
Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups

Robert Boyd* and Peter J. Richerson†

Zentrum für Interdisziplinäre Forschung, Universität Bielefeld, Bielefeld, Germany

Existing models suggest that reciprocity is unlikely to evolve in large groups as a result of natural selection. In these models, reciprocators punish noncooperation by withholding future cooperation, and thus also penalize other cooperators in the group. Here, we analyze a model in which the response is some form of punishment that is directed solely at noncooperators. We refer to such alternative forms of punishment as *retribution*. We show that cooperation enforced by retribution can lead to the evolution of cooperation in two qualitatively different ways. (1) If benefits of cooperation to an individual are greater than the costs to a single individual of coercing the other $n - 1$ individuals to cooperate, then strategies which cooperate and punish noncooperators, strategies which cooperate only if punished, and, sometimes, strategies which cooperate but do not punish will coexist in the long run. (2) If the costs of being punished are large enough, *moralistic* strategies which cooperate, punish noncooperators, and punish those who do not punish noncooperators can be evolutionarily stable. We also show, however, that moralistic strategies can cause any individually costly behavior to be evolutionarily stable, whether or not it creates a group benefit.

1. INTRODUCTION

Human behavior is unique in that cooperation and division of labor occur in societies composed of large numbers of unrelated individuals. In other eusocial species, such as social insects, societies are made up of close genetic relatives. According to contemporary evolutionary theory, cooperative behavior can only be favored by selection when social groups are formed so that cooperators are more

* Permanent address: Department of Anthropology, University of California, Los Angeles, California.

† Permanent address: Institute of Ecology, University of California, Davis, California.

Received May 10, 1990; revised November, 1991.

Address reprint requests to: Robert Boyd, Department of Anthropology, University of California, Los Angeles, CA 90024.

likely to interact with other cooperators than with noncooperators (Hamilton 1975; Brown et al. 1982; Nunney 1985). It is widely agreed that kinship is the most likely source of such nonrandom social interaction. Human society is thus an unusual and interesting special case of the evolution of cooperation.

A number of authors have suggested that human eusociality is based on reciprocity (Trivers 1971; Wilson 1975; Alexander 1987), supported by our more sophisticated mental skills to keep track of a large social system. It seems unlikely, however, that natural selection will favor reciprocal cooperation in sizable groups. An extensive literature (reviewed by Axelrod and Dion 1989, also see Hirschleifer and Martinez-Coll, 1988; Boyd 1988, 1989) suggests that cooperation can arise via reciprocity when *pairs* of individuals interact repeatedly. These results indicate that the evolutionary equilibrium in this setting is likely to be a contingent strategy with the general form: "cooperate the first time you interact with another individual, but continue to cooperate only if the other individual also cooperates." Several recent papers (Joshi 1987; Bendor and Mookherjee 1987; Boyd and Richerson 1988, 1989) present models in which larger groups of individuals interact repeatedly in potentially cooperative situations. These analyses suggest that the conditions under which reciprocity can evolve become extremely restrictive as group size increases above a handful of individuals.

In most existing models, reciprocators retaliate against noncooperators by withholding future cooperation. In many situations other forms of retaliation are possible. Noncooperators could be physically attacked, be made the targets of gossip, or denied access to territories or mates. We will refer to such alternative forms of punishment as *retribution*. It seems possible that selection may favor cooperation enforced by retribution even in sizable groups of unrelated individuals because, unlike withholding reciprocity, retribution can be made only against noncooperators, and, because with retribution, the magnitude of the penalty imposed on noncooperators is not limited by an individual's effect on the outcome of cooperative behavior.

Here, we extend the theory of the evolution of cooperation to include the possibility of retribution. We review the evolutionary models of the evolution of reciprocity in sizable groups, and present a model of the evolution of cooperation enforced by retribution. An analysis of this model suggests that retribution can lead to the evolution of cooperation in two qualitatively different ways.

1. If the long-run benefits of cooperation to a punishing individual are greater than the costs to that single individual of coercing all other individuals in a group to cooperate, then strategies which cooperate and punish noncooperators, strategies which cooperate only if punished, and, sometimes, strategies which cooperate but do not punish coexist at a stable equilibrium or stable oscillations.
2. If the costs of being punished are large enough, "moralistic" strategies

which cooperate, punish noncooperators, and punish those who do not punish noncooperators can be evolutionarily stable.

We also show, however, that moralistic strategies can cause any individually costly behavior to be evolutionarily stable, whether or not it creates a group benefit. Once enough individuals are prepared to punish any behavior, even the most absurd, and to punish those who do not punish, then everyone is best off conforming to the norm. Moralistic strategies are a potential mechanism for stabilizing a wide range of behaviors.

2. MODELS OF THE EVOLUTION OF RECIPROCITY

Models of the evolution of reciprocity among pairs of individuals share many common features. Each assumes that there is a population of individuals. Pairs of individuals are sampled from this population and interact a number of times. During each interaction individuals may either cooperate (*C*) or defect (*D*). The incremental fitness effects of each behavior define a single period prisoner's dilemma, and, therefore, cooperative behavior is altruistic in the sense that it reduces the fitness of the individual performing the cooperative behavior, but increases fitness of the other individual in the pair (Axelrod and Hamilton 1981; Boyd 1988). Each individual is characterized by an inherited *strategy* that determines how it will behave. Strategies may be fixed rules like, unconditional defection ("always defect"), or contingent ones like tit-for-tat ("cooperate during the first interaction; subsequently do whatever the other individual did last time"). The pair's two strategies determine the effect of the entire sequence of interactions on each pair member's fitness. An individual's contribution to the next generation is proportional to its fitness.

Analysis of such models suggests that lengthy interactions between pairs of individuals are likely to lead to the evolution of reciprocity. Reciprocating strategies, like tit-for-tat, leading to mutual cooperation, are successful if pairs of individuals are likely to interact many times. A population in which unconditional defection is common can resist invasion by cooperative strategies under a wide range of conditions. However, there seem to be a variety of plausible mechanisms that allow reciprocating strategies to increase when rare. Axelrod and Hamilton (1981) and Axelrod (1984) have shown that a very small degree of assortative group formation, when coupled with the possibility of prolonged reciprocity, allows strategies like tit-for-tat to invade noncooperative populations. Other mechanisms have been suggested by Peck and Feldman (1985), Boyd and Lorberbaum (1987), and Feldman and Thomas (1987).

Recent work suggests that these conclusions do not apply to larger groups. Joshi (1987) and Boyd and Richerson (1988) have independently analyzed a model in which n individuals are sampled from a larger population, and then interact repeatedly in an n -person prisoner's dilemma. In

this model, cooperation is costly to the individual, but beneficial to the group as a whole. This work suggests that increasing the size of interacting social groups reduces the likelihood that selection will favor reciprocating strategies. As in the two individual cases, if groups persist long enough, both reciprocal and noncooperative behavior are favored by selection when they are common. For large groups, however, the conditions under which reciprocity can increase when rare become extremely restrictive. Bendor and Mookherjee (1987) show that when errors occur, reciprocal cooperation may not be favored in large groups even if they persist forever. Boyd and Richerson (1989) derived qualitatively similar results in which groups were structured into simple networks of cooperation.

Intuitively, increasing group size places reciprocating strategies on the horns of a dilemma. To persist when common, they must prevent too many defectors in the population from receiving the benefits of long-term cooperation. Thus, reciprocators must be provoked to defect by the presence of even a few defectors. To increase when rare, there must be a substantial probability that the groups with less than this number of defectors will form. This problem is not great when pairs of individuals interact; a relatively small degree of assortative group formation will allow reciprocating strategies to increase. As groups become larger, however, both of these requirements can be satisfied only if the degree of assortment in the formation of groups is extreme.

