



www.sciencemag.org/cgi/content/full/328/5978/617/DC1

Supporting Online Material for

Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare

Robert Boyd,* Herbert Gintis,* Samuel Bowles*

To whom correspondence should be addressed. E-mail: rboyd@anthro.ucla.edu (R.B.);
hgintis@comcast.net (H.G.); samuel.bowles@gmail.com (S.B.)

Published 30 April 2010, *Science* **328**, 617 (2010)

DOI: 10.1126/science.1183665

This PDF file includes:

Materials and Methods

Figs. S1 to S7

References

Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare: Supplementary Online Materials

Robert Boyd,^{1*} Herbert Gintis,² Samuel Bowles³

¹Department of Anthropology, UCLA

²Santa Fe Institute and Central European University

³Behavioral Sciences Program, Santa Fe Institute and Department of Economics, University of Siena

*To whom correspondence should be addressed; E-mail: rboyd@ucla.edu.

Table of Contents

Model description	p. 2
Model Analysis: Punishers vs Nonpunishers	p. 2
Liars and the minimum cost of signalling	p. 4
conditional cooperators and Second-order free riding	p. 5
Figure S1: Effect of Varying the Gains to Cooperative Punishment	p. 7
Linkage and the Evolution of Punishment	p. 7
Evolution of τ	p. 7
Numerical Analysis	p. 8
Figure S2: Effect of Varying Economies of Scale in Punishment a	p. 11
Figure S3: Effect of Varying the Cost of Punishment k	p. 11
Figure S4: Effect of Varying the Error Rate e	p. 12
Figure S5 Effect of Varying the Group Size n	p. 13
Figure S6: Effect of Varying the Expected Number of Periods T	p. 13
Figure S7: Effect of Varying the Amount of Assortment r	p. 14
References	p. 16

Model Description Suppose we have a large population of agents of several types in which individuals are sampled with assortment. For any individual in a given group, the probability that a randomly chosen other member of the group is of the same type is $r + (1 - r)x$ where x is the frequency of the type in the whole population. Note that in a genetic version of the model r is equivalent to the average coefficient of relatedness among group members. The probability that a randomly chosen other member is of a different type is then $(1 - r)(1 - x)$. When $r = 0$, this model represents random interaction. When $r > 0$, individuals of the same type are more likely to find themselves together in group than chance alone would dictate.

Individuals interact in a repeated game in which the first round has one signaling, one cooperation, and one punishment stage, and succeeding periods have only single cooperation and punishment stages. All individuals have a baseline fitness W_0 . The “costs” and “benefits” are the incremental effects of behavior on individual fitness. The interaction continues for an expected number T periods, which means that at the end of each period, the probability of dissolving the group is $w = 1 - 1/T$.

We define four strategies that specify behavior. (Note that cooperation is not a strategy it is an elective action by each of the four strategies in response to the expectation of punishment for defection.)

1. **Punishers:** Punishers signal in the signaling stage. We say the quorum is met if the number of signalers is at least $\tau + 1$ ($1 \leq \tau \leq n - 1$). In each cooperation stage, Punishers cooperate if the quorum is met and defect otherwise. In each punishment stage, Punishers collectively punish all agents who defected in the current cooperation stage if the quorum is met. If the quorum is not met, Punishers do not punish defectors. For any period after the first one, the quorum is met if (a) the quorum was met in the previous period and there were no defections in the current cooperation stage, or (b) there was at least one defection in the current cooperation stage and more than τ agents participated in punishing defectors.
2. **Nonpunishers:** In the signaling stage, Nonpunishers do not signal. In the first cooperation stage, Nonpunishers defect. In subsequent periods, Nonpunishers cooperate if the quorum is met. Nonpunishers never punish.
3. **Liars:** Liars signal in the signaling stage. In each cooperation stage, Liars cooperate if and only if the quorum is met. Liars never punish.
4. **Conditional Cooperators:** Conditional cooperators, like all other types, cooperate if they believe that that they will be punished otherwise. Unlike Nonpunishers Conditional Cooperators believe the signal during the first period and therefore cooperate if a quorum is met. Conditional Cooperators do not signal and never punish.

Model Analysis: Punishers vs Nonpunishers Suppose the population consists only of Punishers and Nonpunishers. Let x be the frequency of Punishers in the population, and fix a particular group. For a given focal group member, let j be the number of Punishers among the other $n - 1$ group members. When the focal member is a Punisher, j is distributed binomially with parameters $\Pr(j|P) = r + (1 - r)x$ and $n - 1$. When the focal member is a Nonpunisher, j is distributed binomially with parameters $\Pr(j|N) = (1 - r)x$ and $n - 1$. The expected fitness of a Punisher, W_P , is composed of three terms in addition to baseline fitness: the cost of signaling in the first period, the expected benefits of cooperation net of the costs of cooperating and punishing others in the first period in cases where the quorum is met, and the benefits of cooperation net of the costs of punishing others and being punished in the case of errors summed over the expected duration of the interaction following the first period (which is $T - 1$) in cases where the quorum is met. This is given by

$$\begin{aligned}
W_P(x) = & W_0 - q \\
& + \sum_{j=\tau}^{n-1} \Pr(j|P) \left(-\frac{(e(j+1) + (n-1-j))k}{(j+1)^a} + (1-e) \left(\frac{b(j+1)}{n} - c \right) \right) \\
& + (T-1) \sum_{j=\tau}^{n-1} \Pr(j|P) \left((b-c)(1-e) - \frac{nek}{(j+1)^a} - ep \right) \tag{S1}
\end{aligned}$$

