SOCIAL PUNISHMENT

MARK HOWARD

October 5, 2005

Abstract. This essay presents a version of the public goods game with punishment. In this version of the game contributions and punishment occur simultaneously. Additionally effort exerted during punishment is effort that would otherwise be available for primary public goods provision. This is the key difference between this paper and previous work. The cost of punishment lowers the level of public good produced but not the relative payoff of the punisher. In this analysis agents gain neither utility nor fitness advantage from other players payoff. Here agents interact anonymously in large groups without the benefit of reputation. The results indicate that depending on the punishment norm multiple stable equilibria exist which are immune to mutant invasion. These include the all defect equilibria which motivates much of the literature as well as mixed equilibria with very high levels of contributions. This is the first time cooperation via punishment has been shown to be stable in a single stage *N-player* game with mutation.

1. Introduction

This essay presents a version of the public goods game with punishment. In this version of the game contributions and punishment occur simultaneously. Additionally effort exerted during punishment is effort that would otherwise be available for primary public goods provision. This is the key difference between this paper and previous work. The cost of punishment lowers the level of public good produced but not the relative payoff of the punisher. In this analysis agents gain neither utility nor fitness advantage from other players payoff. Here agents interact anonymously in large groups without the benefit of reputation. The results indicate that depending on the punishment norm multiple stable equilibria exist which are immune to mutant invasion. These include the all defect equilibria which motivates much of the literature as well as mixed equilibria with very high levels of contributions. This is the first time cooperation via punishment has been shown to be stable in a single stage *N-player* game with mutation.

The key difference between this model and others is that the cost of punishment is borne out of time and effort that could otherwise go into production of the public good. The standard treatment followed by all other models surveyed here on indirect reciprocity and costly punishment assumes that either the time and effort associated with punishment is drawn from a separate source other than that available for the initial public good provisioning or punishment occurs after the public good has been created. Here the punishment occurs instead of public good production. This results in reciprocators always doing at least as well as cooperators in the same group.

Trivers (1971) theory of reciprocal altruism uses the repeated prisoner's dilemma which is the standard explanatory tool for understanding the evolution of cooperation between unrelated individuals. He argues that selfish individuals cooperate, in one example save a drowning man, on the assumption that in the future they will be cooperated with, saved from drowning. According to Trivers what it takes for reciprocal altruism to evolve is that the same two players engage repeatedly. Axelrod and Hamilton (1981) put Trivers argument in an evolutionary perspective. They held a simulated tournament again using the repeated prisoner's dilemma and invited people to submit strategies. They got 62 entries from six countries. Anatol Rapoport's *tit for tat*, the simplest of the strategies submitted, won the contest. This strategy cooperates in the first round and then does what ever it's opponent did in the previous round. When all players uses this strategy full cooperation is the result. Further *tit for tat* induces defectors to cooperate by imposing punishment in the form of withheld cooperation in future rounds. This is the first reciprocal strategy in the literature. This type of reciprocal altruism requires repeated play against the same opponent. In the Axelrod tournaments each pair played 200 rounds against each other. Throughout history reciprocal altruism no doubt played a key role in solving many social dilemmas among related individuals within the same tribe for instance. It does not, however, explained why many human agents cooperate in large groups or when there is little chance of repeated play.

Alexander (1987) proposed moral systems of indirect reciprocity as a way to understanding large-scale human cooperation. He wrote,

> (A) In indirect reciprocity, the return is expected from someone other than the recipient of the beneficence. This return may come from essentially any individual or collection of individuals in the group. (B) Indirect reciprocity, involves reputation and status, and results in everyone in the group continually being assessed and reassessed by interactants past and potential, on the basis of their interactions with others . . . (C) the power of indirect reciprocity [is] in the forms of ethical and moral systems, to manipulate the cost and benefits of the acts of individuals.[1]

Using computer simulations Nowak and Sigmund(1998) model indirect reciprocity as a donation game in which individuals never interact with the same partner twice. They model reputation as an 'image score' which is known to all agents. An agents image is good if they donated in the previous round and bad if they did not. They introduce the Discriminator strategy which donates to others with a good image and refuses to donate to others with a bad image. They find that, even

[1]The game presented here shows that a type of social reciprocity could easily evolve which conforms to parts A and C but not B.

when agents meet only once, this strategy cannot be invaded by defectors. If unconditional altruists are added to the pool, cooperation is unstable and defection is the only evolutionarily stable strategy (ESS).