This result should be interpreted with caution. Modeling social interaction as an n -person prisoner's dilemma means that the only way a reciprocator can punish a defector is by withholding future cooperation. There are two reasons to suppose that cooperation might be more likely to evolve if cooperators could retaliate in some other way. First, in the n -person prisoners' dilemma, a reciprocator who defects in order to punish defectors induces other reciprocators to defect. These defections induce still more defections. More discriminating retribution would allow defectors to be penalized without generating a cascade of defection. Second, in the n -person prisoners' dilemma the severity of the sanction is limited by an individual's effect on the whole group, which becomes diluted as group size increases. Other sorts of sanctions might be much more costly to defectors, and therefore allow rare cooperators to induce others to cooperate in large groups.

There is also a problem with retribution. Why should individuals punish? If being punished is sufficiently costly, it will pay to cooperate. However, by assumption, the benefits of cooperation flow to the group as a whole. Thus as long as administering punishment is costly, retribution is an altruistic act. Punishment is beneficial to the group, but costly to the individual, and selection should favor individuals who cooperate but do not punish. This problem is sometimes referred to as the problem of "second order" cooperation (Oliver 1980; Yamagishi 1986).

A recent paper by Axelrod (1986) illustrates the problem of second-order cooperation. Axelrod analyzes a model in which groups of individuals

interact for two periods. During the first period individuals may cooperate or defect in an n -person prisoner's dilemma, and in the second, individuals who cooperated on the first move have the opportunity to punish those individuals who did not cooperate at some cost to themselves. Axelrod shows that punishment may expand the range of conditions under which cooperation could evolve. However, the strategy of cooperating but not punishing was precluded. As Axelrod notes, such second-order defecting strategies would always do better because second-order punishment of non-punishers is not possible.

The problem of second-order cooperation has been partly solved by Hirshleifer and Rasmusen (1989). They consider a game theoretic model in which a two-stage game consisting of a cooperation stage followed by a punishment stage is repeated a number of times. They show that if punishment is costless, then the strategy of cooperating, punishing noncooperators, and punishing non-punishers is what game theorists call a "perfect equilibrium." (The perfect equilibrium is a generalization of the Nash equilibrium which is useful in repeated games. See Rasmusen (1989) for an excellent introductory discussion of game theoretic equilibrium concepts.) Because it is a game theoretic model it does not provide information about the evolutionary dynamics. It also seems possible that if the model were extended to an infinite number of periods, a similar strategy would be evolutionarily stable even if punishment is costly.

Here we consider evolutionary properties of an infinite period model of cooperation with the possibility of punishment that is similar to Hirshleifer and Rasmusen's. We will perform the analysis in three stages. First, we describe the basic structure of the model. Then, we consider populations in which there are cooperators who punish defection, and a variety of strategies which initially defect and then respond to punishment in different ways. The goal is to investigate the evolutionary dynamics introduced by retribution without the complications introduced by second-order defection and second-order punishment. Finally, we consider the effects of these complications.

3. DESCRIPTION OF THE MODEL

Suppose that groups of size n are sampled from a large population, and interact repeatedly. The probability that the group persists from one interaction to the next is w , and thus the probability that it persists for t or more interactions is w^{t-1} . Each interaction consists of two stages, a cooperation stage followed by a punishment stage. During the cooperation stage an individual can either cooperate (C) or defect (D). The incremental effect of a single cooperation stage on the fitness of an individual depends on that individual's behavior and the behavior of other members of the group as fol-

lows: Let the number of *other* individuals choosing *C* during a particular turn be *i*. Then the payoffs to individuals choosing *C* and *D* are:

$$V(C|i) = (b/n)(i + 1) - c \quad (1)$$

$$V(D|i) = (b/n)i \quad (2)$$

where $b > c$ and $c > b/n$. Increasing the number of cooperators increases the payoff for every individual in the group, but each cooperator would be better off switching to defection. (This special case of the *n*-person prisoners' dilemma has been used in economics and political science to represent provision of public goods (Hardin 1982). It is also identical to the linear model of social interactions used in most kin selection models.) During the punishment stage any individual can punish any other individual. Punishing another individual lowers the fitness of the punisher an amount *k*, and the fitness of the individual being punished an amount *p*.

Each individual is characterized by an inherited "strategy" which specifies how it will behave during any time period based on the history of its own behavior and the behavior of other members of the group up to that point. The strategy specifies whether the individual will choose cooperation or defection during the cooperation stage, and which other individuals, if any, it will punish during the punishment stage. Strategies can be unconditional rules like the asocial rule "never cooperate/never punish." They can also be contingent rules like "always cooperate/punish all individuals who didn't cooperate during the cooperation stage."

We assume that individuals sometimes make errors. In particular we suppose that any time an individual's strategy calls for cooperation there is a probability $e > 0$ that the individual will instead defect "by mistake." This is the only form of error we investigate. Individuals who mean to defect always defect, and individuals always either punish or do not punish according to the dictates of their strategy.

Groups are formed according to the following rule: the conditional probability that any other randomly chosen individual in a group has a given strategy S_j given that the focal individual also has S_i is given by:

$$\Pr(S_j|S_i) = r + (1 - r)q_j \quad (3)$$

where q_i is the frequency of the strategy S_i in the population before social interaction, and $0 < r < 1$. The conditional probability that any other randomly chosen individual in a group has some other strategy S_j given that the focal individual has S_i is given by:

$$\Pr(S_j|S_i) = (1 - r)q_j \quad (4)$$

When $r = 0$, social interaction occurs at random. When $r > 0$, social interaction is assortative. There is a chance r of drawing an individual with the same strategy as the focal individual and a chance $1 - r$ of picking an individual at random from the population (who will also be identical to the focal individual with probability equal to the frequency of the focal indivi-

dual's strategy in the population). If strategies are inherited as haploid sexual traits, r is just the coefficient of relatedness. For other genetic models, r is not exactly equal to the coefficient of relatedness. However, it is a good approximation for rare strategies and thus is useful for determining the conditions under which a rare reciprocating strategy can invade a population in which all defection is common.

After all social interactions are completed, individuals in the population reproduce. The probability of reproduction is determined by the results of social behavior. Thus, the frequency of a particular strategy, S_i , in the next generation, q'_i , is given by:

$$q'_i = \frac{q_i W(S_i)}{\sum_j q_j W(S_j)} \quad (5)$$

where $W(S_i)$ is the average payoff of individuals using strategy S_i in all groups weighted by the probability that different types of groups occur. (As argued by Brown et al. 1982, this assumption is consistent with haploid genetic inheritance of strategies and some simple forms of cultural transmission.) We then ask, which strategies or combinations of strategies can persist?

4. RESULTS

4.1. No Second-Order Defection

First we analyze the evolutionary dynamics of retribution with second-order defection is excluded. To do this, we consider a world in which only the following two strategies are possible.

Cooperator-punishers (P). During each interaction (1) cooperate, and (2) punish all individuals who did not cooperate during the cooperation stage.

Reluctant cooperators (R_1). Defect until punished once, then cooperate forever. Never punish.

We temporarily exclude strategies which cooperate, but do not punish to eliminate the possibility of second-order defection. We also exclude strategies that continue to defect after one act of punishment. This latter assumption is *not* harmless. We show in the appendix that if R_1 is replaced by unconditional defection, then (1) cooperation is much less likely to evolve, and (2) R_1 may not be able to invade a population in which unconditional defection is common. The present analysis is justified for two reasons: First, it provides a best case for the evolution of cooperation, and second, there is abundant empirical evidence that organisms do respond to punishment.