Nonpunishers are punished in the first period, receive a share of the public good produced by Punishers, and then experience the long term benefits of cooperation only if the quorum is met. Thus if $\tau < n - 1$, $W_N(x)$ is given by

$$\begin{aligned}
W_N(x) = & W_0 + \sum_{j=\tau+1}^{n-1} \Pr(j|N) \left(-p + \frac{(1-e)jb}{n} \right) \\
& + (T-1) \sum_{j=\tau+1}^{n-1} \Pr(j|N) ((b-c)(1-e) - ep) \tag{S2}
\end{aligned}$$

If $\tau = n - 1$ so that Punishers only punish if all other members of their group are Punishers, then $W_N(x) = W_0$ because Non-punishers do not find themselves in a group in which either punishment or cooperation occurs.

Individual reproductive success is proportional an individual's payoff. This means that the change in frequency of Punishers who use strategy P_τ is given by:

$$\Delta x = x(1-x) \frac{W_P(x) - W_N(x)}{xW_P(x) + (1-x)W_N(x)} \tag{S3}$$

Thus, equilibria occur at $x = 0$, $x = 1$, and values of x is such that $W_P(x) = W_N(x)$.

The fitness difference is then

$$\begin{aligned}
\Delta(x) &= W_P(x) - W_N(x) \\
&= (T-1)((b-c)(1-e) - ep) \left(\sum_{j=\tau}^{n-1} \Pr(j|P) - \sum_{j=\tau+1}^{n-1} \Pr(j|N) \right) \\
&\quad - (T-1) \sum_{j=\tau}^{n-1} \frac{ekn}{(j+1)^a} \Pr(j|P) \\
&\quad + p \sum_{\tau+1}^{n-1} \Pr(j|N) - q \\
&\quad + \frac{b(1-e)}{n} \left(\sum_{j=\tau}^{n-1} \Pr(j|P) - \sum_{j=\tau+1}^{n-1} \Pr(j|P) \right) - \left(c - \frac{b}{n} \right) (1-e) \sum_{j=\tau}^{n-1} \Pr(j|P) \\
&\quad - \sum_{j=\tau}^{n-1} \frac{k(n-1-j+e(j+1))}{(j+1)^a} \Pr(j|P). \tag{S4}
\end{aligned}$$

The first two terms determine the long-run behavior of the model when T is large so that groups on average last for many periods. The first term is the expected long-run net gain to being a Punisher. This term is never negative, is small for small x unless r is large. For sufficiently large r , this term is close to the payoff to full cooperation, $(T-1)((b-c)(1-e) - ep)$. The second term is the long-run cost of punishing, This term is small for low error rates, but for error rates of the magnitude $1/n$ this term can be substantial.

Suppose that groups are formed at random. Then the long-term net gain from being a Punisher is given by becomes

$$(T-1) \left(((b-c)(1-e) - ep) \Pr(\tau|P) - \sum_{j=\tau}^{n-1} \frac{ekn}{(j+1)^a} \Pr(j|P) \right) \tag{S5}$$

where $\Pr(j|P) = \Pr(j|N)$. In a population of Nonpunishers (where x is near zero), if $\tau > 0$ then there will almost never be a quorum, but Punishers must pay q to signal, so Punishers cannot invade. If $\tau = 0$, so a lone Punisher disciplines the rest of the group, we have

$$\Delta(x) = (T-1)((b-c)(1-e) - ep) - k(n-1) + p - q, \tag{S6}$$

so a lone Punisher can invade if p is sufficiently large and k sufficiently small. We consider this an implausible situation, and henceforth will assume $\tau > 0$.

Liars and the minimum cost of signalling Next consider the fitness of rare Liars. Liars signal so incur the cost q . They are not punished during the first period. In subsequent periods

they act just like a Nonpunisher and so get the payoff of Nonpunishers in later periods. When groups are formed at random, rare Liars will be in groups without another Liar. Thus, the expected payoff of a rare liar when groups are formed at random is:

$$W_L(x) = W_0 - q + (T - 1) \sum_{j=\tau+1}^{n-1} \Pr(j|N) ((b - c)(1 - e) - ep) \quad (\text{S7})$$

At equilibrium $W_N = W_P$, and thus liars can invade this equilibrium unless

$$q > p \sum_{j=\tau+1}^{n-1} \Pr(j|N) \quad (\text{S8})$$

That is, the cost of signalling must be greater than the expected first-period cost of being punished. Note that this depends both on the frequency of punishers and the punishment threshold. However, as we will see below, the expected cost of being punished during the first period is often only slightly less than p , especially when there are increasing returns to punishment. Thus as a first approximation, p provides a lower bound on q . Punishers can always deter Liars if the cost of signalling is the same as the cost of being punished. Throughout we assume that $q = p$.

