) \right) \tag{41}$$

2.

$$c_p \sum_{k_{ef}=0}^{k^*-2} W(k_{ef}) \leq (c_2 - d_2 - c_r) W(k^* - 1) \tag{42}$$

3.

$$c_r \sum_{k_{ef}=k^*} W(k_{ef}) \leq (c_2 - d_2 - c_r + c_p) W(k^* - 1) \tag{43}$$

The next corollary follows directly from condition 41:

Corollary 5 *If*

$$\sup_{\substack{\gamma_{rc} \in [0,1] \\ (\gamma_+ + \gamma_-) \in [0, 1-\gamma_{rc}]}} \left(W(k^* - 1) + W(k^*) \right) < \frac{\min\{c_p, c_r\}}{c_2 - d_2} \tag{44}$$

then there is no stable equilibrium with $\gamma_{rc} > 0$, i.e. the stable equilibria of proposition 9 are the only ones.

We now analyze under which conditions there exists a stable equilibrium consisting only of

self-interested and reciprocal preferences, i.e. $\gamma_+ = \gamma_- = 0 = \gamma_+ + \gamma_-$. Notice that

$$W_{N-1,0,\gamma_{rc}}(k_{ef}) = \begin{cases} 0 & \text{if } k_{ef} \text{ odd} \\ \frac{(N-1)!}{\left(\frac{k_{ef}}{2}\right)!\left(N-1-\frac{k_{ef}}{2}\right)!}\gamma_{rc}^{\left(\frac{k_{ef}}{2}\right)}\left(1-\gamma_{rc}\right)^{N-1-\frac{k_{ef}}{2}} & \text{if } k_{ef} \text{ even.} \end{cases} \tag{45}$$

It follows directly that for even k^* condition 42 and condition 43 cannot be both fulfilled. Hence for even k^* there is no stable equilibrium consisting only of self-interested and reciprocal preferences.

For odd k^* on the other hand things are different. If an equilibrium with a positive fraction of reciprocal and self-interested preferences is stable in the subspace of this two types, it is stable also against invasion of preferences for rewarding or preferences for punishing. This can be seen directly from average payoffs. In such an equilibrium the probability $W_{N-1,0,\gamma_{rc}}(k_{ef} = k^*)$ of a group with characteristic $k_{ef} = k^*$ is 0, but these are the only groups where preferences for rewarding or preferences for punishing perform better than self-interested ones. To check existence of a stable equilibrium in the subspace of self-interested and reciprocal preferences we just have to consider differences in average material payoffs of this two types, i.e.

$$\begin{aligned} &\bar{u}_{rc}(\gamma_{rc}) - \bar{u}_s(\gamma_{rc}) = \\ &c_p \sum_{k=0}^{k^*-2} B_{N-1,\gamma_{rc}}(k) + (c_2 - d_2 - c_r)\left(B_{N-1,\gamma_{rc}}(k^*-1) + B_{N-1,\gamma_{rc}}(k^*)\right) + c_r \sum_{k=k^*+1}^{N-1} B_{N-1,\gamma_{rc}}(k) \end{aligned} \tag{46}$$

For $k^* \geq 2$ this difference is negative for γ_{rc} sufficiently close to zero or one. Hence, there exists a stable equilibrium if and only if

$$\sup_{\gamma_{rc}\in[0,1]} \left(\bar{u}_{rc}(\gamma_{rc}) - \bar{u}_s(\gamma_{rc})\right) > 0. \tag{47}$$

This is summarized in the following

Corollary 6 *Consider case 3:*

a) *For even k^* there is no stable equilibrium consisting only of reciprocal and self-interested preferences.*

b) *For odd k^* and $k^* \geq 2$ there exists a stable equilibrium of only self-interested and reciprocal preferences if and only if*

$$\sup_{\gamma_{rc}\in[0,1]} \left(\bar{u}_{rc}(\gamma_{rc}) - \bar{u}_s(\gamma_{rc})\right) > 0. \tag{48}$$

Finally, we give the intuition why there are no stable equilibria with a mixture of three or all four different preference-types in case 3. An equilibrium with positive fractions of self-interested preferences and preferences for punishing is not stable because preferences for

punishing become more successful than self-interested ones if a small deviation in their favor occurs. Similarly, an equilibrium with positive fractions of preferences for rewarding and reciprocal preferences is not stable because small deviation in favor of reciprocal preferences make reciprocal preferences more successful than preferences for rewarding. Hence, in case 3 maximally two preference types coexist in equilibrium.

References

[1] Benaim, M. and Weibull, J.W. (2003),"Deterministic Approximation of Stochastic Evolution in Games" *Econometrica* Vol. 71, 873-903

[2] Bergstrom, Theodore C. (2001), "The Algebra of Assortative Encounters and the Evolution of Cooperation", to appear in: *International Game Theory Review*

[3] Bergstrom, Theodore C. (2002), "Evolution of Social Behavior: Individual and Group Selection", *Journal of Economic Perspectives* Vol. 16, 2, pp. 67-88

[4] Bester, Helmut and Güth, Werner (1998),"Is altruism evolutionary stable?", *Journal of Economic Behavior and Organization* 34, 193-209

[5] Bolton, G. and Ockenfels, A. (2000), "ERC: A theory of equity, reciprocity and competition", *American Economic Review*, 90, 166-193

[6] Bowles, Samuel and Gintis, Herbert (1998),"The Evolution of Strong Reciprocity",Santa Fee Institute Working Paper 98-08-073E.