When groups are formed at random ($r = 0$), such a population can persist at one of three stable equilibria (or ESSs):

- All individuals are R_1 —no one cooperates.

- All individuals are *P*—everyone cooperates.
- Most individuals are *R*₁, but a minority are *P*—most are induced to cooperate by the punishing few.

In what follows we describe and interpret the conditions under which each of these ESSs can exist. Proofs are given in the appendix.

Reluctant cooperators resist invasion by the cooperating, punishing strategy whenever the cost to a cooperator-punisher of cooperating and punishing $n - 1$ reluctant cooperators exceeds the benefit to that punisher that results from the cooperation that is induced by its punishment. It can be shown that the responsive defecting strategy *R*₁ can be invaded by the cooperating, punishing strategy *P* as long as:

$$k(n - 1) + (c - b/n) < \frac{1}{1 - w} \left(w(b - c) - \frac{ek(n - 1)}{1 - e} \right) \quad (6)$$

initial cost of cooper-
ating and punishing
long run benefit induced
by punishing

When cooperator-punishers are rare, and groups are formed at random, virtually all cooperator-punishers will find themselves in a group in which the other $n - 1$ individuals are defectors. The left-hand side of (6) gives the fitness loss associated with cooperating, and then punishing $n - 1$ defectors during the first interaction. The right-hand side of (6) gives the long-term net fitness benefit of the cooperation that results from punishment. The term $w(b - c)/(1 - w)$ is the long-term fitness benefit from the induced cooperation by *R*₁ individuals and the term proportional to e is the long-run cost that results from having to punish erroneous defections. Thus, if this term is positive, *P* can invade if w is large enough.

If the cooperator-punisher strategy, *P*, can increase when rare, punishing is not altruistic. Retribution induces cooperation that creates benefits sufficient to compensate for its cost. The longer groups persist, the larger the benefit associated with cooperation. Thus, as long as error rates are low or the benefits of cooperation are large, longer interactions will permit cooperative strategies to invade, even if groups are formed at random. Also notice that the condition for *R*₁ to be invaded does not depend on p , the cost of being punished. As one would expect, increasing the group size or the error rate make it harder for the cooperative strategy to invade.

The cooperating-punishing strategy, *P*, is evolutionarily stable as long as

$$p(n - 1) > c - b/n + \frac{ek(n - 1)}{(1 - w)(1 - e)} \quad (7)$$

cost of being punished
cost of cooperating and punishing

The first term on the right-hand side of (7) gives the cost of cooperating during one interaction; the term on the left-hand side is the cost of being punished by $n - 1$ other individuals, and the second term on the right-hand side is the cost of punishing mistakes over the long run. The rare *R*₁ indi-

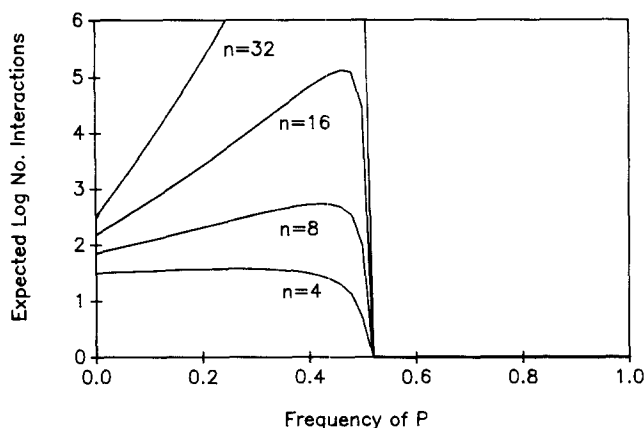


FIGURE 1. The equilibrium frequency of P for a given expected number of interactions for different group sizes ($n = 8, 16, 32$) assuming that $e = 0.001$. For these parameter values populations consisting of all P are always at a stable equilibrium. Populations without P individuals are also always an equilibrium, but it may be either stable or unstable. To find the polymorphic equilibria, pick a number of expected interactions and group size, and then determine the frequencies of P at which the horizontal line at that value of $\log(1/(1-w))$ intersects the curve at that value of n . If the horizontal line lies below the curve for some q_P , then the frequency of P increases; if it lies above the curve, the frequency of P decreases. Thus, if there is only one polymorphic equilibrium (e.g., $n = 4$, $\log(1/(1-w)) = 1$), it is unstable and $q_P = 0$ is stable. If there are two polymorphic equilibria (e.g., $n = 16$, $\log(1/(1-w)) = 3$), the polymorphic equilibrium with the lower frequency of P is stable, and the other polymorphic equilibrium and $q_P = 0$ are both unstable. Finally, if there are no polymorphic equilibrium (e.g., $n = 8$, $\log(1/(1-w)) = 4$), the only stable equilibrium is $q_P = 1$.

vidual suffers the cost of punishment, but avoids the cost of cooperating on the first turn and the cost of punishing erroneous defection over the long run. Notice that this condition is independent of the long-run expected benefit associated with cooperation (because it does not contain terms of the form $b/(1-w)$). It depends only on the cost of the cooperation to the individual and the costs of punishing and being punished. Thus, retribution can stabilize cooperation, but this stability does not result from the mutual benefits of cooperation.

There is a stable internal equilibrium at which both P and R_1 are present whenever (1) neither R_1 nor P are ESSs, or (2) R_1 is not an ESS but P is, and the condition (24) (given in the appendix) is satisfied. We have not been able to derive an expression for the frequency of P at the internal equilibrium. Figure 1 shows the frequency of P at this equilibrium determined numerically as a function of the expected number of interactions ($\log(1/(1-w))$) for various group sizes. When groups persist for only a few interactions, both P and R_1 are ESSs. Increasing the number of interaction eventually destabilizes R_1 , and allows a stable internal equilibrium to exist. Further increases in the expected number of interactions destabilize P leaving the internal equilibrium as the only stable equilibrium.

Without second-order defection, cooperation can persist at two qualitatively different equilibria: either cooperative strategies coexist with non-cooperative strategies at a polymorphic equilibrium, or all individuals in the population are cooperative. When the cooperator-punisher strategy is very rare, it will increase whenever the benefit from long-run cooperation to an individual punisher exceeds the cost of the punishment necessary to induce reluctant cooperators to cooperate. As cooperator-punishers become more common, more reluctant cooperators find themselves in groups with at least one cooperator-punisher, and thus they enjoy the benefits of long-run cooperation without bearing the costs associated with punishing. Thus the relative fitness of cooperator-punishers declines. As cooperator-punishers become still more common, reluctant cooperators are punished more harshly during the initial interaction and their relative fitness declines.

Assortative group formation has both positive and negative effects on the conditions under which punisher-cooperators evolve. When there is assortative group formation, individuals are more likely to find themselves in groups with others like themselves than chance alone would dictate. Such assortment decreases the cost of cooperating and punishing because cooperators are more likely to receive the benefits that result from the cooperative acts of others than are noncooperators, and because cooperator-punishers need to punish fewer noncooperators on the first interaction. However, assortment decreases the long-run benefit associated with punishment because cooperator-punishers are more likely to be punished for erroneous defection. (Assortment increases the amount of punishment which an inadvertently defecting cooperator-punisher receives.) The second effect becomes more pronounced the longer groups last because cooperator-punishers will make more errors. The negative effect will predominate whenever the following condition is satisfied:

$$(1 - e)(b/n + k) < \frac{ep}{1 - w} \quad (8)$$

When (8) is satisfied, assortment increases the range of conditions under which R_1 is an ESS, decreases the range of conditions under which P is an ESS, and, if a stable internal equilibrium exists, decreases the frequency of P at that equilibrium. Note that the negative effects increase as the expected number of interactions increase. When (8) is not satisfied, increasing r decreases the range of parameters under which R_1 is an ESS, increases the range under which P is an ESS, and may either increase or decrease the frequency of P at internal equilibria.