Contingent cooperators and Second-order free riding Contingent Cooperators receive the same payoff as Nonpunishers in groups where the quorum is not met, and do strictly better in groups where the quorum is met. Thus Contingent Cooperators can invade a Punisher-Nonpunisher equilibrium. As illustrated in figure S1, Contingent Cooperators replace nonpunishers without much changing the equilibrium frequency of Punishers. This effect will be even greater for parameter values for which the equilibrium fraction of Punishers is considerable, or where groups are very long-lived because Contingent Cooperators in cooperative groups have a payoff advantage only in the first period, and the size of this advantage is on the order of magnitude c , the cost of cooperating.

A world of all Nonpunishers is a plausible ancestral state. Nonpunishers do not cooperate, nor do they respond to unproven threats. Once they are punished, however, they cooperate to avoid further punishment. We have seen that rare Punishers can invade such a population as long as there is some assortment, interactions go on long enough, and there are economies of scale in punishment. When Punishers are common, however, it selection will favor Contingent Cooperators who take the signal seriously, and hence cooperate in the first period when the quorum is met. Contingent Cooperators “invade” a Punisher-Nonpunisher equilibrium but displace the nonpunishers rather than the Punishers. Contingent Cooperators are the optimal second-order free rider. They only cooperate when they will be punished, and never punish themselves so the stable equilibrium mix of Punishers and Contingent Cooperators cannot be invaded by any other cooperative strategies.

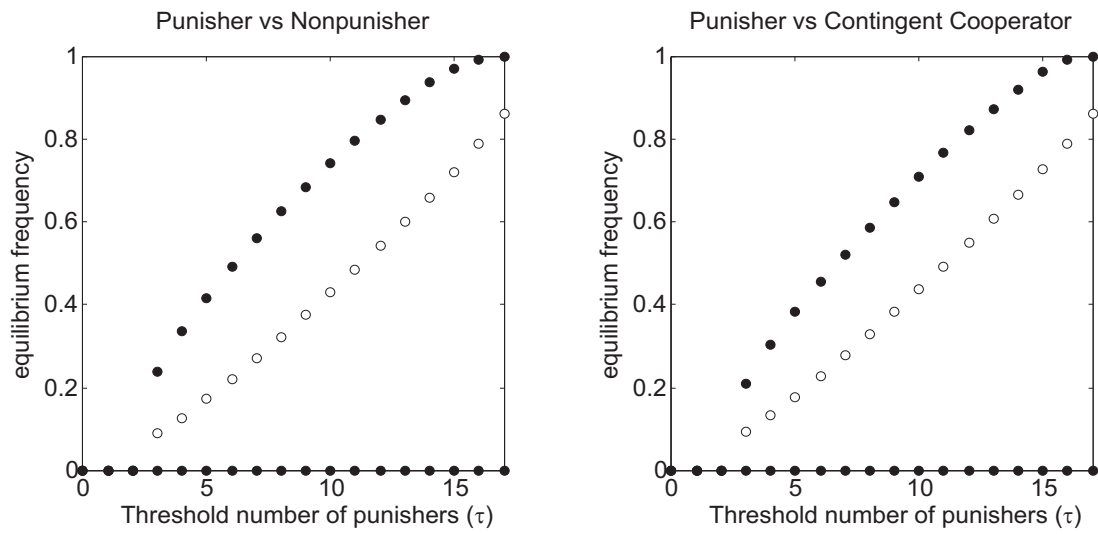


Figure S1: The equilibrium frequencies of Punishers as a function of τ when Punishers compete with Nonpunishers and when Punishers compete with Contingent cooperators. Solid circles represent the frequency of Punishers at stable equilibria, and open circles represent the frequency of Punishers at unstable equilibria. Notice that the equilibrium values are very similar. Base line parameter values: $b = 2c$, $a = 2$, $k = p = q = 1.5c$, $T = 25$, $e = 0.1$, $n = 18$, and $r = 0.0$.

Linkage and the Evolution of Punishment It has been argued that punishment can evolve to support cooperation only when the two traits are genetically linked *l*. Linkage allows persistent genetic correlation between the two traits, and this means that the selective benefits of cooperation can increase the frequency of punishment. In the model presented above there is some linkage between Punishment and Cooperation because Punishers cooperate in groups above the threshold while Nonpunishers do not cooperate. However, similar results obtain in a model in which there is no linkage at all between cooperation and punishment. In this model, the ordering of behaviors during the first stage is different. First, there is the cooperation stage, then the signalling stage, and finally the punishment stage. Both Punishers and Nonpunishers defect during the first period cooperation stage. Punishers then signal and if the number of signallers is greater than $\tau + 1$ punish those who do not signal during the first stage. Noncooperators also defect, do not signal and do not punish. During subsequent periods both types cooperate if punishment has occurred in the past. Thus, there is no linkage at all between cooperation and punishment. Both types cooperate only when punishment has occurred. The equilibrium behavior of this model is very similar to the behavior of the model presented in this paper and slightly more favorable to the evolution of punishment.

The Evolution of τ In this section we consider the evolution of τ by asking whether punishing strategies with larger or smaller values of τ can invade a mixed equilibrium with P_τ and N .