[7] Boylan, R. (1992),"Laws of large numbers for dynamical systems with randomly matched individuals", *Journal of Economic Theory* 57, 473-504

[8] Cooper, Ben and Wallace, Chris (2001), "Group Selection and the Evolution of Altruism", Oxford Discussion Paper Series, Nr. 67

[9] Fehr, Ernst and Gächter, Simon (2000),"Fairness and retaliation: The economics of reciprocity", *Journal of Economic Perspectives*, 14, 159-181

[10] Fehr, Ernst and Schmidt, Klaus (2000),"Theories of Fairness and Reciprocity - Evidence and Economic Applications", (paper prepared for the invited session of the 8th World Congress of the Econometric Society)

[11] Frank, R.H. (1987),"If homo economicus could choose his own utility function, would he want one with a conscience?", *American Economic Review* 77, 593-604

[12] Friedman, D. and Singh, N. (1999), "On the viability of vengance", UC Santa Cruz, Mimeo

[13] Gintis, Herbert (2000), "Strong Reciprocity and Human Sociality", Journal of Theoretical Biology 206, 169 - 179

[14] Güth, Werner (1995), "An Evolutionary Approach to Explaining Cooperative Behavior by Reciprocal Incentives", *International Journal of Game Theory* 24, 323-44

[15] Güth, Werner and Yaari (1992), Menahem, "An evolutionary approach to explain reciprocal behavior in a simple strategic game," in: U. Witt(Editor), *Explaining Process and Change: Approaches in Evolutionary Economics*, Ann Arbor: The University of Michigan Press, 23-34

[16] Hart, Sergiu (2000), "Evolutionary dynamics and backward induction", *Games and Economic Behavior* 41, 227-264

[17] Höffler, Felix (1999),"Some play fair, some don't. Reciprocal fairness in a stylized principal-agent problem", *Journal of Economic Behavior & Organization*, Vol.38, 113-131

[18] J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis and R. McElreath (2001), "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies", *American Economic Review*

[19] Huck, S. and Oechssler, J. (1999),"The Indirect Evolutionary Approach to Explaining Fair Allocations", *Games and Economic Behavior* 28, 13-24

[20] M. Kandori, G. Mailath and R.Rob (1993) "Learning, Mutation, and Long-Run Equilibria in Games.", *Econometrica*, 61:29-56

[21] Kuzmics, C. (2002), "Stochastic Evolutionary Stability in Generic Extensive Form Games of Perfect Information", Mimeo

[22] Levine, D.K., (1998),"Modeling altruism and spitefulness in experiments", *Review of Economic Dynamics*, 1, 593-622

[23] Nöldeke, G. and Samuelson, L. (1993), "An evolutionary analysis of backward and forward induction", *Games and Economic Behavior*, 5, 425-454

[24] Ok, E.A. and Vega-Redondo,F. (2001),"On the evolution of individualistic preferences: an incomplete information scenario", *Journal od Economic Theory*, 97, 231-254

[25] Price, G.R. (1970),"Selection and Covariance", *Nature* 277, 520-521

[26] Samuelson,L.(2001),"Analogies, Adaptation, and Anomalies", *Journal of Economic Theory*, 97, 320-366

[27] Sethi, R. (1996), "Evolutionary stability and social norms", *Journal of Economic Behavior and Organization*, 29, 113-140

[28] Sethi, Rajiv and Somananthan, E. (1996), "The Evolution of Social Norms in Common Property Resource Use", *American Economic Review* Vol. 86, 4, 766-788

[29] Sethi, Rajiv and Somananthan, E. (2001), "Preference Evolution and Reciprocity", *Journal of Economic Theory*, Vol. 97, 273-297

[30] Sethi, Rajiv and Somananthan, E. (2003), "Understanding Reciprocity", *Journal of Economic Behavior and Organization*, 50, 1-27

[31] Sober, E. and Wilson, D.S. (1998),"UNTO OTHERS - The Evolution and Psychology of Unselfish Behavior", Cambridge(M.A.): Harvard University Press

[32] Weibull, Jörgen W. (1995) "Evolutionary Game Theory", MIT Press

F Notation

c_1	:	material payoff of player 1 if he cooperates and is not rewarded
c_2	:	material payoff of player 2 if player 1 cooperates and player 2 does not reward
d_1	:	material payoff of player 1 if he defects and is not punished
d_2	:	material payoff of player 2 if player 1 defects and player 2 does not punish
o_1	:	material payoff of player 1 if he chooses his outside option
o_2	:	material payoff of player 2 if player 1 chooses his outside option
c_r	:	costs for player 2 when rewarding
c_p	:	costs for player 2 when punishing
r	:	material gain of being rewarded for player 1
p	:	material loss of being punished for player 1
N	:	number of players 2 in each group
γ_+	:	fraction of rewarders in the total population
γ_-	:	fraction of punishers in the total population
γ_{rc}	:	fraction of reciprocal players in the total population
γ_s	:	fraction of selfish players in the total population
γ^{eq}	:	fraction of rewarders in the stable mixed equilibrium of case 1
γ^{cut}	:	fraction of punishers in the unstable mixed equilibrium of case 2
k_+	:	number of rewarding players 2 in a specific group
k_-	:	number of punishing players 2 in a specific group
k_{rc}	:	number of reciprocal players 2 in a specific group
k_s	:	number of selfish players 2 in a specific group
k^*	:	number of rewarding players 2 $\geq k^* \Rightarrow$ players 1 cooperate (in case 1)
k^{**}	:	number of punishing players 2 $\geq k^{**} \Rightarrow$ players 1 cooperate (in case 2)
$\bar{u}_+(\cdot)$	:	average material payoff of a rewarding player 2
$\bar{u}_-(\cdot)$	:	average material payoff of a punishing player 2
$\bar{u}_{rc}(\cdot)$	:	average material payoff of a reciprocal player 2
$\bar{u}_s(\cdot)$	:	average material payoff of a selfish player 2