

# Coordinated Contingent Punishment is Group-Beneficial and Can Proliferate When Rare

Robert Boyd<sup>1,2</sup>, Herbert Gintis<sup>2,4</sup>, and Samuel Bowles<sup>2,3</sup>

**In the past few years, the long-standing debate on the evolution of human cooperation has homed in on the central role of punishment of defectors; but the evolution of punishment has proven difficult to explain. Theoretical work indicates that while punishment of non-cooperators can persist once common, its initial emergence is precluded because rare punishers bear substantial costs and generate little cooperation. Moreover, experiments show that even when punishment is common, its cost often exceeds the resulting benefits of cooperation. In both models and experiments punishment has been implemented as an unconditional uncoordinated individual action. By contrast, in natural settings punishment of shirkers is typically contingent and coordinated. As a result, punishment is targeted on shirkers, deters counter-punishment, and is undertaken only when cost-effective, thus facilitating its proliferation when rare. Here we analyze a model of the evolution of contingent and coordinated punishment, showing that, under parameters approximating ancestral human conditions, punishment can proliferate when rare and enhances group-average payoffs when common.**

Humans are a uniquely cooperative species. In even the simplest societies, we cooperate in large groups of genealogically distant individuals (1–3). In the laboratory, subjects routinely cooperate in situations where selfish agents would defect (4,5). Recent theoretical studies provide an evolutionary explanation for such cooperative behavior (6–9). Punishment reduces the payoffs of defectors, and so groups with more punishers can sustain more cooperation. Punishment is costly, but unlike unconditional altruism, its costs are greatly reduced when punishers are common, because punishment then occurs at very low frequency. As a result, a modest advantage of groups in which cooperation is sustained by the presence of punishers is sufficient to stabilize a high level of cooperation.

There are two important problems with this explanation of human cooperation: First, punishment can be counter-productive; the costs of punishing may exceed the gains from cooperation (5,10,11). This problem is exacerbated when punishers target cooperative group members. This sometimes occurs in experiments (11) because the defector retaliates against those suspected of punishing him in the past. Second, the initial emergence of punishment remains a puzzle because in a world without punishment, a rare punisher must punish a large number of defectors (12–15).

We suggest that these problems are artifacts of the unrealistic way that punishment is implemented in existing models and experiments—as an unconditional, uncoordinated individual action automatically triggered by defection. By contrast, ethnographic evidence indicates that punishment is coordinated by communication among punishers, and is contingent on the

expected effectiveness of punishment in inducing cooperation (16-18). Solitary individuals rarely attempt to punish those who violate social norms. As a result, when it occurs, punishment is usually collective and is administered only when seen as legitimate by most group members. Consistent with the anthropological evidence, in behavioral experiments with communication or with a choice of a collective punishment strategy by subjects, punishment is often effective in raising group average payoffs (19).

Here, we analyze a model of the evolution of punishment that incorporates two empirically based features absent from previous work. First, punishment is characterized by increasing returns to scale, so the total cost of punishing declines as the number of punishers increases. Second, punishment is coordinated among many group members so it is contingent on the number of others predisposed to participate in the punishment. This means that when individuals willing to punish are rare, they demur and so bear only the cost of signaling their willingness to punish. They thus avoid the costs of punishing when it does not pay. Adding these two novel features resolves the problems with previous models. Our model shows that for levels of relatedness consistent with recent genetic data from hunter-gather populations (20), punishment can proliferate when rare, and when it is common it increases average group fitness.

Consider a large population in which individuals interact repeatedly in groups of size  $n$ . Groups are randomly formed so there is no genetic assortment. Later, we will introduce an empirically plausible degree of genetic assortment among groups. The model and its simulations are fully described in (21). The first period of an interaction has three stages. First is a cooperation stage, during which individuals can choose to cooperate or defect. Individuals who choose to cooperate mistakenly defect with probability  $e$ . Individuals who choose to defect always defect. Cooperation costs the cooperator  $c$  and benefits each member of the group  $b/n$ . There follows a signaling stage in which individuals can signal (at a cost  $q$ ) their intent to punish any defector. The cost of signaling is high enough to guarantee a cheater-proof way of communicating the intent to punish. Finally, there is a punishment stage in which individuals can cooperate with other punishers to administer punishment costing the target  $p$  at an expected cost  $k/n_p^a$  to each punisher, where  $n_p$  is the number of punishers. When  $a > 1$  there are increasing returns to scale so the total costs of a punishment episode decreases as the number of punishers increases. The interaction continues to another period with probability  $(1 - 1/T)$ , so  $T$  is the expected number of periods that the group persists. During subsequent periods, there are only cooperation and punishment stages.