First, we show that such an equilibrium cannot be invaded by rare individuals whose strategy is characterized by a larger value of τ . As before, define P_τ as the punisher strategy with a threshold value of τ where $\tau = 0, 1, \dots, n - 1$. We assume that P_τ and N have reached a stable internal equilibrium value and derive the conditions under which $P_{\tau+1}$ can invade. Since this is unlikely to be possible unless there is some assortment we need to allow for the possibility that $r > 0$. This means that in calculating the fitness of invaders we cannot assume that invaders are by themselves but instead have to allow for the possibility that rare invaders will be in groups with some of their own type. Consider a focal $P_{\tau+1}$ invader when this strategy is rare. The probability that a given individual among the other $n - 1$ is P_τ is $(1 - r)\hat{x}$ where \hat{x} is the equilibrium frequency of P_τ in competition with N . Similarly, the probability that another individual is N is $(1 - r)(1 - \hat{x})$, and the probability that another individual $P_{\tau+1}$ is r . Let the number of P_τ individuals among the other $n - 1$ equal i , the number of $P_{\tau+1}$ individuals among the other $n - 1$ equal j , and the number of N individuals equal l . These are distributed as a multinomial with parameters $(1 - r)x$, r , $(1 - r)(1 - x)$ and $n - 1$.

Now compute the expected fitness of a focal $P_{\tau+1}$ when this type is rare. Both punishing types signal. Thus each punisher perceives $i + j$ signals from others. If $i + j < \tau$, or $l > n - 1 - \tau$, neither type punishes or cooperates. If $i + j \geq \tau + 1$, or equivalently $l \leq n - 1 - (\tau + 1)$, the number of perceived signals is greater than $\tau + 1$ so both types punish on the first interaction and cooperate and punish on subsequent interaction. If $i + j = \tau$, or equivalently $l = n - 1 - \tau$,

the none of the $j + 1$ $P_{\tau+1}$ individuals cooperates or punishes during the first period. The i P_{τ} individuals punish defectors, including the $P_{\tau+1}$ individuals, during the first period. During subsequent periods neither type cooperates or punishes. Thus the expected fitness of a rare $P_{\tau+1}$ individuals is:

$$W_{\tau+1}(x) = W_0 - q - p(1 - r^{tau}) \Pr(l = n - 1 - \tau | P_{\tau+1}) - \sum_{l=0}^{n-2-\tau} \Pr(l | P_{\tau+1}) \frac{lk}{(n-l)^a} \\ + T \sum_{l=0}^{n-2-\tau} \Pr(l | P_{\tau+1}) \left((b-c)(1-e) - \frac{nek}{(n-l)^a} - ep \right)$$

Now compare this to the value of W_{τ} now expressed as a function of l ,

$$W_{\tau}(x) = W_0 - q - \sum_{l=0}^{n-1-\tau} \Pr(l | P_{\tau}) \frac{lk}{(n-l)^a} \\ + T \sum_{l=0}^{n-1-\tau} \Pr(l | P_{\tau}) \left((b-c)(1-e) - \frac{nek}{(n-l+1)^a} - ep \right)$$

Now, note that $\Pr(l | P_{\tau+1}) = \Pr(l | P_{\tau})$ since the probability that each individual among the other $n - 1$ individuals in both cases is $(1 - r)(1 - x)$. Thus

$$W_{\tau}(x) - W_{\tau+1}(x) = \Pr(l = n - 1 - \tau) \left(-p - \frac{(n - 1 - \tau)k}{(\tau + 1)^a} \right. \\ \left. + T \left((b-c)(1-e) - \frac{nek}{(\tau + 1)^a} - ep \right) \right)$$

That is, $W_{\tau+1}$ individuals avoid the costs of punishing and give up the long run benefit of cooperation in groups in which there are exactly τ punishers. In addition they suffer the cost of being punished. But this is exactly the only group size in which W_{τ} individuals have a payoff advantage relative to N . So the benefits of cooperation in this group size have to exceed the costs of punishing in this group size. Individuals with higher thresholds cannot invade.

Next, we show that a stable polymorphic equilibrium P_{τ} and N can be invaded by rare individuals whose strategy is characterized by a smaller value of τ only very restrictive conditions. As before, we assume that P_{τ} and N have reached a stable internal equilibrium value and derive the conditions under which $P_{\tau-1}$ can invade. Consider a focal $P_{\tau-1}$ invader when this strategy is rare. The probability that a given individual among the other $n - 1$ is P_{τ} is $(1 - r)\hat{x}$ where \hat{x} is the equilibrium frequency of P_{τ} in competition with N . Similarly, the probability that another individual is N is $(1 - r)(1 - \hat{x})$, and the probability that another individual $P_{\tau-1}$ is r . Let the number of P_{τ} individuals among the other $n - 1$ equal i , the number of $P_{\tau-1}$ individuals among the other $n - 1$ equal j , and the number of N individuals equal l . These are distributed as a multinomial with parameters $(1 - r)x$, r , $(1 - r)(1 - x)$ and $n - 1$.

