WHEN IN ROME, DO AS THE ROMANS DO: THE COEVOLUTION OF ALTRUISTIC PUNISHMENT, CONFORMIST LEARNING, AND COOPERATION

RICARDO ANDRÉS GUZMÁN
INSTITUTO DE ECONOMÍA
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

CARLOS RODRÍGUEZ-SICKERT
INSTITUTO DE SOCIOLOGÍA
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

ROBERT ROWTHORN *
FACULTY OF ECONOMICS
UNIVERSITY OF CAMBRIDGE

WORD COUNT: 3838.

* Corresponding author. Faculty of Economics, Sidgwick Avenue, Cambridge, CB3 9DD, UK. Tel: 44-(0) 1223-335230. E-mail: rer3@econ.cam.ac.uk.

ABSTRACT

We model the coevolution of social learning rules and behavioral strategies in the context of a cooperative dilemma, a situation in which individuals must decide whether or not to subordinate their own interests to those of the group. There are two learning rules in our model, *conformism* and *payoff-dependent imitation,* which evolve by natural selection; and three behavioral strategies, *cooperate*, *defect*, and *cooperate and punish defectors*, which evolve under the influence of the prevailing learning rules. Group and individual level selective pressures drive evolution.

We also simulate our model for conditions that approximate those in which early hominids lived. Contrary to previous claims, we find that conformism can evolve when the only problem individuals face is a cooperative dilemma. Furthermore, the presence of conformists dramatically increases the group size for which cooperation can be sustained. The results of our model are robust: they hold even when migration rates are high, and when conflict among groups is infrequent.

1.0 INTRODUCTION

We are a cooperative species. Experimental evidence and field data show that humans often sacrifice resources in order to benefit non-relatives, even when those who benefit are not expected to return the favor (Gintis et al. 2003). People sometimes use “altruistic punishment” to enforce cooperation, whereby they pay a cost in order to punish non-cooperators whom they will never meet again (Fehr & Gaechter, 2000, 2002; Ostrom, Walker & Gardner, 1992). The combination of unrequited cooperation between non-relatives and altruistic punishment is known as “strong reciprocity” (Gintis 2000). Both of these components of strong reciprocity pose a puzzle for the standard evolutionary theories of cooperation: kin-selection (Hamilton, 1964) and reciprocal altruism (Trivers, 1971; Axelrod & Hamilton, 1981).

Some authors argue that human cooperation may be explained by the selection of cultural traits at the group level (Bowles et al., 2003; Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981; Sober & Wilson 1994). Assuming that cooperative groups outcompete less cooperative ones in the struggle for survival, then it may be possible for group level selective pressure to outweigh the maladaptive nature of altruism at the individual level. For this to occur, either noncooperative individuals must invade cooperative groups infrequently or else the amount of intergroup conflict must be very high.

Analytical models suggest that two factors play a crucial role in the emergence of cooperation: altruistic punishment and conformism (i.e., the tendency of individuals to imitate the most common form of behavior; see Boyd & Richerson, 1985, and Henrich & Boyd, 1998). Gintis (2000) proves that, when a group faces the threat of extinction, a small number of altruistic punishers may induce selfish individuals to behave cooperatively. Henrich and Boyd (2001) show that an arbitrarily small amount of conformism may permit altruistic punishment to persist. Boyd et al. (2003) report simulations that mimic the environment in which early hominids lived. They show that altruistic punishment enhances cooperative behavior when social learning takes the form of payoff-dependent imitation (i.e., when individuals imitate the most successful forms of behavior). However, this mixture of group selection and punishment cannot sustain cooperation in large groups if the migration rate between groups is high and conflict between groups is low.

Boyd & Richerson (2005) argue that cultural group selection is especially strong in human populations due to the fact that variation amongst human groups is maintained by an unusual combination of strong reciprocity and conformist social learning. Following their lead, this paper uses a group selection approach to explore the coevolution of social learning rules and behavioral strategies in the context a “cooperative dilemma”. By cooperative dilemma we mean a situation in which an individual must choose whether or not to behave cooperatively, and benefit the group, or uncooperatively, and benefit himself. In our model, there are two social learning rules, conformism and payoff-dependent imitation, which evolve by natural selection; and three behavioral strategies, cooperate, defect, and cooperate and punish defectors, which evolve under the influence of the prevailing learning rules.