4.2. Second-Order Defection

When punishers are common, cooperation is favored because cooperative individuals avoid punishment. Thus, if punishment is costly, punishment may be an altruistic act. It is costly to the individual performing the pun-

ishment, but benefits the group as a whole. This argument suggests that individuals who cooperate, but do not punish, should be successful. In the previous model (and that of Axelrod 1986) cooperators always punish non-cooperators, and thus this conjecture could not be addressed. To allow for second-order defection, consider a model in which P and R_1 compete with the following strategy.

Easy-going cooperator (E). Always cooperate, never punish.

When second-order defection is possible, neither E nor P is ever an ESS. A population in which P is common can always be invaded by E , because easy-going cooperators get the benefits of cooperation without incurring the cost of enforcement. A population in which E is sufficiently common can always be invaded by R_1 , because reluctant cooperators can enjoy the benefits of cooperation without fear of punishment.

R_1 is an ESS whenever punishment does not pay (i.e., (6) is not satisfied). At this ESS there is no cooperation because reluctant cooperators behave as unconditional defectors. If the long-run benefits of cooperation to an individual are not sufficient to offset the cost of coercing all the other members of the group to cooperate, the noncooperators can resist invasion by punishing or cooperating strategies. Persistent noncooperation is not the only possible outcome, however, under this condition. If P can resist invasion by R_1 (i.e., (7) is satisfied), then simulation studies indicate that there may be persistent oscillations involving all three strategies. Such oscillations seem to require that the cost of being punished is much greater than the cost of punishing ($p \gg k$) and the benefits of cooperation barely exceed the cost ($b \approx c$).

If punishment does pay, the long-run outcome is a mix of reluctant cooperators who coexist with punisher cooperators and, sometimes, easy-going cooperators. This can happen in three different ways:

- There can be a stable mix of reluctant cooperators and cooperator-punishers. Such a stable equilibrium exists anytime there is a stable polymorphic equilibrium on the $R_1 - P$ boundary in the absence of E . If, in addition, P is not an ESS in the absence of E , this mixture of reluctant cooperators and cooperator-punishers is the only equilibrium, and numerical simulations suggest that the polymorphic equilibrium is globally stable. Thus, at equilibrium populations will consist of a majority of reluctant cooperators with a minority of cooperator-punishers. E cannot invade because rare E individuals often find themselves in groups without a cooperator-punisher, and thus pay the cost of cooperation without receiving the long-run benefits of cooperation. Punishers in all groups received the benefits of long-term cooperation.
- If there is no polymorphic equilibrium on the $R_1 - P$ boundary (i.e., in the absence of E), then there is a single interior equilibrium point at which all three strategies are present. We have not been able to

derive an expression for the frequencies of the three traits at these interior equilibria, nor determine when they are stable. Numerical simulation indicates that when an interior equilibrium exists, it is almost always stable.

- The mixture of all three strategies can oscillate. When P is stable in the absence of E , the frequencies of the three strategies may oscillate indefinitely. Simulation studies suggest that this outcome only occurs under relatively rare parameter combinations.

In each case, as group size increases, the average frequency of cooperative strategies typically declines to a quite low level. However, the average frequency of groups with at least one P individual and therefore groups in which cooperation occurs over the long run, can remain at substantial levels even when groups are large. One must keep in mind, however, that this conclusion presupposes that individual punishers can afford to punish every noncooperator in the group. A model in which the capacity to punish is limited would presumably stabilize at some higher frequency of punishers as group size increased.

4.3. Moralistic Strategies

The results of the previous section suggest that strategies which attempt to induce cooperation through retribution can always be invaded when they are common by strategies which cooperate but do not punish. However, such is not the case. Consider the strategy:

Moralists (M). Always cooperate, and punish individuals who are not in “good standing.” Individuals are in good standing if they have behaved according to M since the last time they were punished or the beginning of the interaction.

Thus, moralists punish individuals who do not cooperate. But they also punish those who do not punish noncooperators and those that do not punish nonpunishers. Each M individual punishes others at most once per turn. Once an individual is punished, it can avoid further punishment by cooperating, punishing noncooperators, and punishing nonpunishers (thus returning to good standing).

Moralists can resist invasion by reluctant cooperators (R_1) whenever

$$(n - 1)p \left(1 - \frac{w}{1 - w} (1 - (1 - e)^{n-1}) \right) > c - b/n + \frac{e(n - 1)k}{(1 - w)(1 - e)}$$

cost of being punished
cost of cooperating and punishing

(9)

The left-hand side of (9) gives the cost to an R_1 individual of being punished. It is proportional to the number of interactions because such reluctant co-

operators are punished every time there is an error. The right-hand side is the cost of cooperating and punishing. Thus, as long as the error rate is not exactly zero, moralists can resist invasion by R_1 under a wider range of conditions that can P .

Moralists can resist invasion by easy-going cooperators (E) whenever the following condition is satisfied:

$$(1 - (1 - e)^n)wp > ek \quad (10)$$

If errors occur only infrequently ($ne \ll 1$), then this condition simplifies to become $nwp > k$. Thus, unless punishing is much more costly than being punished, moralists can resist invasion by easy-going cooperators.

In fact, as Hirshleifer and Rasmusen (1989) have pointed out, moralistic aggression of this kind is a recipe for stabilizing *any* behavior. Notice that neither condition (9) or (10) involves terms representing the long run benefits of cooperation (i.e., terms of the form $b/(1 - w)$). When M is common, rare individuals deviating from M are punished; otherwise they have no effect on the behavior of the group. Thus, as long as being punished by all the other members of the group is sufficiently costly compared to the individual benefits of not behaving according to M , M will be evolutionarily stable. It does not matter whether or not the behavior produces group benefits. The moralistic strategy could require any arbitrary behavior—wearing a tie, being kind to animals, or eating the brains of dead relatives. Then M could resist invasion by individuals who refuse to engage in the arbitrary behavior unless punished as long as (9) was satisfied (where $c - b/n$ is the cost of the behavior), and resist invasion by individuals who perform the behavior but do not punish others as long as (10) is satisfied.

5. DISCUSSION

Our results suggest that problems of second-order cooperation can be overcome in two quite different ways: First, even though retribution creates a group benefit, it need not be altruistic. If defectors respond to punishment by a single individual by cooperating, and if the long-run benefits to the individual punisher are greater than the costs associated with coercing other group members to cooperate, then the strategy which cooperates and punishes defectors can increase when rare, and will continue to increase until an interior equilibrium is reached. At this equilibrium the punishing strategy coexists with strategies which initially defect but respond to punishment by cooperating, and, sometimes, strategies which cooperate but do not punish. For plausible parameter values, the punishing strategy is rarer than the other two strategies at such an equilibrium. However, since a single punisher is sufficient to induce cooperation, cooperating groups are nonetheless quite common.

Increasing group size reduces the likelihood that this mechanism will

lead to the evolution of cooperation because it increases the cost of coercion. This effect, however, is not nearly so strong as previous models in which defection was punished by withdrawal of cooperation. In those models (Joshi 1987; Boyd and Richerson 1988, 1989) a linear increase in group size requires an exponential increase in the expected number of interactions necessary for cooperation to increase when rare. In the present model, the same increase in group size requires only a linear increase in the expected number of interactions.