Now consider the focal fitness of an $P_{\tau-1}$ individual. Both punishing types signal. Thus each punisher perceives $i + j$ signals from others. If $i + j < \tau - 1$, or $l > n - 1 - \tau$, neither type punishes or cooperates. If $i + j \geq \tau$, or equivalently $l \leq n - 1 - \tau$, the number of perceived signals is greater than τ so both types punish on the first interaction and cooperate and punish on subsequent interaction. If $i + j = \tau - 1$, or equivalently $l = n - 1 - (\tau - 1)$, the none of the $i P_\tau$ individuals punish during the first period and only the $j P_{\tau-1}$ individuals punish defectors during the first period. Thus, unless $j = n - 1$ there is no punishment during subsequent periods. Thus the expected fitness of a rare $P_{\tau-1}$ individuals is:

$$\begin{aligned}
W_{\tau-1}(x) = & W_0 - q - \sum_{j=0}^{\tau-1} \Pr(j, l = n - j | P_{\tau-1}) \frac{(n - 1 - (\tau - 1))k}{(j + 1)^a} \\
& - \sum_{l=0}^{n-1-\tau} \Pr(l | P_{\tau-1}) \frac{lk}{(n - l)^a} \\
& + T \sum_{l=0}^{n-1-\tau} \Pr(l | P_{\tau-1}) \left((b - c)(1 - e) - \frac{nek}{(n - l)^a} - ep \right) \\
& + Tr^{\tau-1} \left((b - c)(1 - e) - \frac{nek}{(n - l)^a} - ep \right)
\end{aligned}$$

Thus

$$\begin{aligned}
W_{\tau-1}(x) - W_\tau(x) = & - \sum_{j=0}^{\tau-1} \Pr(j, l = n - j | P_{\tau-1}) \frac{(n - j)k}{(j + 1)^a} \\
& + Tr^{\tau-1} \left((b - c)(1 - e) - \frac{nek}{(n - l)^a} - ep \right)
\end{aligned}$$

That is the long run advantage in groups in which all of the punishers have a threshold of $\tau - 1$ has to be larger than the uncompensated cost of punishment in all groups in which there is a mix of the common type punishers and the invading type. Since the advantage declines geometrically with the small fraction r this condition is quite restrictive, and is not satisfied for any of the combinations of parameter values given below.

Numerical Analysis We have not been able to use the fitness functions presented above to derive expressions for the equilibrium frequencies of Punishers and Non-punishers. However, it is easy to compute these numerically, and in this section we present a range of results obtained in this way.

We chose the parameter values both to reflect what is known about the demographic and other conditions of Late Pleistocene humans and to adequately represent the costs of the punishment strategy so as not to bias the results in favor of its evolution.

- **Cost of cooperation (c):** This sets the units of the payoff parameters used in the simulation, so all further parameters will be in units of c . We chose $c = 0.01$, with a baseline fitness of $W_0 = 1$ so that selection coefficients are on the order of a few percent. Selection against a deleterious allele imposing a cost of 0.01 on its bearer would reduce its presence in a population from 90 percent to 10 percent in fewer than 500 generations
- **Per capita benefit of full cooperation (b):** We consider two values, weakly beneficial, $b = 2c$, which would only be favored by kin selection among full sibs, and $b = 4c$, so that altruism would be favored by kin selection among half sibs. Group benefits from other kinds of cooperation, in predation and defense or sharing information of danger or food sources, for example, could substantially exceed 4. The likely frequency of both environmental shocks and intergroup conflict among Pleistocene foragers (2,3), coupled with the importance of cooperation in surviving those challenges suggests that we have not overestimated the range of likely benefits.
- **Cost of being punished (p):** This value must be greater than $c - b/n$ in order to motivate individuals to cooperate. We assume that $p = 1.5c$. We thus assume “surplus” punishment, accompanied by costs to the punisher in excess of that strictly needed to induce the target to cooperate.
- **Cost of punishing to a single punisher (k):** Many people think that punishment is less costly to the punisher than to the target either because the punisher has the advantage of initiating the conflict or because the target is often outnumbered. Our baseline assumption is that this is not the case. We simulate over three values: $k = p/2$, $k = p$, and $k = 2p$.
- **Cost of signalling (q):** The theory above suggests that the cost of signalling must be greater than the expected cost of being punished. This is guaranteed if $q = p$ as was shown above.
- **Effect of number of punishers on the cost of punishing (a):** We use a probabilistic version of Lanchesters Law on the relationship between numbers on opposing sides of a conflict and success in contests as the basis for our expected cost of punishment for those engaging in a punishment episode against a target (4–6). The expected cost of engaging in a punishing party depends on the likely outcome of the encounter with the target. We take account of this aspect of punishment in the following way. Suppose the cost of punishing, k , is an injury or other cost borne by a randomly selected member of the group of punishers if and only if the punishing episode is a standoff (neither the punishers nor the target “win” so that both target and punishers bear costs) and that a standoff occurs with probability $1/n$ where n is the number of punishers against a lone target. Thus, for instance, a single punisher against a single target always results in a standoff. The expected cost of joining a group of n punishers is thus the probability of a standoff ($1/n$) times the expected cost that a member of the punishing party will bear in the case of a standoff (k/n) or k/n^2 . Using k/n^a , we simulate two values: $a = 1$, which means constant

per capita cost of punishing, a base case that allows comparison with previous models, and $a = 2$, which means that the cost of punishing decreases linearly with the number of punishers. Our baseline $a = 2$ is quite unfavorable to the punishment strategy. In the example above, it implies that even in an encounter in which there are nine punishers and a single target, the target will achieve a standoff 10 percent of the time.