To the extent that our analysis is concerned with competing learning rules, it relates to the literature on endogenous learning. There is, however, one important difference. This literature is primarily concerned with social and individual learning as alternative ways to acquire information about the natural environment. Within such a framework, Boyd and Richerson (1985) demonstrate how the balance between social and individual learning depends on the accuracy of learning and the variability of the environment. Feldman et al. (1996) show that social learning can evolve if there is a fixed fitness cost to learning errors, whilst Henrich and Boyd (1998) show that social learning can evolve as long as the environment is not too variable. Using an experimental approach, Efferson et al. (2006) explore the choice between alternative forms of social learning. They find that this choice depends on the type of information available to the individual. Conformism is preferred when the individual has information about the frequencies of different kinds of behavior, whereas payoff-dependent imitation is preferred when the individual has information on the highest or lowest payoffs. However, the authors do not examine how individuals will choose between or combine the two forms of social learning when both kinds of information are available. Nor do they address how these alternative forms of social learning coevolve in an environment in which individual decisions involve strategic interaction with others.

The aims of this paper are as follows: first, to determine if conformist transmission can evolve within the context of a cooperative dilemma, and secondly, to explore the impact of conformism on cooperation. Contrary to previous claims (Henrich, 2004; Henrich & Boyd, 2001), we find that conformism can indeed evolve when the only problem individuals face is a cooperative dilemma. Furthermore, the presence of conformists dramatically increases the group size for which cooperation can be sustained.

2.0 MODEL

We shall now develop a model in which evolution determines both the learning rules which individuals adopt and the behavioral strategies that they follow. The learning rules evolve at the biological level and the actions chosen by individuals at any time are based on these rules. Our model builds on the work by Boyd et al. (2003), but departs from it by allowing conformist learning, and by making learning rules endogenous.

There are G groups, each of which has N members. Every year the members of a particular group play a societal game. This game is divided into five phases: hunting, war, learning, reproduction, and migration.

During the hunting phase, each individual follows one of three possible behavioral strategies: *cooperate* (C), *defect* (D), and *cooperate and punish defectors* (P). Denote by $\sigma(s) \in [0,1]$ the fraction of the group that chooses strategy $s \in \{\mathrm{C,D,P}\}$. Someone who intends to cooperate may erroneously defect with probability *e*, so the ex post fraction of defectors will be $\sigma(\mathrm{D}) + e[\sigma(\mathrm{C}) + \sigma(\mathrm{P})]$. We assume that punishers who unintentionally fail to cooperate continue to punish. Let $\pi(\mathrm{s},\sigma)$ be the payoff of an individual who follows strategy *s* when the distribution of types in his group is $\sigma(\cdot)$. We define $\pi(\mathrm{s},\sigma)$ as follows:

$$\pi(\mathrm{D},\sigma) = -p\sigma(P) + z,$$

$$\pi(\mathrm{C},\sigma) = -(1-e)c - ep\sigma(P) + z,$$

$$\pi(\mathrm{P},\sigma) = -(1-e)c - ep\sigma(P) + k\left\{\sigma(\mathrm{D}) + e[\sigma(\mathrm{C}) + \sigma(\mathrm{P})]\right\} + z,$$

where $z = \max\left\{(1-e)c + k,\ p\right\}$. The positive constants *c*, *k*, and *p* capture the costs of cooperating, punishing, and being punished. The inclusion of z in the payoff function guarantees that payoffs will always be positive.

In each period, all groups pair at random. Every pair of groups makes war with probability ε. Only one group in each warring pair survives. Suppose groups *g* and g' enter into conflict. Group *g* will survive with probability $\frac{1}{2}[1 + \sigma'(\mathrm{D}) - \sigma(\mathrm{D})]$, where $\sigma(\mathrm{D})$ is the fraction of defectors in group *g* and $\sigma'(\mathrm{D})$ is the fraction of defectors in group g'. The surviving group fissions and repopulates the site of the extinct group in the following fashion. First, every individual in the surviving group produces a clone of himself. Second, individuals and their clones intermingle and are randomly reassigned to the site of the surviving group or to the site of the extinct one, creating two new groups of size N.