Panchanathan and Boyd (2003)[2] show that indirect reciprocity cannot be based on an image scoring when players make mistakes. Alternatively they base reputation on standing following Sugden (1986).They call this strategy a reputation discrimination. Trivers pointed to a 'classic problem in social science' which is, should altruism be defined in terms of actions or intentions? The idea of standing attempts to take account of intention. Like an image score standing can be good or bad. All individuals who cooperate in the previous round will have good standing. However defection will put a player in bad standing only if it is not justified. A justified defection will not change the actor's standing. Defection is justified if the other player is in bad standing and unjustified when the other player is in good standing. Panchanathan and Boyd show that this type of indirect reciprocity may be evolutionarily stable given mutation. The standing model as well as the image score model of indirect reciprocity do not require that the individuals be related or meet repeatedly but they do require that they have knowledge of all their different partners recent history. And in the

[2]See also (Leimar and Hammerstein 2001)

case of the standing model their partner's last partner's. This does not explain why humans cooperate with strangers whose reputation they do not know and who they will likely never meet again. Furthermore these models both assume pairwise interactions while humans often find themselves in social dilemmas in large groups.

One solution is to allow many players in the PD game. A multi player PD game is a public goods game. In this game with many players in a group withholding future cooperation after defection means a few defectors in a large group can cause mutual defection forever. So this type of punishment is not well suited to public goods games. This problem has be dealt with by allow cooperative players to punish defectors directly. This is usually done in a second stage, after the initial public goods game. When there are only few selfish types unconditional cooperators earn higher payoffs than punishers. Unconditional cooperators therefore, have higher fitness and displace punishers in the population. Once the portion of punishers falls below a certain level selfish types can then invade. Solutions to this second order problem include costly signaling Gintis, Smith and Bowles (2001) and group selection via wars Boyd, Gintis, Bowles and Richerson (2003) or ostracism Bowles and Gintis (2004a), Cinyabuguma, Page and Putterman (2004a). .

Panchanathan and Boyd (2004) show that indirect reciprocity can lead to cooperation in large groups without the second order free rider problem. They propose a two stage game. The first stage is a public goods game where players can contribute to the public good which is equally valuable to all agents but at a cost that is greater than that value to the individual contributor. The second stage agents play a multi period mutual aid game where in each period one randomly chosen individual can either be helped by all the other members of the group or not. Here rewarding in the second stage, or more precisely withholding reward can enforce cooperation since, rather than costing punisher relative to unconditional cooperators it increases their relative payoff. Again Panchanathan and Boyd use standing as the basis for aid among what they term Shunners. Shunners aid those in good standing and refuse aid to those in bad standing. Here failure to contribute to the public good in the first stage puts one in bad standing forever while failure to aid a reputable player in the second stage puts one in temporary bad standing which can be reverse after aid has been given to a member in good standing. Thus model allows for mistaken non aid but not mistaken non contribution to be forgiven. This model is typical of most attempts to solve the second order free rider problem in that it adds pairwise interactions to the standard multiplayer public goods game.

Individual punishment fails for at least two reasons. First the well documented second order problem is that punishment itself is a public good and in a two stage game the dominant strategy in the second round is not punishing. The other is that defectors who are punished retaliate in future rounds by punishing the the full contributors.(Cinyabuguma, Page and Putterman 2004b)

Fowler (2005) and Hauert, Haiden and Sigmund (2004) have shown that if individuals have an outside option or activity available which yields a positive reward which is less than the cooperators payoff but greater than the all defect payoff then a cycle emerges which is then stabilized at a interior equilibrium given altruistic punishers.

Sethi and Somanthan(1996) look at the Common Pool Resource Problem again within an evolutionary-game-theoretic framework. They show that cooperation guided by a norm of restraint and punishment can be stable against invasion by selfish types. Their model is very similar to the one presented here. There as here the problem of second order punishment is nuanced by assuming punishment and cooperation occur simultaneously.This theory of strong reciprocity relies on the ability of reciprocators to make a credible commitment to sanction norm violators even at a personal cost. In both that model and this one, one of the stable states is all defect. The current model differs from Sethi and Somanthan in that here agents are playing a Nplayer public goods game as opposed to a common pool resource extraction game. Further in their model as well as virtually all other models of cooperation via punishment, the cost of punishment does not lessen the amount of the public good produced. Here the act of punishment by the individual lowers the amount that individual can provide to the public good. The result is that in the Sethi and Somanthan model within any group punishers have a lower payoff than do cooperators(altruist) given any defectors in the population. Where as in the alternative presented here they have the same payoff.