- **Error rate (e):** We interpret the error rate to be the frequency with which an individual does not cooperate even when punishment for defection is anticipated. This could occur by mistake, or because idiosyncratic circumstances in which the individual's cost of cooperating is exceptional (e.g., sick children at home that must be cared for) or another reason not included in the model for incurring the cost of being punished. We simulate three values $e = 0.01$, $e = 0.1$, and $e = 0.2$.
- **Group size (n):** The relevant size is the number of individuals of both sexes in a single breeding generation. The average census size of non-arctic, non-equestrian foraging groups in the ethnographic record that face conditions similar to must human groups in the Late Pleistocene is 37 (8), which implies an n of about 12 to 18, given that the fraction of a foraging population that is of reproductive age may be between a third and a half. To provide a strong test of our model, we simulate the upper end of this range and two other group sizes, $n = 18$, $n = 36$, and $n = 72$.
- **Number of interactions (T):** The appropriate number of interactions depends on the nature of the cooperative activity in question. Opportunities to cooperate in hunting and sharing prey occur thousands of times in a generation, while cooperation in predation or defense may occur but a few times per decade. We assume $T = 10$, $T = 25$, and $T = 50$.
- **Within-group relatedness (r):** We simulate four values: $r = 0$ (random group formation), $r = 0.035$, $r = 0.07$, and $r = 0.125$. The value $r = 0.07$ is the average relatedness in a sample of hunter gatherer groups (9).

These are varied one parameter at a time from the base case parameter set $\{b = 2c, p = k = q = 1.5c, a = 2, e = 0.1, n = 18, T = 25, r = 0.0\}$. For the most part we set the base case so that it is the worst case for cooperation, i.e., the benefit cost ratio is only 2, the cost of punishment is the same as the cost of being punished which is 50% higher than necessary to motivate individuals to cooperate, and 10% of the cooperators defect by mistake meaning that there are significant costs of punishing even in the long run.

Each figure gives the possible equilibria as a function of the τ value used by Punishers. Solid circles are stable equilibria, and open circles are unstable equilibria marking the boundaries between the basins of attraction of adjacent stable equilibria.

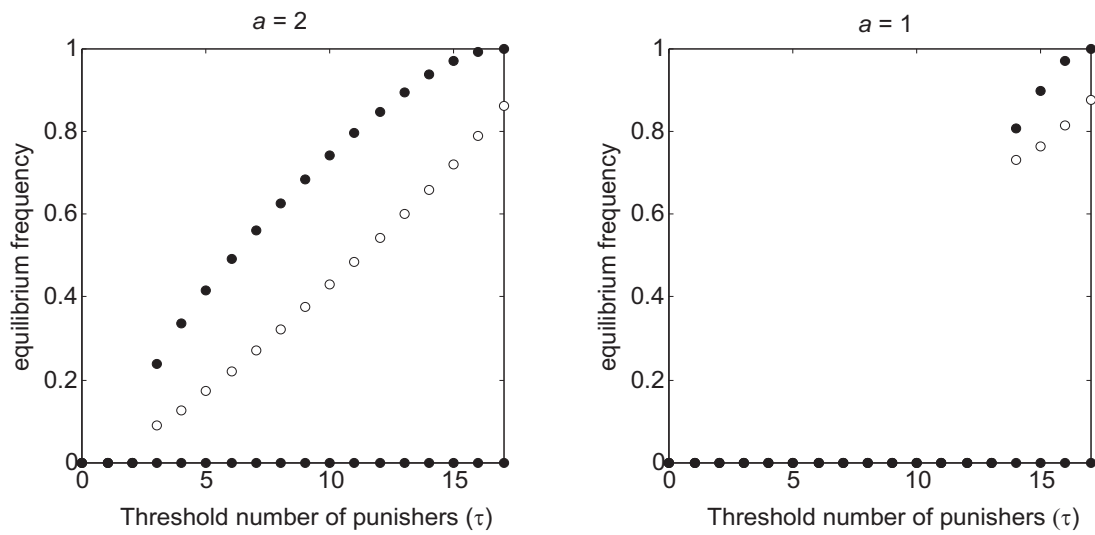


Figure S2: The scale economies in punishment (a) vary. When $a = 1$ there are no economies of scale in punishment. When $a = 2$ doubling the number of punishers halves the total cost of punishment. Without economies of scale, punishment is much more costly when τ is small, and thus Punishment can persist at equilibrium only for large value of τ .