Individuals have one of two heritable strategies, “Punisher” and “Nonpunisher.” During the first interaction, Punishers defect, and signal they are willing to punish. If at least  $\tau$  ( $0 \leq \tau \leq n - 1$ ) other group members signal, Punishers then punish any individual who did not signal. (We refer to Punishers with a threshold of  $\tau$  as “ $\tau$ -Punishers”). Non-punishers defect, but do not signal and do not punish, and as a result are punished if there are more than  $\tau + 1$  Punishers in the group. During subsequent stages, both types attempt to cooperate if they were punished during the first stage. Punishers punish defectors if at least  $\tau$  other individuals punished the last time a defection occurred. The cost of being punished,  $p$ , is greater than the cost of cooperating,  $c$ , so it is adaptive to choose cooperation if punishment is anticipated.

Note that cooperation is a facultative choice, not an inherited behavior. Punishers and Nonpunishers cooperate under exactly the same conditions, so there can be no linkage between

cooperation and punishment. After social interaction, individuals reproduce at a rate that is proportional to their payoff compared to the population average payoff. This allows the derivation of equations (21) that describe how natural selection changes the frequencies of the two types through time.

*Figure 1 about here.*

In the absence of genetic assortment, there are two longrun evolutionary outcomes (Figure 1). First, a population of all Nonpunishers is evolutionarily stable as long as solitary Punishers do not punish (i.e.,  $\tau > 0$ ). When Punishers are rare in the entire population, they will most often be alone in a group. Thus, they pay the cost of signaling but do not reap the benefits of cooperation, and as a result will have lower fitness than Nonpunishers. If Punishers are willing to punish alone ( $\tau = 0$ ), then they cannot invade a population of all Nonpunishers unless the benefits from cooperation are so large that a single punisher can recoup the costs of signaling and punishing everyone else in the group. Here, we assume that this condition is not satisfied so that only punishment by two or more Punishers pays.

Mixtures of Punishers and Non-punishers can also be evolutionarily stable. Punishers have an advantage over Non-punishers only in groups in which there are exactly  $\tau + 1$  Punishers because in such “threshold groups” each Punisher is necessary to sustain punishment and therefore cooperation. In groups with fewer Punishers, Punishers pay the cost of signaling, but because they do not punish, they enjoy no cooperative benefits. In groups with larger numbers of Punishers, a Punisher who switched to non-punishing would enjoy the same payoff from cooperation without paying the costs of signaling and punishment. This means that selection cannot favor  $\tau$ -Punishers unless they are in groups in which there are  $\tau + 1$  Punishers, and the benefits from cooperation are enough to compensate Punishers for the costs of signaling and punishment. Moreover, the advantage enjoyed by Punishers in these critical groups must be large enough to offset the payoff disadvantages suffered in groups with fewer or more punishers.

*Figure 2 about here*

The existence of a stable mixture of Punishers and Nonpunishers depends on the value of the punishment threshold,  $\tau$ . When the threshold is too low, punishment does not pay even at the threshold, and non-punishment is the only evolutionarily stable strategy. At higher thresholds, punishment does pay in threshold groups, and this means that punishment may be favored if such groups are sufficiently common. Thus, as the frequency of Punishers increases from zero, the fraction of groups with the threshold number of cooperators increases, and so does the expected fitness of Punishers (Figure 2). Once the fraction of threshold groups is high enough, the Punishers' advantage in these groups offsets their disadvantage in all other groups. Then natural selection will increase the frequency of Punishers. This marks the unstable (blue dots) equilibria shown in Figure 1 and the left-most zero intercept on the horizontal axis for each of the functions in Figure 2.

Further increases in the frequency of Punishers eventually decrease the fraction of threshold groups. When, as a result, the fitness of Punishers and Nonpunishers is equalized, there is a stable, polymorphic equilibrium (the upper red dots in Figure 1 and the right most horizontal axis intersection in Figure 2). As  $\tau$  increases, the frequency of Punishers at the polymorphic

equilibrium also increases, but the minimum initial frequency of punishers required for selection to move a population to this equilibrium increases, making it less accessible if Punishers are initially rare.