Here we start with the simplest model of public goods provisions with no punishment. Then we add punishment as is done in (Sethi and Somanathan 1996, Bowles and Gintis 2004b) and see how the second order free rider problem occurs. We see that when mutation is allowed the hanging valley near the R-C equilibrium segment is unstable since given the positive number of defectors C always does better than R. Then we modify the analysis so punishment and contributions flow from the same pool of resources. In this case since the payoff to C and R are the same within a group mutation creates a hanging valley parallel to the C-R segment and this is stable.

2. The Public Goods Game

Assume each player can either Cooperate (C) which means in this case to contribute amount c to the public good, or Defect(D) which means to contribute nothing to the public good. The portion of C types is y and the portion of D types is z so $y + z = 1$. Given the number of players in a group, n, and the marginal private return on one unit contributed to the public good m. The game is an n player PD game if $m < 1$ and $m \cdot n > 1$ which is assumed here. The individual benefit from the public good is $b = y \cdot m \cdot n$. Further we assume that n is large enough so that the individuals effect on b when comparing his alternatives is negligible. Let π_i be the payoff to strategy i and let $\overline{\pi}$ be the average payoff in the population.

$$\pi_C = b - c$$
$$\pi_D = b$$
$$\overline{\pi} = y\pi_C + z\pi_D$$

If we add $(c-b)$ to both π_C and π_D the resulting dynamic is unchanged and we get,

$$\pi_C = 0$$
$$\pi_D = c$$
$$\overline{\pi} = y \cdot 0 + zc = zc$$

From which we can write the *replicator equations* following(Gintis 2000a). On the interior

$$\dot{y} = y(\pi_C - \overline{\pi})$$
$$\dot{z} = z(\pi_D - \overline{\pi})$$

$$\dot{y} = -yzc$$
$$\dot{z} = yzc$$

The payoff to defecting is always better by c than to contributing and the only nash equilibrium and the only stable rest point is at all defect. This is the starting point for virtually all of the literature reviewed here.

Next, following the literature we add the reciprocator strategy R with population frequencies x. The Reciprocator contributes c and also randomly punishes one Defector in his group if one is present at a cost of p. This punishment is multiplied by a factor s (stick size) and the result is deducted from the randomly chosen Defectors score. The benefit from the public good is now $b = (y + x) \cdot m \cdot n$. The rescaled payoffs are now,

$$\pi_C = 0$$
$$\pi_D = c - sp\tfrac{x}{z}$$
$$\pi_R = \begin{cases} -p & if \quad z > 0 \\ 0 & otherwise \end{cases}$$

Clearly All D is still a nash. Additionally anywhere on the the R-C segment where $z = 0$ is nash. Again we can write the replicator equations

$$\bar{\pi} = y \cdot 0 + x \cdot (-p) + z(c - sp\tfrac{x}{z})$$
$$x + y + z = 1$$
$$\dot{x} = x(-p + xp - zc + spx)$$
$$\dot{y} = y(0 + xp - zc + spx)$$
$$\dot{z} = z(c - sp\tfrac{x}{z} + xp - zc + spx)$$

We let $p = 1, c = s = 2$ and get.