Individuals come in two genetic types which differ according to their learning rules: *payoff-dependent imitators* and *conformists*. Every individual uses the same learning rule throughout his life. The evolution of learning rules is governed by natural selection. Individuals die with probability q. A dead individual is replaced by a son of some member of his group. The probability that a dead individual will be replaced by a son of i is given by

$$\frac{\pi_i}{\sum_{j=1}^{N} \pi_j}.$$

The newborn son will be an exact replica of his father. Thus he will have the same genetically-determined learning rule as his father, and will start life with his father's behavioral strategy. With probability ν the son will immediately mutate and adopt a random type and strategy.

During the learning phase, each payoff-dependent imitator meets a role model from his group. Let s be the strategy used by the imitator, and let s' be the strategy used by the role model. The probability that the imitator will adopt the strategy of the role model is

$$\frac{\pi(\mathrm{s}',\sigma)}{\pi(\mathrm{s},\sigma)+\pi(\mathrm{s}',\sigma)}.$$

After meeting the role model, the imitator may still decide to innovate and switch to a randomly chosen strategy with probability μ. Conformists do not innovate and just play their group's modal strategy s^*, where

$$s^* = \underset{s\in\{\mathrm{C,D,P}\}}{\arg\max}\,\sigma(s).$$

In order to introduce a migration-like force, we assume that each individual meets a stranger from another group with probability m. Let π be the last payoff of the individual, and let π' be the last payoff of the stranger. The individual will be replaced by a clone of the stranger with the following probability:

$$\frac{\pi'}{\pi+\pi'}.$$

Finally, we assume that at the beginning of time there are $G-1$ groups of payoff-dependent imitators who all use the behavioral strategy *defect*, and one group of conformists that all use the behavioral strategy *cooperate and punish*.

3.0 RESULTS

3.1 Baseline Scenario

Following Boyd et al. (2003), we simulate the model of the previous section for conditions that approximate those in which early hominids lived. Each simulation spans 2000 years of model time. Baseline parameters are given in Table 1. Our model introduces two new parameters which are absent in Boyd et al. (2003): the death rate and the mutation rate. We set the death rate at $q = 0.1$, which implies a reproductive life of ten years. The mutation rate is assumed to be one order of magnitude lower than the innovation rate.

TABLE 1 ABOUT HERE

Figure 1 presents the results of our model for the baseline parameters (the solid square lines), along with the results of three other models: one in which punishment is allowed to evolve, but not conformism (the empty square lines); one in which conformism is allowed to evolve, but not punishment (the empty triangle lines); and one in which neither punishment or conformism are allowed to evolve (the empty circle lines). The case with punishment but no conformism corresponds to the model in Boyd et al. (2003). The figure plots averages of frequencies over the final 1000 years of 20 simulations.

FIGURE 1 ABOUT HERE

To understand these results, it is convenient to analyze first the dynamics of the societal game for a group that lives in isolation, subject to no mutation, no migration and no war, and is comprised entirely of payoff-dependent imitators. In such a group there are no conformists. Under these conditions, the societal game will have two kinds of equilibria: one composed entirely of defectors and one with no defectors at all. In the latter type of equilibrium the condition $\sigma(\mathrm{P}) > a$ must be satisfied, where $a = c^{-1}p$ is the fraction of punishers such that cooperation and defection yield the same payoff. If this condition is not satisfied, then defectors can invade and eventually take over. Consider an equilibrium in which the fraction of punishers is equal to $\sigma_0(\mathrm{P}) > a$. If someone innovates and becomes a defector he will be driven out by punishers. However, this will require a finite period of time during which punishers will incur the extra cost of policing defectors and hence will be less fit than cooperators. During the transition period to the new equilibrium, the ratio of punishers to cooperators will therefore decrease. When the population restabilizes after the innovator has been driven out, this will be in a new equilibrium with $\sigma_1(\mathrm{P}) < \sigma_0(\mathrm{P})$. Eventually, as a result of successive innovations $1, 2, ..., j$, there will come a point where $\sigma_j(\mathrm{P}) < a$, and from then onwards defectors will prosper and take over. In consequence, the only stable equilibrium of the societal game is the one in which everybody defects.