Moralistic strategies which punish defectors, individuals who do not punish noncooperators, and individuals who do not punish nonpunishers can also overcome the problem of second-order cooperation. When such strategies are common, rare noncooperators are selected against because they are punished. Individuals who cooperate but do not punish are selected against because they are also punished. In this way, selection may favor punishment, even though the cooperation that results is not sufficient to compensate individual punishers for its costs.

It is not clear whether moralistic strategies can ever increase when rare. We have not presented a complete analysis of the dynamics of moralistic strategies because to do so in a sensible way would require the introduction of additional strategies, a consideration of imperfect monitoring of punishment, and a consideration of more general temporal patterns of interaction. We conjecture, however, that the dynamics will be roughly similar to the dynamics of P and R_1 in the case in which there is no stable internal equilibrium: Both defecting and moralistic strategies will be evolutionarily stable. Increasing the degree of assortment will mean that moralists will have fewer defectors to punish, but will be punished more when they err. Assortative social interaction will not interact with group benefits in a way that will allow moralistic strategies to increase.

It is also interesting that moralistic strategies stabilize any behavior. The conditions which determine whether M can persist when rare are independent of the magnitude of the group benefit created by cooperation. The moralistic strategy could stabilize any behavior equally well, whether it is beneficial or not. If our conjecture about the dynamics of M is correct, then the dynamics will not be strongly effected by whether or not the sanctioned behavior is group beneficial.

This result is reminiscent of the "folk theorem" from mathematical game theory. This theorem holds that in the repeated prisoner's dilemma with a constant probability of termination (the case analyzed by Axelrod and most other evolutionary theorists) strategies leading to *any* pattern of behavior can be a game theoretic perfect equilibrium (Rasmusen 1989). The proof of this theorem relies on the fact that if there is enough time available (on average) for punishment, then individuals can be induced to adopt any pattern of behavior. Thus in games without a known endpoint, game theory may predict that anything can happen. This result, combined with the fact that nobody lives forever, has led many economists to restrict their analyses

to games with known endpoints. The diversity of equilibria here and in the nonevolutionary analysis can be regarded as a flaw or embarrassment for the analysis.

We prefer to take these results as telling us something about the evolution of social behavior. Games without a known endpoint seem to us to be a good model for many social situations. Although nobody lives forever, social groups often persist much longer than individuals. When they do, individuals can expect to be punished up until their own last act. Even dying men are tried for murder, and in many societies one's family is also subject to retribution. If one accepts this argument, then it follows that moralistic punishment is inherently diversifying in the sense that many different behaviors may be stabilized in exactly the same environment. It may also provide the basis for stable among-group variation. Such stable among-group variation can allow group selection to be an important process (Boyd and Richerson 1985, 1990 a,b) leading to the evolution of behaviors that increase group growth and persistence.

6. CONCLUSIONS

Cooperation enforced by retribution is strikingly different from reciprocity in which noncooperation is punished by withdrawal of cooperation. We think two features of this system are interesting and warrant further study:

1. Cooperation may be possible in larger groups than is the case with reciprocity. This effect invites further study of the limitations on the ability of single individuals to punish, and how coalitions of punishers might or might not be able to induce reciprocity in very large groups.
2. In the model studies here, punishers collect private benefit by inducing cooperation in their group that compensates them for punishing, while providing a public good for reluctant cooperators. There are often polymorphic equilibria in which punishers are relatively rare, generating a simple political division of labor reminiscent of the "bigman" systems of New Guinea and elsewhere. This finding invites study of further punishment strategies. Consider, for example, strategies that punish but do not cooperate. Such individuals might be able to coerce more reluctant cooperators than cooperator-punishers, and therefore support cooperation in still larger groups. If so, such models might help understand the evolution of groups organized by full-time specialized, "parasitical" coercive agents like tribal chieftains.

The importance of the study of retribution can hardly be underestimated. The evolution of political complexity in human societies over the last few thousand years depended fundamentally on the development of a variety of coercive strategies similar to those we have investigated here.

We thank Alan Rogers for useful comments and for carefully checking every result in this paper. Joel Peck also provided helpful comments.

REFERENCES

- Alexander, R.D. *The Biology of Moral Systems*, New York: Aldine De Gruyter, 1987.
- Axelrod, R. *The Evolution of Cooperation*, New York: Basic Books, 1984.
- . The evolution of norms. *American Political Science Review* 80: 1095–1111, 1986.
- , and Dion, D. The further evolution of cooperation. *Science* 232: 1385–1390, 1989.
- , and Hamilton, W.D. 1981. The evolution of cooperation. *Science* 211: 1390–1396, 1981.
- Bendor, J. and Mookherjee, D. Institutional structure and the logic of ongoing collective action. *American Political Science Review* 81: 129–154, 1987.
- Boyd, R. Is the repeated prisoner's dilemma game a good model of reciprocal altruism? *Ethology and Sociobiology* 9: 211–221, 1988.
- . Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *Journal of Theoretical Biology* 136: 47–56, 1987.
- , and Lorberbaum, J. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature* 327: 58–59, 1987.
- , and Richerson, P.J. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press, 1985.
- and —. The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology* 132: 337–356, 1988.
- and —. The evolution of indirect reciprocity. *Social Networks* 11: 213–236, 1989.
- and —. Culture and cooperation. In *Beyond Self Interest*, Jane Mansbridge (Ed.). Chicago: University of Chicago Press, 1990a, pp. 111–132.
- and —. Group selection among alternative evolutionarily stable strategies. *Journal of Theoretical Biology* 145: 331–324, 1990b.
- Brown, J.S., Sanderson, M.J., and Michod, R.E. Evolution of social behavior by reciprocation. *Journal of Theoretical Biology* 99: 319–339, 1982.
- Feldman, M.W. and Thomas, E.R.C. Behavior dependent contexts for repeated plays of the repeated prisoner's dilemma II: Dynamical aspects of the evolution of cooperation. *Journal of Theoretical Biology* 129: 297–314, 1987.
- Hamilton, W.D. Innate social aptitudes of man; An approach from evolutionary genetics. In *Biosocial Anthropology*, R. Fox (Ed.). London: Malaby, 1975, pp. 135–232.
- Hardin, R. *Collective Action*, Baltimore: Johns Hopkins University Press, 1982.
- Hirshleifer, D. and Rasmusen, E. Cooperation in a repeated prisoner's dilemma with ostracism. *Journal of Economic Behavior and Organization*, 12: 87–106, 1989.
- , and Martinez-Coll, J. What strategies can support the evolutionary emergence of cooperation? *Journal of Conflict Resolution* 32: 367–398, 1988.
- Joshi, N.V. Evolution of cooperation by reciprocation within structured demes. *Journal of Genetics* 1: 69–84, 1987.
- Michod, R.E. and Sanderson, M.J. Behavioural structure and the evolution of cooperation. In *Evolution: Essays in Honor of John Maynard Smith*, Cambridge: Cambridge University Press, 1985.
- Nunney, L. Group selection, altruism, and structured deme models. *American Naturalist* 126: 212–230, 1985.
- Oliver, P. Rewards and punishments as selective incentives for collective action: Theoretical investigations. *American Journal of Sociology* 85: 1356–1375, 1980.
- Peck, J. and Feldman, M.W. The evolution of helping behavior in large, randomly mixed populations. *American Naturalist* 127: 209–221, 1985.
- Rasmusen, E. *Games and Information: An Introduction to Game Theory*, Oxford: Basil Blackwell, 1989.
- Trivers, R. The evolution of reciprocal altruism. *Quarterly Review of Biology* 46: 35–57, 1971.
- Wilson, E.O. *Sociobiology: The New Synthesis*, Cambridge: Belknap/Harvard University Press, 1975.