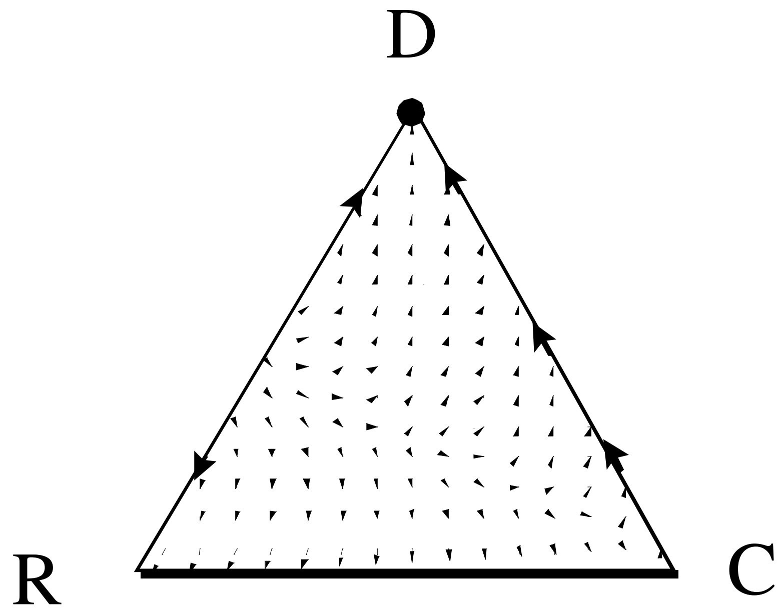

FIGURE 1. **Punishment 1**

$$\dot{x} = x(-1 + x - z2 + 2x) = x(3x - 2z - 1)$$
$$\dot{y} = y(0 + x - z2 + 2x) = y(3x - 2z)$$
$$\dot{z} = z(2 - 2\frac{x}{z} + x - z2 + 2x) = z(3x + 2 - 2\frac{x}{z} - 2z)$$

There is no interior $(x, y, z > 0)$ rest point since if $z > 0$ then $\pi_C > \pi_R$. And when$x = \frac{cz}{sp}$ we get $\pi_D = \pi_C > \pi_R$. While when $z = 0$ $\pi_C = \pi_R$ forms a set of rest points on the boundary where any mutation pushes the system towards $y = 1$, after which it goes to $z = 1$. This is identical to the analysis in (Sethi and Somanathan 1996, Bowles and Gintis 2004b). When $x = \frac{c}{c+sp}$ and $y = 0$ $\pi_C = \pi_D$. This defines the unstable equilibria on the boundary. Next we use Mathematica [3] to plot the Simplex and the vector field defined above to get Figure 1.

[3]and the SimplexDrawGraphics Package from Carl T. Bergstrom

3. Perturbed Replicator Dynamics

Given

$$\bar{\pi} = y \cdot 0 + x \cdot (-p) + z(c - sp\tfrac{x}{z})$$
$$x + y + z = 1$$

We follow Hopkins(2001) and write the perturbed replicator equations as

$$\dot{x} = x(-p + xp - zc + spx) + \tfrac{1-3x}{\lambda}$$
$$\dot{y} = y(0 + xp - zc + spx) + \tfrac{1-3y}{\lambda}$$
$$\dot{z} = z(c - sp\tfrac{x}{z} + xp - zc + spx) + \tfrac{1-3z}{\lambda}$$

where λ inversely controls the error/mutation rate.

Given mutation this system will eventually go to the all defect outcome, $z = 1$ as shown by Hauert, Haiden and Sigmund (2004).

3.1. **The Alternative Public Goods Game.** The standard treatment of public good provision with punishment posits a two stage interaction. Contributions to the public good are provided by the agents in the first stage and the public good is created. Punishment is inflicted by agents in the second. Two models (Bowles and Gintis 2004a, Sethi and Somanathan 1996) have a single stage game of sufficient length so that punishment and contributions(extraction of the resource) occur in a single period. In both these models however punishment doesn't change amount of the public good produced. In the modified game here commitments to punishment and contributions to the public good occur simultaneously and most importantly are funded out of the same pool of resources.

In this case we can modify the payoffs above as follows

$$\pi_{R'} = 0$$
$$\pi_{C'} = 0$$
$$\pi_{D'} = c - sp\tfrac{x}{z}$$

where

$$c = sp\tfrac{x}{z}$$
$$x^* = x = \tfrac{cz}{sp}$$

Now we have a unstable set of interior rest points where $x = x^*$. Again all D is still a Nash as is anywhere on the the R-C segment where $z = 0$. Also now the set of rest points on the boundary where z = 0 will not move towards y = 1 given positive mutation.

$$\dot{x} = x(0 - z2 + 2x) = x(2x - 2z)$$
$$\dot{y} = y(0 - z2 + 2x) = y(2x - 2z)$$
$$\dot{z} = z(2 - 2\tfrac{x}{z} - z2 + 2x) = z(2x + 2 - 2\tfrac{x}{z} - 2z)$$

Assuming again that $s = c = 2, p = 1$ we can plot this vector field as is done in Figure 2. Next we add Mutation to get the perturbed