Note that at the stable polymorphic equilibrium, punishment is not altruistic: averaged over all groups, the long run benefits of punishment exactly compensate for the costs. However, it is mutually beneficial to the group (Figure 3) in that populations with the equilibrium frequency of punishers have higher average fitness than populations without Punishers. We show below that modest amounts of positive assortment in the formation of groups allow for the evolution of altruistic punishment.

*Figure 3 about here*

The results presented so far depend critically on two parameters, the extent of economies of scale in punishment,  $a$ , and the cost Punishers have to pay to signal their willingness to punish,  $q$ . Considering the first, were we to assume  $a = 1$  the total cost of punishing defectors would be independent of the number of punishers and much higher frequencies of punishment are required before punishment becomes evolutionarily stable (21). This supports the intuition that increasing returns is crucial, and therefore, the notion of coordinated punishment is important.

To determine the minimum cost of signaling,  $q$ , necessary to insure that the signal is honest, we introduce a third strategy, “Liar”. During the first period Liars defect, then signal that they are Punishers, thereby incurring the signaling cost, and avoiding punishment during the first period. In subsequent periods, Liars count the number of other group members that signaled in the first period, and cooperate if the number of Punishers is greater than  $\tau + 1$ . Liars never punish. This means that after the first period, Liars behave like Non-punishers and so receive the Non-punisher payoff. At equilibrium, Punishers and Nonpunishers have the same fitness, and thus Liars can invade if the cost of signaling is less than the expected punishment cost that Non-punishers suffer in the first period. Since some Nonpunishers avoid punishment during the first period, the expected cost of being punished is always less than  $p$  and thus Liars are excluded if  $q > p$ . In all results presented here we assume  $q = 1.5 p$ . Note too that there is no second-order free rider problem here. Individuals who always cooperate, but do not signal and do not punish always do worse than Non-punishers because they pay the cost of cooperation, but get punished anyway because they do not signal.

While punishment is evolutionarily stable in this model, so is non-punishment. A complete account of the evolution of cooperation must explain how punishing strategies can increase when rare. In their classic work on pair-wise reciprocity, Axelrod and Hamilton (22) suggest that a small amount of non-random assortment, such as interaction between weakly related group members, destabilizes uncooperative equilibria, but not cooperative ones. Subsequent work has shown that this principle holds in a wide range of pairwise cooperative interactions, but not in interactions in larger groups (12-14).

*Figure 4 about here*

To explore the effects of genetic assortment, we assume that the relatedness within groups is  $r$ . When  $r = 0$ , as we have assumed thus far, groups are formed at random. When  $r > 0$ , individuals

are more likely than by chance to interact with individuals like themselves. Figure 4 shows the equilibrium behavior assuming that  $r = 0.07$ , which is a rough estimate of the average relatedness within human foraging groups (20). For low thresholds ( $\tau \leq 3$ ) the only stable equilibrium is a mixture of Punishers and Non-punishers so Punishers can invade when rare. And due to the population structure (between group genetic differences) punishment may also be altruistic at the polymorphic equilibrium.

This result persists when groups are much larger ( $n = 72$ ) and for lower levels of relatedness if the benefit-cost ratio is somewhat higher (21). However, modest assortment does not allow punishment strategies with higher thresholds to invade populations of Punishers with lower thresholds, so there is no evolutionary process that would ratchet up the threshold levels. Thus consistent with ethnographic observation, the model predicts that only some individuals will engage in punishment. Note (from Figure 3), however, that even with a threshold of 3, groups achieve about two thirds of the maximum gains from cooperation attainable with higher thresholds.