Yamagishi, T. The provisioning of sanctioning as a public good. *Journal of Personality and Social Psychology* 51: 100–116, 1986.

A. APPENDIX

A.1. Sensitivity of the Model to the Response to Punishment

The effects of punishment on the evolution of cooperation are strongly affected by the extent to which defector responds to punishment by cooperating. To see this consider a game in which cooperator-punishers (*P*) compete with the following nonresponsive strategy.

Unconditional defectors (U). Never cooperate. Never punish.

Many of the evolutionary properties of the two-person repeated prisoner’s dilemma can be derived considering a model in which only tit-for-tat (TFT, cooperate on the first move, and punish each defection by defecting) and ALLD (always defect) are present. Our strategies *P* and *U* seem like the natural generalizations of TFT and ALLD to the *n*-person game with punishment, and one might (as we did) expect that their evolutionary dynamics would be similar. This expectation is largely incorrect. Understanding why provides useful insight into the evolutionary effects of punishment. For simplicity, we assume that there are no errors ($e = 0$) throughout this section.

Let *j* be the number of the other $n - 1$ individuals in the group who are *P*. The expected fitness of *U* individuals given *j* is:

$$V(U|j) = \frac{(b/n - p)j}{1 - w} \tag{11}$$

Similarly, the expected fitness of *P* individuals given *j* is:

$$V(P|j) = \frac{b/n(j + 1) - c - (n - 1 - j)k}{1 - w} \tag{12}$$

The expected fitness of *U* individuals averaged overall groups is:

$$\begin{aligned} W(U) &= \sum_{j=0}^{n-1} m(j|U)V(U|j) \\ &= E(j|U) \frac{b/n - p}{1 - w} \end{aligned} \tag{13}$$

where $m(j|U)$ is the probability that there are *j* other cooperator-punishers given that the focal is an unconditional defector, and $E(j|U)$ is the expected value of *j* conditioned on the focal individual being *U*. An analogous calculation shows that

$$W(P) = \frac{(b/n)(E(j|P) + 1) - c - (n - 1 - E(j|P))k}{1 - w} \tag{14}$$

When groups are formed at random $E(j|P) = E(j|U) = (n - 1)q$ where q is the frequency of P in the population just before groups are formed. To determine when U is an ESS, let $q \rightarrow 0$ and determine when $W(U) > W(P)$. To determine when P is an ESS, let $q \rightarrow 1$ and determine when $W(U) < W(P)$. When groups are formed assortatively and P is rare $E(j|P) = (n - 1)r$ and $E(j|U) = 0$. Combining these expressions yields the condition for P to increase when rare (16).

It follows from these expressions for the fitness of U and P that (1) unconditional defection is always an ESS, and (2) P is an ESS only if:

$$c - b/n < (n - 1)p \quad (15)$$

The left-hand side of (15) is the per period cost to an individual of cooperating, and the right-hand side is the per period cost of being punished by $n - 1$ individuals.

Superficially these properties seem analogous to the competition between always-defect and tit-for-tat in the two-person model. Always-defect is always an ESS, tit-for-tat is an ESS only under certain conditions. However, notice that (15) does not depend on the parameter w which measures the average number of interactions. Thus, if (15) is satisfied, P is stable even if individuals interact only once! In contrast, tit-for-tat is stable against always-defect only if w is large enough that the long-run benefit of reciprocal interaction is greater than the short-term benefit of cheating. Tit-for-tat is never stable if individuals interact only once.

The qualitative difference between the two models is made clearer if we consider the effect of assortative group formation. In the two-person case, assortative group formation makes it easier for tit-for-tat to increase when rare, and if w is near one, even a small amount of assortment is sufficient. In the present model, the punishing strategy, P , can increase when rare if

$$(b/n)[r(n - 1) + 1] - c > (n - 1)(1 - r)k \quad (16)$$

inclusive fitness punishment

The left-hand side gives the inclusive fitness advantage of cooperators relative to defectors. If P individuals are sufficiently likely to interact with other P individuals ($r \rightarrow 1$), then P can increase in frequency even when it is rare in the population because P individuals benefit from the cooperation of other P individuals in their groups. The right-hand side gives the effect of punishment on the fitness of P individuals. Notice that this term is always positive. This means that cooperation supported by punishment is *harder* to get started in a population than unconditional cooperation.

Why are these two models so different? In models without retribution, reciprocal strategies such as tit-for-tat are favored because they lead to assortative interaction of cooperators (Michod and Sanderson 1985). Even if individuals are paired at random, the fact that tit-for-tat individuals convert

to defection if they experience acts of defection from others, causes a non-random distribution of cooperative behavior: tit-for-tat individuals are more likely to receive the benefits of cooperation than are always-defect individuals. In contrast, in the present model, punishment has no effect on who receives the benefits of cooperative behavior. P individuals continue to cooperate while they punish, and U individuals do not respond to punishment by cooperating—they keep defecting. Models of reciprocity without punishment suggest that the strategy of punishing defectors by withdrawing cooperation is unlikely to work in large groups (Joshi 1987; Boyd and Richerson 1988). However, it is not unreasonable to imagine that a kind of conditional defector might respond to punishment by cooperating much as tit-for-tat responds to cooperation with more cooperation.

A.2. Should Defectors Respond to Punishment?

Should defecting individuals respond to punishment by cooperating? To address this question, we consider the conditions which R_1 can invade a population in which the strategy U is common. We further assume that groups are formed at random.

Unfortunately, the answer to this question does not depend on the fitness consequences of alternative behaviors alone. It also depends on what kinds of punishing strategies are maintained in the population by nonadaptive processes like mutation and nonheritable environmental variation. In a population in which only U and R_1 are present (and every individual accurately follows its strategy) U and R_1 will have the same expected fitness. Both will defect forever, and never be punished because no punishing strategies are present. The strategies U and R_1 will have different expected fitnesses only if there are punishing strategies present in the population. If U is common, however, the expected fitness of any rare punishing strategy must be less than the expected fitness of U . This means that any punishing strategies present in the population must be maintained by nonadaptive processes like errors or mutation. R_1 may or may not be able to invade depending on the mix of punishing strategies maintained by such forces.

We conjecture that the most plausible source of nonadaptive variation is mistakes about the behavioral context. Modelers typically assume that there is a single behavioral context, with given costs and benefits, and an unambiguous set of behavioral strategies. However, in the real world, there are many behavioral contexts each with its own appropriate strategy. Before deciding how to behave individuals must categorize a particular situation as belonging to one context or another. It seems plausible that individuals sometimes miscategorize situations in which punishment is not favored, and thus mistakenly punish others. Suppose, for example, selection favors individual retaliation if others damage personal property. Then individuals might sometimes punish others who damage commonly held property because they mistakenly miscategorize the behavior.

To prove that R_1 may or may not be about to invade U , consider the second punishing strategy:

Timid punishers (T_1). Always cooperate. Punish each defector the first time it defects, but only the first time.