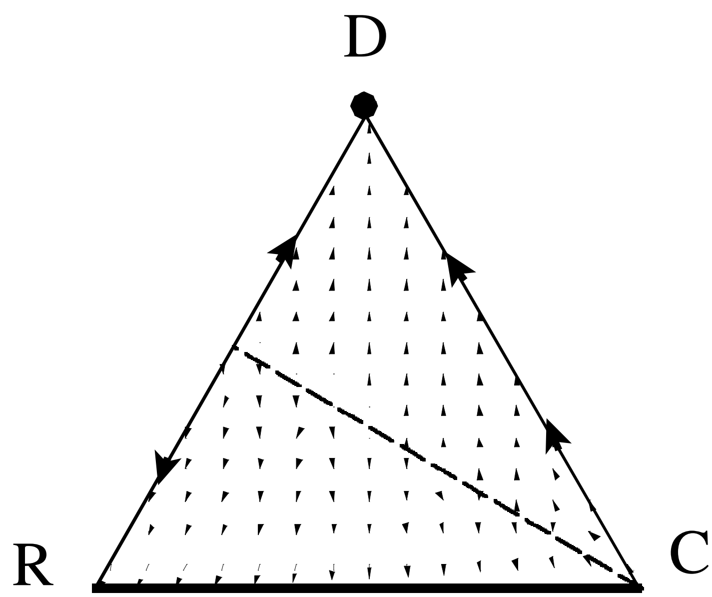

FIGURE 2. **Punishment 2**

replicator equations following . This results in the R-C segment being unstable for the standard model while the same segment is stable under the modified version presented here.

4. SIMULATION

At the beginning of each round agents are randomly placed in groups. The agents then play the following game. Each agent is given *2* dollars either of which he can place in one of three accounts, only whole dollars may be deposited. Given (e,f,g) where e is the contribution to the public account, f is the amount kept in the private account and g is the contribution to the sanction account. The 6 possible types are

- Reciprocator - S1 -(1,0,1)
- Spiteful - S2 -(0,0,2)
- Altruist - S3 - (2,0,0)
- Masochist - S4 - (0,1,1)
- Sneaky - S5 - (1,1,0)
- Selfish - S6 - (0,2,0)

Earnings are updated every period as follows: Each private account is directly added to that individual agents earnings . The public account proceeds are totaled for each group and then get multiplied by a fixed factor and shared equally by that group. Sanctioning also takes place at the group level. The Sanction account proceeds for each group get multiplied by a fixed factor then sanctioning takes place as follows: one agent per round from that group may be sanctioned by subtracting the total sanction amount from the earnings of the highest numbered type in the group except for type S1,S2, and S3[4]. If there are two of the highest type number in a group one will randomly get the full sanction. If the whole population is of type S1,S2, or S3 no agents are punished and the funds vanish. Agents carry no memory of past actions or outcomes with them. Therefore in this simple version agents cannot condition their behavior on either the perceived distribution of types or other individuals pasted play. The game is repeated for R periods. After R periods (the end of a generation) agents reproduce/imitate according to Schlags' imitation method. Then the next generation

[4]An Alternative punishment norm includes punishing S2 and S3.

begins and earnings for all agents are set to zero. The simulation continues for G (number of generations) generations.

The punishment method here randomly picks an agent from among those in each group with the highest type (S4...S6) and subtract from that agents earnings the sum of all sanctions in that group times the strick size. This is not the standard treatment. For instance Bowles and Gintis (2004) have each reciprocator randomly pick one agent to monitor and potentially punish in each round while Sethi and Somanathan (1996) have each reciprocator punish all the defectors in the group. These punishment methods are much less likely than the one used here to result in a few defectors with very high scores while most have very low. The problem is in large groups if only one is getting punished many more will be getting away without punishment. The result is that if the group size is much bigger than the number of rounds in a generation then there could be a small few defectors who make it through without being punished and in so doing earn above the average level.

In the simulation results presented below we use the SPOR imitation rule from Schlag (1998) which sequentially and proportionally imitates the strategy with the higher payoff in pairwise comparisons through a given sample. When all players use the SPOR rule, actions yielding above average payoffs increase in frequency and those with below average payoffs decrease. The increase in play of a strategy in this case is a function of the product of its frequency and of the difference between it's own and the average expected payoff in the population. The resulting learning adjustment path is a discrete-time version of the replicator dynamic. Increasing the sampling size in the SPOR method speeds up the dynamic. As the sample size goes to infinity the learning dynamic turns into the adjusted replicator dynamic of Maynard Smith (1982).

The simulation results confirm the analysis above.