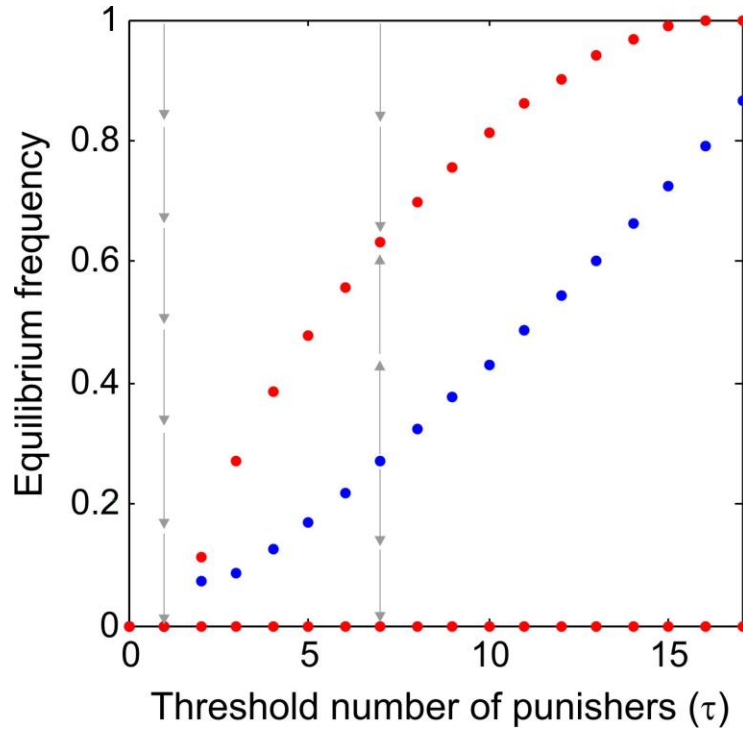
In our model the initial proliferation of punishment occurs under plausible levels of group genetic differences, it is based on its contingent nature, and it results in persistent and high levels of cooperation. It thus differs from the model of Hauert et al. (23) in which the population cycles between periods of cooperation, defection, and opting-out of the interaction entirely, the latter strategy invading the all-defect phase of the cycle and subsequently being invaded by cooperators. While their model applies to some forms of cooperation, we think that the present model is a more realistic representation of the nature and dynamics human cooperation (24,25).

1) The authors declare no competing interests. Affiliations: <sup>1</sup>Department of Anthropology, University of California, Los Angeles, CA 90064, USA. <sup>2</sup>Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA. <sup>3</sup>Department of Economics, University of Siena Piazza S.Francesco,7 - I 53100, Italy. <sup>4</sup>Department of Economics, Central European University, Nádor u. 9., 1051 Budapest, Hungary. Acknowledgements: We would like to thank the Behavioral Sciences Program of the Santa Fe Institute, the U.S. National Science Foundation, the European Science Foundation and the University of Siena for research support.

## References

1. C. Boehm, *Hierarchy in the Forest: The Evolution of Egalitarian Behavior* (Harvard University Press, Cambridge, MA, 2000).
2. P. Wiessner, *Evolution and Human Behavior* **23**, 407 (2002).
3. H. Kaplan, K. Hill, *Current Anthropology* **26**, 223 (1985).
4. E. Fehr, U. Fischbacher, *Nature* **425**, 785 (2003).
5. E. Fehr, S. Gächter, *American Economic Review* **90**, 980 (2000).
6. R. Boyd, H. Gintis, S. Bowles, P. J. Richerson, *Proceedings of the National Academy of Sciences* **100**, 3531 (2003).
7. R. Andrés Guzmán, C. Rodríguez-Sickert, R. Rowthorn, *Evolution and Human Behavior*, **28**, 112, (2008).
8. K. Sigmund, *Trends in Ecology and Evolution*, **11**, 593 (2007).

9. J. Henrich, R. Boyd. *Journal of Theoretical Biology*, **11**,79 (2001).
10. O. Bochet, T. Page, L. Putterman, *Journal of Economic Behavior and Organization* **60**, 11 (2006).
11. B. Herrmann, C. Thoni, S. Gächter, *Science* **319**, 1362 (2008).
12. R. Boyd, P. J. Richerson, *Journal of Theoretical Biology*, **132**: 337 (1988).
13. K. Panchanathan, R. Boyd, *Nature*, **224**:115 (2003).
14. R. Boyd, P. J. Richerson, *Ethology and Sociobiology*, **13**, 171 (1992).
15. Gardner, S. West. *American Naturalist*, **164**, 753 (2004).
16. P. Wiessner, *Human Nature* **16**, 115 (2005).
17. C. Boehm, *Current Anthropology* **34**, 227 (1993).
18. N. Q. Mahdi, *Ethology and Sociobiology* **7**, 295 (1986).
19. Ertan, T. Page, L. Putterman, *European Economic Review* **3**, 495 (2009).
20. S. Bowles, *Science* **314**, 1569 (2006).
21. Boyd, R., H. Gintis and S. Bowles, *Supporting Online Materials*
22. R. Axelrod, W. D. Hamilton, *Science*, **211**, 1390 (1981).
23. C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, *Science* **316**, 1905 (2007).
24. R. Boyd, S. Matthew, *Science* **316**, 1858 (2007).
25. S. Matthew, R. Boyd, *Proceedings of the Royal Society (B)* **276**, 1167 (2008).