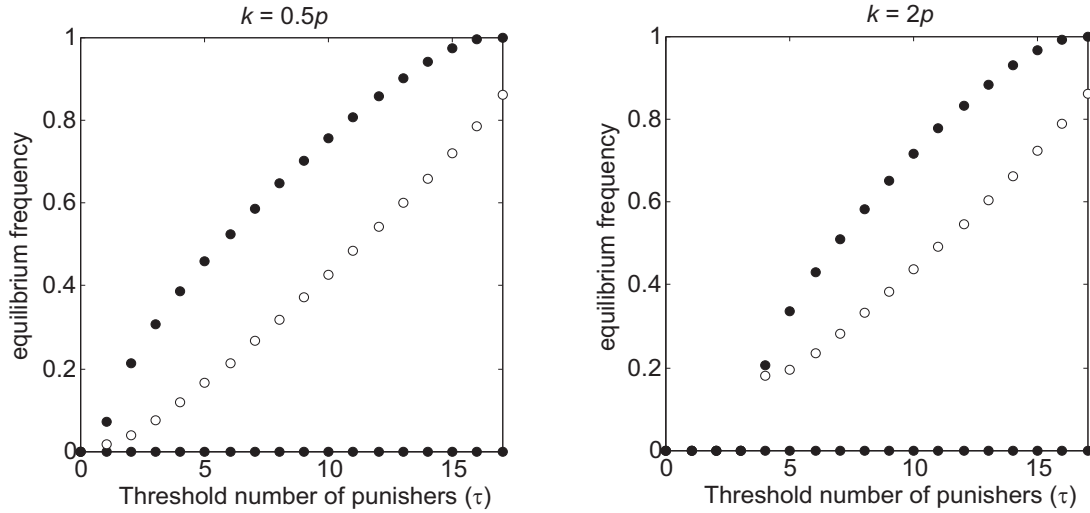


Figure S3: The cost of punishment (k) varies. Increasing the cost of punishment increases the cost of inducing Non-punishers to cooperate, and also increases the cost of punishing errors in the longer run. Thus increasing the cost of punishment reduces the range of conditions under which punishment can evolve.

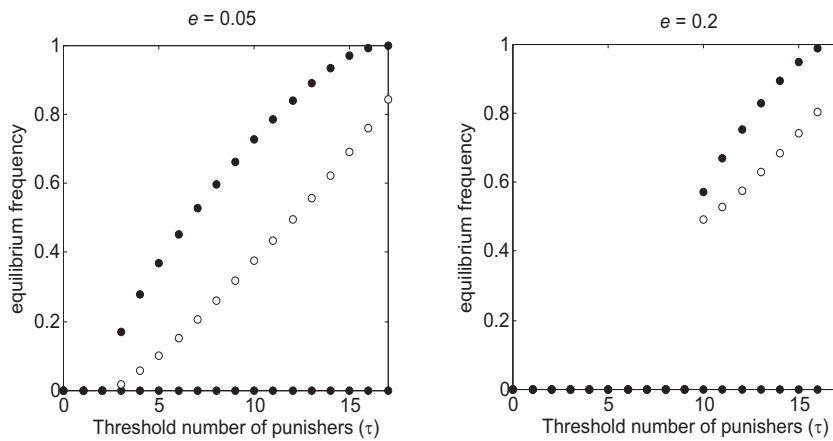


Figure S4: The error rate (e) varies. Higher error rates increase the frequency and thus the cost of punishment in the long run and thus decrease the benefit of cooperation induced by punishment. As a result, as error rates increase, cooperative equilibria exist only at higher values of τ .

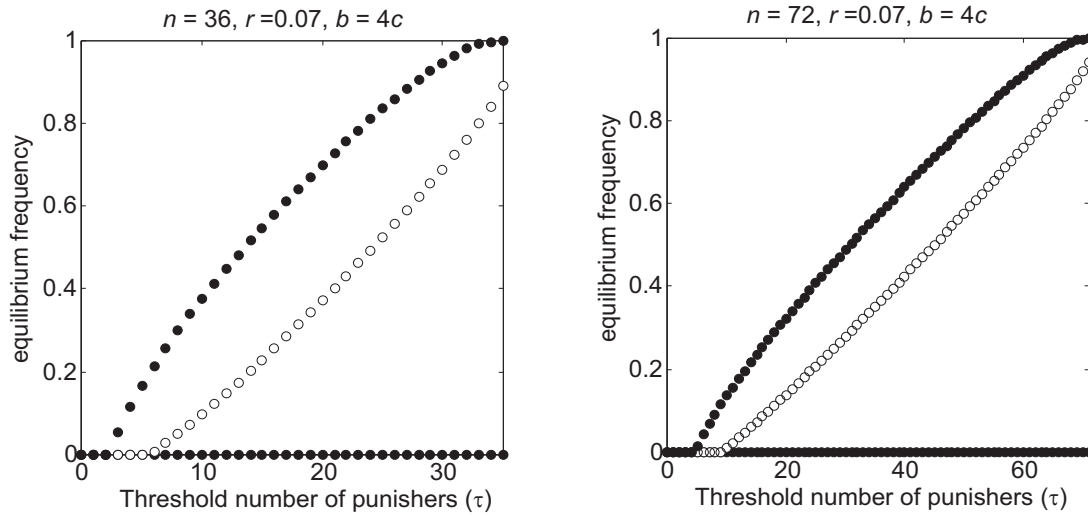


Figure S5: Group size (n) varies. Larger groups increase the cost of punishment of Non-punishers during the initial period and errors during later periods. Thus increasing n decreases the range of conditions which allow for punishment and cooperation at equilibrium.