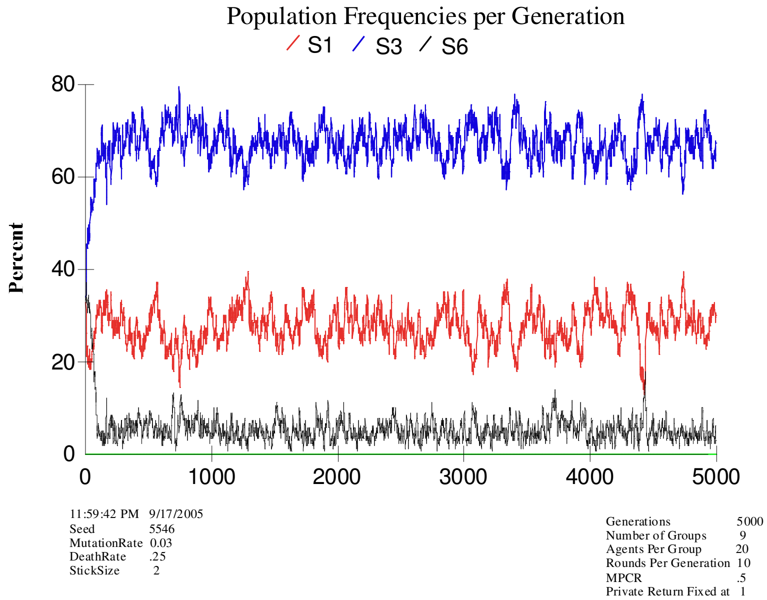

FIGURE 3. **Simulation 1**

REFERENCES

Alexander, R.D., *The Biology of Moral Systems*, New York: Aldine de Gruyter, 1987.

Axelrod, R and W.D.Hamilton, "The Evolution of Cooperation," *Science*, 1981, *211*, 1390–1396.

Bowles, Samuel and Herbert Gintis, "The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations," *Theoretical Population Biology*, 2004, *65*, 17–28.

______ and ______, "Persistent Parochialism: Trust and Exclusion in Ethnic Networks," *Journal of Economic Behavior and Organization*, 2004, *55*, 1–23.

Boyd, Robert, Herbert Gintis, Samuel Bowles, and Peter J. Richerson, "The Evolution of Altruistic Punishment," *PNAS*, 2003, *100*, 3531–3535.

Cinyabuguma, Matthias, Talbot Page, and Louis Putterman, "Cooperation Under the Threat of Expulsion in a Public Goods Experiment," 2004.

______, ______, and ______, "On Perverse and Second-Order Punishment in Public Goods Experiments with Decentralized Sanctioning," 2004.

Gintis, Herbert, *Game Theory Evolving*, Princeton, New Jersey: Princeton University Press, 2000a.

———, **Eric Alden Smith, and Samuel Bowles**, "Costly Signaling and Cooperation," *Journal of Theoretical Biology*, 2001, *213*, 103–119.

Hauert, Christoph, Nina Haiden, and Karl Sigmund, "The Dynamics of Public Goods," *DISCRETE AND CONTINUOUS DYNAMICAL SYSTEMS SERIES B*, 2004, *4*, 575–587.

Hopkins, Ed, "Two Competing Models of How People Learn in Games," *NAJ Economics Peer Reviews of Economics Publications*, 2001, *1.*

Leimar, Olof and Peter Hammerstein, "Evolution of Cooperation Through Indirect Reciprocity," *Proc. R. Soc. Lond. B*, 2001, *268*, 745–753.

Maynard Smith, John, *Evolution and the Theory of Games*, Cambridge, UK:: Cambridge University Press, 1982.

Nowak, M.A. and K. Sigmund, "Evolution of Indirect Reciprocity by Image Scoring," *Nature*, 1998, *393*, 573–577.

Panchanathan, Karthik and Robert Boyd, "A Tale of Two Defectors: The Importance of Standing for Evolution of Indirect Reciprocity," *Journal of Theoretical Biology*, 2003, *224*, 115–126.

Schlag, Karl, "Justifying Imitation," 1998. Working Paper.

Sethi, Rajiv and E. Somanathan, "The Evolution of Social Norms in Common Property Resource Use," *American Economic Review*, 1996, *86*, 766–88.

Sugden, R., *The Economics of Rights*, Oxford: Basil Blackwell, 1986.

Trivers, Robert L., "The Evolution of Reciprocal Altruism," *The Quarterly Review of Biology*, 1971, *46*, 35–57.