**Figure 1. Equilibrium frequencies of Punishers with a threshold frequency of  $\tau$  when group members are unrelated.** For each value of  $\tau$ , the red dots give locally stable equilibrium frequencies of the punishing type, and the blue dots give interior unstable equilibrium frequencies. The arrows indicate the effect of natural selection at points above and below the red and blue dots. For  $\tau < 4$  the only stable equilibrium is a population without punishers. For larger thresholds, there are two stable equilibrium frequencies, zero, and a mixed strategy at which punishers and non-punishers coexist. In these cases, the unstable equilibria (blue dots) mark the frequency that Punishers must achieve before they are favored by selection. Benchmark parameters:  $c = 0.01$ ,  $b = 2c$ ,  $q = k = p = 1.5c$ ,  $r = 0$ ,  $a = 2$ ,  $e = 0.1$ ,  $n = 18$ ,  $T = 25$ .

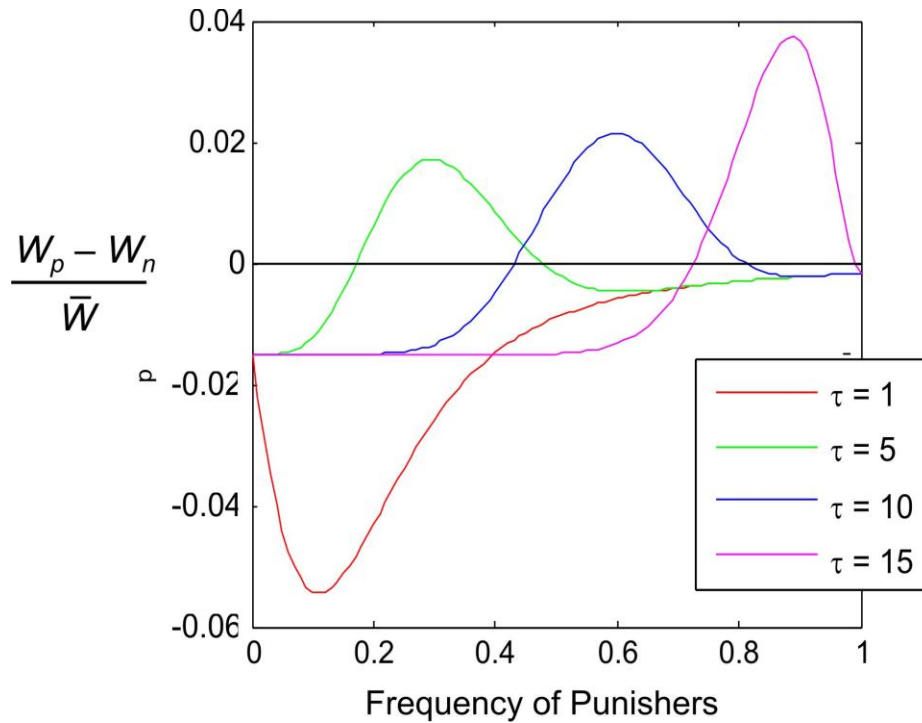
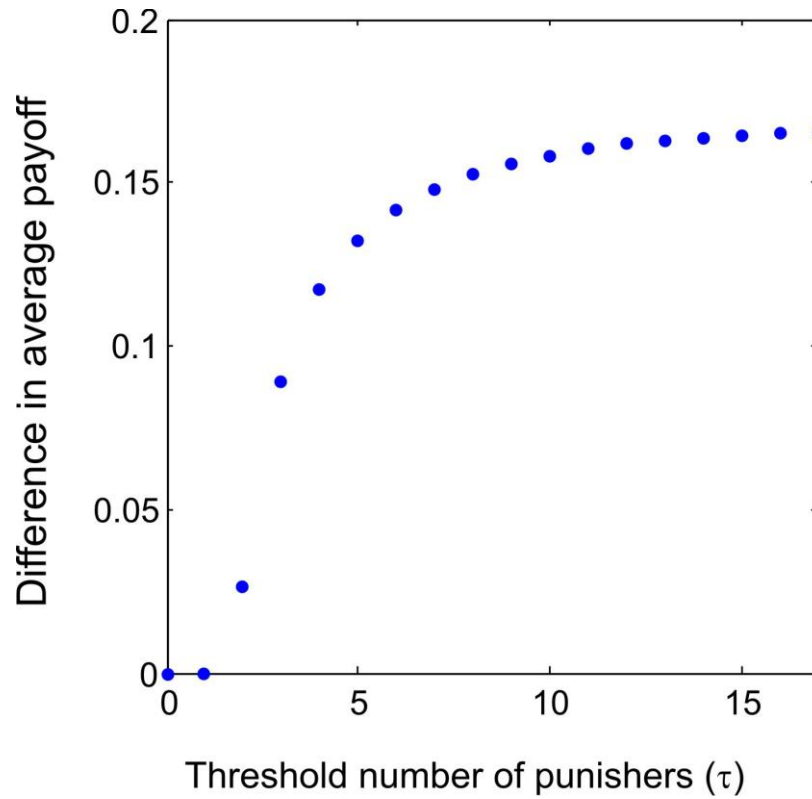
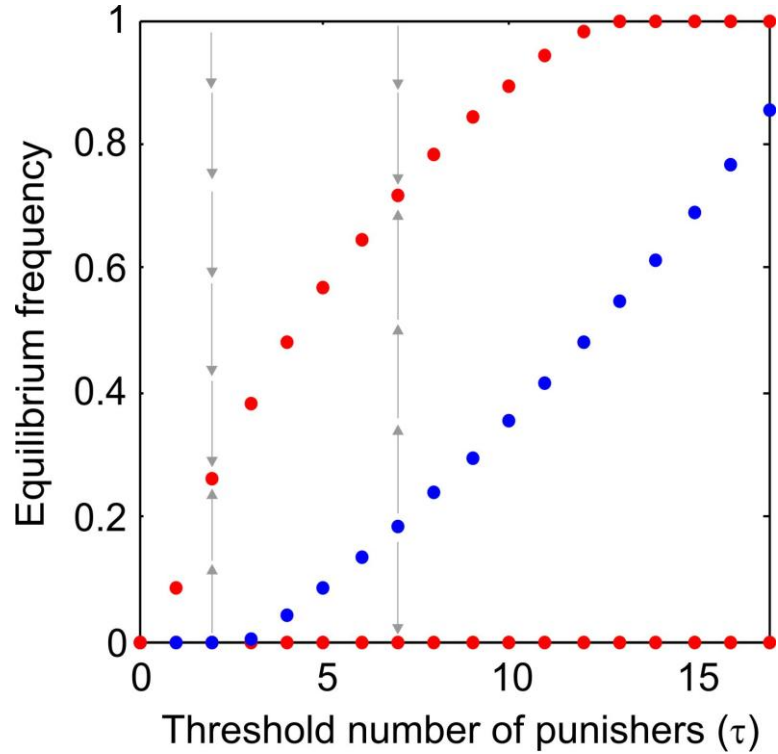