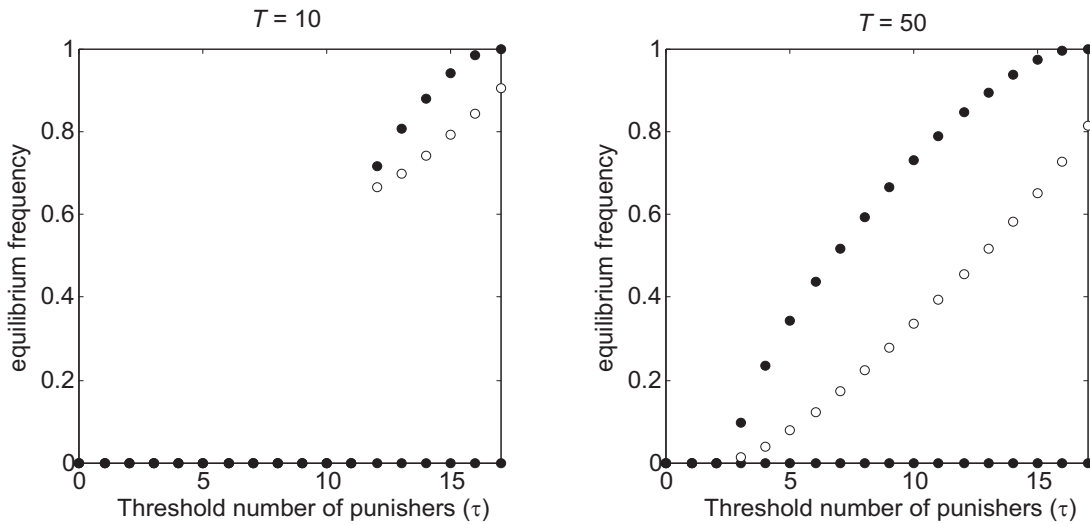


Figure S6: The expected number of periods (T) in an interaction varies. Increases in the expected number of interactions increase the long run benefits of cooperation and thus increase the range of conditions which allow for stable punishment and cooperation.

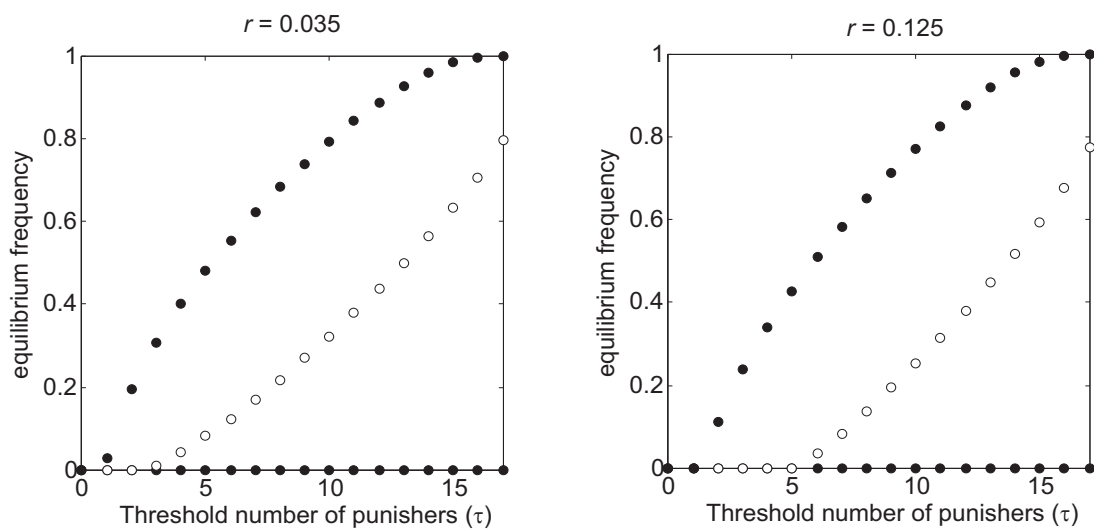


Figure S7: The relatedness (r) among group members vary. In the base case, groups are formed at random, and that means that Punishers are rare, they are alone in groups and can only induce cooperation if they are willing to act alone. Increasing r means that there is a positive probability that more than one punisher will be together in groups even when Punishers are rare. Thus for small values of τ , punishment can increase when rare even when r is assumed to be between 0.035 and 0.07 the latter being the mean value measured among contemporary foraging populations.

These results are surprisingly insensitive to variation in most of the parameter values. For most of the range of parameters simulated, there is a minimum value of τ . For values slightly greater than that value the minimum frequency necessary for Punishers to increase is low, and nearly maximum average fitness is achieved.

For the most part variation among results, can be understood as resulting from a tradeoff between the first-period costs of inducing Non-punishers to cooperate, and the long run benefit from the cooperation thereby induced. Note that the magnitude of p is determined by c —the cost of being punished only has to be sufficiently large to induce Non-punishers to cooperate. The cost of punishing is scaled relative to k . The other parameters fit into two categories. The parameters a , e , b , n and T affect the cost of punishing and the long run benefit. Their effect can be understood in terms of how they affect this balance. The effect of a is especially strong because it has a very large effect on how the cost of punishing varies as the number of Punishers in the groups increases. The parameter r reflects the population structure, and has large effects when Punishers are rare. Small positive values of r create a plausible set of conditions allows punishment to increase when rare.

References

1. L. Lehmann, F. Rousset, D. Roze, L. Keller. *The American Naturalist* 170, 21 (2007)
2. S. Bowles, Did Warfare among Ancestral Hunter-gatherers Affect the Evolution of Human Social Behaviors?, *Science: Supporting on line materials*, (2009)
3. North Greenland Ice Core Project Members, *Nature*, **431**, 147 (2004)
4. F.W. Lanchester, *Aircraft in warfare*, New York: Appleton, (1916)
5. J. H. Engel, *Operations Research*, **2**, 163 (1954)
6. S. Hwang, Large Groups may Alleviate Collective Action Problems, University of Massachusetts, (2009)
7. F. Marlowe, *Evolutionary Anthropology*, **14**, 54 (2005)
8. S. Bowles, *Science*, **314**, 1569 (2006)