Suppose that both U and R occasionally mistakenly play one of the punishing strategies. This could occur because individuals mistake the behavioral context for one in which they would normally punish. The relative fitness of U and R_1 depends on which of these two punishing strategies is present. Suppose that individuals occasionally play T_1 by mistake. R_1 can invade if a focal R_1 individual has higher fitness than a focal U individual in groups with one T_1 individual among the other $n - 1$. In such groups

$$W(U) = \frac{b/n}{1 - w} - p \tag{17}$$

$$W(R_1) = \frac{b/n}{1 - w} - p + \frac{w}{1 - w} (b/n - c) \tag{18}$$

and thus U is always favored if cooperation is costly. In contrast, when P is present as a result of errors, the fitnesses of the two types are

$$W(U) = b/n - p + \frac{w}{1 - w} (b/n - p) \tag{19}$$

$$W(R_1) = b/n - p + \frac{w}{1 - w} (2b/n - c) \tag{20}$$

and thus R_1 is favored whenever the costs of punishment exceed the cost of cooperating.

We think that this result is likely to be quite general. Consider a strategy that begins cooperating only after being punished some number of times. Such a strategy will have higher fitness than an unresponsive strategy only if the punishing strategies present in the population continue to punish on subsequent turns. If they do not, the unresponsive strategy gets the benefit without paying the cost. When should punishing strategies give up? The answer to this question depends on whether the defecting strategies will respond. If defecting strategies are unresponsive, costly punishment provides no benefits.

A.3. Equilibria when R_1 and P Compete

Let j be the number of P individuals among the other $n - 1$ individuals in a group. Then the expected fitnesses of the two types are:

$$\begin{aligned} W(P) = & (1 - e)[(b/n)(E(j|P) + 1) - c] - k[n - 1 - (1 - e)E(j|P)] \\ & - epE(j|p) + \frac{w}{1 - w} [(1 - e)(b - c) - ek(n - 1) - epE(j|P)] \end{aligned} \tag{21}$$

$$\begin{aligned}
 W(R_1) = & [(b/n)(1 - e) - p]E(j|R_1) \\
 & + \frac{w}{1 - w} [(1 - e)(b - c) - epE(j|R_1)] \\
 & - \frac{w\Pr(j = 0|R_1)}{1 - w} (1 - e)(b - c) \quad (22)
 \end{aligned}$$

where $\Pr(j = 0|R_1)$ is the probability that an R_1 individual finds itself in a group with exactly zero P individuals.

When groups are formed at random $E(j|P) = E(j|R_1) = (n - 1)q$ and $\Pr(j = 0|R_1) = (1 - q)^{n-1}$ where q is the frequency of P . Making these substitutions leads to the following condition for R_1 to increase:

$$\begin{aligned}
 (k + p)(n - 1)(1 - q) - \frac{w}{1 - w} (1 - q)^{n-1} (b - c) \\
 - b/n + c - p(n - 1) + \frac{ek(n - 1)}{(1 - w)(1 - e)} > 0 \quad (23)
 \end{aligned}$$

The condition for R_1 to be an ESS (7) is derived by setting $q = 0$ in (23). The condition for P to be an ESS (6) is derived by setting $q = 1$ in (23).

To derive the necessary conditions for a stable internal equilibrium, first notice that the left-hand side of (23) is a concave function with, at most, a single internal maximum. Thus, if neither R_1 or P is an ESS, then there is a single internal equilibrium point. If R_1 is not an ESS but P is, then there are two internal equilibria, one stable and the other unstable, if and only if the value of the left-hand side at that maximum is greater than zero. The value of q that maximizes the left-hand side of (23) can be found by differentiation. Substituting this back into (23) yields the following necessary condition for the existence of two internal equilibria:

$$\begin{aligned}
 \left(\frac{(1 - w)(k + p)}{w(b - c)} \right)^{1/(n-2)} (n - 2)(k + p) > p(n - 1) \\
 - c + b/n - \frac{ek(n - 1)}{(1 - w)(1 - e)} \quad (24)
 \end{aligned}$$

If this condition is not satisfied then P is the only ESS.

To derive the condition for R_1 to increase when groups are formed assortatively let $E(j|P) = (n - 1)(r + (1 - r)q)$ and $E(j|R_1) = (n - 1)(1 - r)q$ and proceed in the same way as above.

A.4. Equilibria when R_1 , E , and P Compete

Let i and j be the numbers of E and P individuals among the other $n - 1$ individuals. Then

$$\begin{aligned}
 W(P) = & (1 - e)[(b/n)(E(i|P) + E(j|P) + 1) - c] \\
 & - k[n - 1 - (1 - e)(E(i|P) + E(j|P))] - epE(j|P) \quad (25) \\
 & + \frac{w}{1 - w} \times [(1 - e)(b - c) - ek(n - 1) - epE(j|P)]
 \end{aligned}$$

$$\begin{aligned}
 W(R_1) &= (b/n)(1 - e)(E(i|R_1) + E(j|R_1)) - pE(j|R_1) \\
 &\quad + \frac{w}{1 - w} \Pr(j > 0|R_1)[(1 - e)(b - c) - epE(j|R_1 \& j > 0)] \\
 &\quad + \frac{w}{1 - w} (1 - e)(b/n)\Pr(j = 0|R_1)E(i|R_1 \& j = 0)
 \end{aligned} \tag{26}$$

$$\begin{aligned}
 W(E) &= (1 - e)[(b/n)(E(i|E) + E(j|E) + 1) - c] - epE(j|E) \\
 &\quad + \frac{w}{1 - w} \Pr(j > 0|E)[(1 - e)(b - c) - epE(j|E \& j > 0)] \\
 &\quad + \frac{w}{1 - w} (1 - e) \Pr(j = 0|E)[(b/n)(E(i|E \& j = 0) + 1) - c]
 \end{aligned} \tag{27}$$

Assume that groups are formed at random so that $E(j|E) = E(j|P) = E(j|R_1) = (n - 1)q_P$, $E(i|E) = E(i|P) = E(i|R_1) = (n - 1)q_E$, $\Pr(j = 0|E) = \Pr(j = 0|R_1) = (1 - q_P)^{(n-1)}$, and $E(i|R_1 \& j = 0) = E(i|E \& j = 0) = (n - 1)(q_E/(1 - q_P))$ where q_E and q_P are the frequencies of E and P . When $q_E = 1$, $W(E) < W(R_1)$ and when $q_P = 1$, $W(P) < W(E)$.

First, we derive conditions for the existence of an internal equilibrium, and show that if such an equilibrium exists, it is unique.

It is useful to define the following functions which give the difference in fitness between each pair of strategies as a function of q_P , and q_E :

$$d_{PE}(q_P, q_E) = W(P) - W(E) \tag{28}$$

$$d_{RE}(q_P, q_E) = W(R_1) - W(E) \tag{29}$$

$$d_{PR}(q_P, q_E) = W(P) - W(R_1) \tag{30}$$

Using Eqs. (25), (26), and (27) and the assumption of random group formation yields the following expression for d_{RE} :

$$\begin{aligned}
 d_{RE}(q_P, q_E) &= -p(1 - e)(n - 1)q_P \\
 &\quad + (1 - e)(c - b/n) \left(1 + \frac{w}{1 - w} (1 - q_P)^{n-1} \right)
 \end{aligned} \tag{31}$$

Notice that the relative fitness of R_1 and E depends only on q_P . Further, note that (1) $d_{RE}(0, q_E) > 0$, (2) $d_{RE}(1, q_E) < 0$ as long as $c - b/n < (n - 1)p$ which is true by assumption, and (3) d_{RE} is a monotonically decreasing function of q_P . Thus, the value of q_P at equilibrium is unique and can be found by finding the root of $d_{RE} = 0$ as shown in Figure 2. Let this value of q_P be \hat{q}_P .