Figure 2. **The difference in fitness of (unrelated) Punishers ( $W_p$ ) and Non-punishers ( $W_n$ ) as a function of the frequency of Punishers.** When this difference is positive, Punishers increase in frequency, and when it is negative Punishers decrease in frequency. Equilibria occur when this difference is zero (evolutionarily stable when the function intersects the horizontal axis from above and unstable otherwise). When  $\tau = 1$ , punishment at the threshold does not pay for any frequency of punishers, and thus increasing the frequency of punishers from zero decreases their relative fitness. For larger values of  $\tau$ , punishment at the threshold does pay, and thus increasing the frequency of punishers increases their fitness. This leads to a stable polymorphic equilibrium at which punishers and non-punishers coexist. Parameters as in Figure 1.



**Figure 3. The difference in average fitness between the polymorphic equilibrium at which Punishers are present and the monomorphic nonpunishing equilibrium.** Whenever the polymorphic equilibrium exists it has higher average fitness, but near maximum benefit differences occur for relatively low thresholds. Benchmark parameters (Figure 1).



**Figure 4. Equilibrium frequencies of Punishers with a threshold frequency of  $\tau$  with modest assortment ( $r = 0.07$ ).** As in Figure 1, for each value of  $\tau$  the red dots give locally stable equilibrium frequencies of the punishing type, and the blue dots give unstable equilibrium frequencies. For  $\tau \leq 3$  rare Punishers can invade a population of Non-punishers, and the only stable equilibrium is a mixture of Punishers and Non-punishers in which cooperation is sustained in most groups. For larger thresholds, there are two stable equilibrium frequencies, zero, and a mixed strategy at which punishers and non-punishers coexist. In these cases, the unstable equilibria (blue dots) mark the frequency that Punishers must achieve before they are favored by selection. Benchmark parameters except  $r = 0.07$ .