Once again using Eqs. (25), (26), and (27) and the assumption of random group formation yields the following expression for d_{PE} :

$$\begin{aligned}
 d_{PE}(q_P, q_E) &= \frac{-ke(n - 1)}{1 - w} + k(1 - e)(n - 1)(1 - q_E - q_P) \\
 &\quad + \frac{wb(n - 1)(1 - e)(1 - q_P)^{n-2}(1 - q_P - q_E)}{n(1 - w)}
 \end{aligned}$$

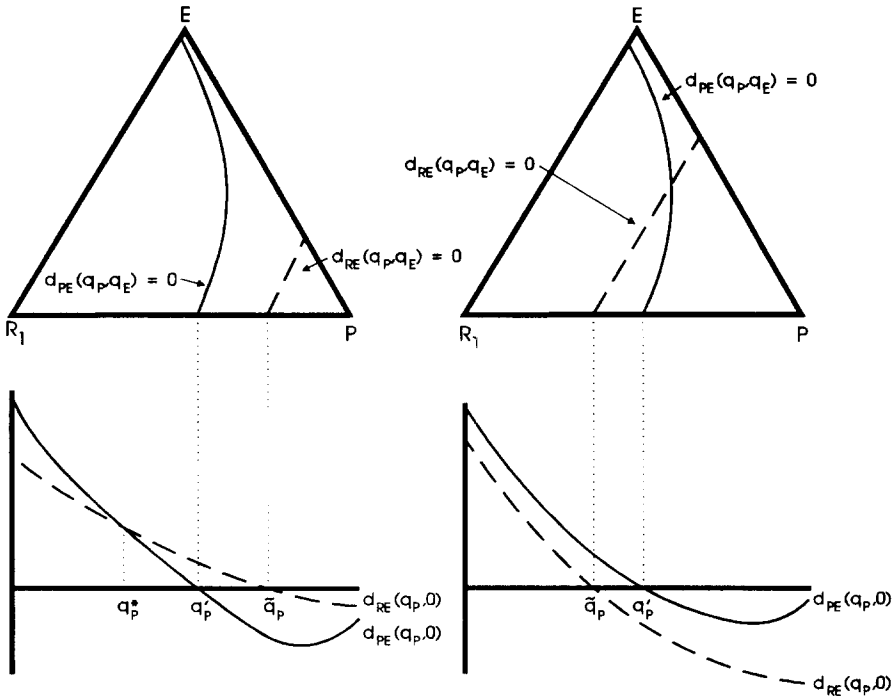


FIGURE 2. This figure illustrates the logic of the proofs given in this section. The left-hand pair of figures represent a situation in which there is a single polymorphic equilibrium on the $R_1 - P$ boundary. The lower figure shows $d_{RE}(q_P, 0)$ and $d_{PE}(q_P, 0)$. These curves intersect only once since there is a single polymorphic equilibrium. Thus, we know that $q'_P < \bar{q}_P$. The upper figure shows how the forms of $d_{RE}(q_P, q_E)$ and $d_{PE}(q_P, q_E)$ guarantee that there is no internal equilibrium in this case. The right-hand pair of figures represent the situation in which there is no polymorphic equilibrium on the $R_1 - P$ boundary because P increases for all values of $q_P < 1$.

Assume that q_E is fixed at some value. Then

$$d_{PE}(1 - q_E, q_E) = \frac{-ke(n - 1)}{1 - w} < 0$$

and

$$d_{PE}(0, q_E) = \frac{-(n - 1)ke}{1 - w} + (1 - q_E)(1 - e)(n - 1) \left(\frac{wb}{n(1 - w)} - k \right)$$

Thus, $d_{PE}(0, q_E) > 0$ if

$$q_E < 1 - \frac{ke}{(1 - e)(w(b/n) - k(1 - w))} \tag{32}$$

and

$$w(b/n) - k(1 - w) > 0 \tag{33}$$

Otherwise $d_{PE} \leq 0$. Differentiating shows that d_{PE} is a convex function of q_P . Thus if (32) and (33) are satisfied, $d_{PE}(q_P, q_E) = 0$ has a unique root for each q_P as illustrated in Figure 2. Let this root be $q'_P(q_E)$. Increasing q_E leads to a decrease in $q'_P(q_E)$. Thus there is an internal equilibrium value if and only if $q'_P(0) > \bar{q}_P$, and if it exists, such an equilibrium is unique. This result is shown graphically in Figure 2.

From the above we know that $d_{RE}(q_P)$ is monotonically decreasing and has one root in the interval (0,1) whenever R_1 is potentially present, and that $d_{PE}(q_P)$ has at most one root and is monotonically decreasing in the interval that contains the root.

Next, we show that if there is no stable polymorphic equilibrium on the $P - R_1$ boundary in the absence of E , then there is an internal equilibrium. If there is no stable equilibrium on the boundary in the absence of E , it follows from the results of the previous section that

$$d_{PR}(q_P, 0) > 0$$

for all q_P . Next, note that

$$d_{PR}(q_P, q_E) = d_{PE}(q_P, q_E) - d_{RE}(q_P, E) \tag{34}$$

Thus, there is an internal equilibrium since $d_{PE}(q_P, 0) > d_{RE}(q_P, 0)$ for all values of q_P . This situation is shown in the right-hand pair of figures in Figure 2.

Next, we show that if there is a stable, polymorphic boundary equilibrium such that $q_E = 0$ and $W(R_1) > W(P)$ for $q_P = 1$, then there is no internal equilibrium. Let q_P^* be the frequency of P at a polymorphic equilibrium on the $P - R_1$ boundary. Then $d_{PR}(q_P^*, 0) = 0$ which implies that $d_{PE}(q_P^*, 0) = d_{RE}(q_P^*, 0)$. The fact that the equilibrium is stable in the absence of E implies that $\partial d_{PR}(q_P, 0) / \partial q_P < 0$ at q_P^* . Since $d_{PR}(1, 0) < 0$, it follows that $d_{PE}(q_P, 0) < d_{RE}(q_P, 0)$ for $q_P^* < q_P < 1$. But this means that $q'_P(0) < \bar{q}_P$, and, therefore, there is no internal equilibrium as shown in the left-hand pair of figures in Figure 2.

It is important to note that there may be no internal equilibrium even if $W(R_1) < W(P)$. When this is the case there is a second, unstable internal equilibrium on the $R_1 - P$ boundary. Anytime that $d_{PE} = d_{RE} < 0$ at this equilibrium, there will be no internal equilibrium, and numerical studies suggest that this is what actually occurs at the vast majority of parameter combinations.

A.5. M is an ESS against P and E

Assume that M is common. When groups are formed at random, M can resist invasion by rare R_1 individuals if the average payoff of M in groups with $n - 1$ other M individuals, $V(M|n - 1)$ is greater than the average payoff of R_1 in groups in which the other $n - 1$ individuals are M , $V(R_1|n - 1)$:

$$V(M|n - 1) = \frac{1}{1 - w} ((b - c)(1 - e) - e(n - 1)(k + p)) \quad (35)$$

$$V(R_1|n - 1) = (n - 1)(b/n)(1 - e) - (n - 1)p \quad (36)$$

$$+ \frac{w}{1 - w} [(1 - e)(b - c) - p(n - 1)(1 - (1 - e)^n)]$$

Substituting these expressions and simplifying yields (9). Similarly, the expected fitness of an E individual in a group of $n - 1$ M individuals, $V(E|n - 1)$, is:

$$V(E|n - 1) = (1 - e)(b - c) - e(n - 1)p + \frac{w}{1 - w}$$

$$\times [(1 - e)(b - c) - e(n - 1)p - p(n - 1)(1 - (1 - e)^n)] \quad (37)$$

This expression is used to determine when $V(M|n - 1) > V(E|n - 1)$ yields (10).