Now consider the case with migration and war between groups. As before, assume there is no mutation and that all individuals are payoff-dependent imitators, but this time suppose that no peer-to-peer sanctioning is available. In this scenario there are no conformists and no punishers, and the only behavioral strategies available are cooperation and defection. The long run values of cooperation in this scenario are depicted by the circle line in Fig. 1A. In small groups, moderate levels of cooperation are achieved by group selection alone. When two groups enter into conflict, the one with more cooperators is more likely to win and repopulate the site of the other. In this way cooperation will spread between groups. For group selection to produce high levels of cooperation, however, there must be enough inter-group variation to contain the proliferation of free riders in the years between wars. The extent of inter-group variation between groups depends on the balance between the homogenizing effect of migration between groups and the diversity arising from innovation and fissioning within groups. When group size is small, innovation and fissioning can generate enough inter-group diversity to offset the homogenizing effect of migration. In larger groups, however, the law of large numbers comes into play so that innovation and fissioning produce less variation, with the result that diversity arising from this source is no longer sufficient to offset migration and preserve the inter-group variation required to sustain cooperation.

As can be observed from the empty square line in Fig. 1A, the addition of punishers ameliorates the negative effect of group size. With a high proportion of punishers the first order free-riding problem —the irruption of defectors— is solved. Although a second order free-riding problem emerges —cooperators failing to punish defectors— this problem is less serious: whereas the payoff advantage of defectors over cooperators does not depend on the frequency of defection, the payoff advantage of cooperators over punishers decreases as defectors become rare.

Even when peer-to-peer sanctioning is available, random variation is still needed to sustain high levels of cooperation. To see why, suppose that all groups are in a cooperative equilibrium without defectors, and let $\sigma_0(\mathrm{P}) > a$ be the fraction of punishers in the overall population. Also suppose the homogenizing effect of migration has operated long enough so that the share of punishers is the same in all groups. If groups are large, the law of large numbers entails that the same fraction of every group will innovate and start defecting. Punishers will drive them out, but during the transition period the share of punishers in all groups will decrease to $\sigma_1(\mathrm{P}) < \sigma_0(\mathrm{P})$. Since this process will generate no inter-group variation, when war happens, group selection will have nothing to select. As in the isolated group case, the share of punishers will eventually fall to the point where innovating defectors can successively invade and cooperation will break down. Even if groups are too small for the law of large number to operate effectively, migration may still reduce inter-group differences, thereby undermining cooperation.

The triangle lines in Fig. 1 show that conformism and cooperation coevolve in our model even when no peer-to-peer sanctioning is available. The mere presence of conformists raises the frequency of cooperation in comparison to the no conformism and no punishment scenario, and makes cooperative behavior possible in much larger group. To see why, imagine a group of cooperative conformists which is colonized by a foreign defector. Since cooperation will still be the modal behavior of the group, conformists won't react to the payoff advantage of the newcomer; they will just keep on cooperating. In this example, conformism acts as a force against the homogenization of groups, reinforcing the effect of innovation and fissioning.

The solid square lines in Fig. 1 show what happens in our baseline model which contains both conformism and punishment. In this model, cooperation achieves a very high level and is an increasing function of group size. The combination of conformism and punishment encourages cooperation in several ways. Consider a group in which punishment is the modal strategy. Over the course of time, such a group will absorb a stream of "newcomers" in the form of immigrants and newborns, together with existing members who modify their behavioral strategies by innovating. If the newcomer is a conformist, he will adopt the modal strategy and become a punisher who reinforces group cooperation. However, if he is a pay-off dependent imitator then, according to his past history, he may adopt another course of action. He may defect, in which case he will directly weaken the group, or else he may simply cooperate, but fail to punish defectors, thereby encouraging defection by others. In a group where punishment is the modal strategy, conformist newcomers will immediately start to punish, whereas pay-off dependent imitators may choose some other form of behavior. In such a group, conformism stabilizes punishment and reinforces cooperation.

Conformism also has another positive effect on co-operation. Consider a conformist-defector who migrates into a population consisting mainly of punishers. On arriving in his new group he will immediately switch to the modal behavior, so that punishers will have no reason to punish him. This benefits both the group and the newcomer, who avoids being punished. That conformism is convenient for immigrants is no new discovery. On the contrary, it was long ago captured by conventional wisdom: when in Rome, do as the Romans do.

In sum: conformism preserves between group variation and stabilizes punishment; punishment protects groups from the spread of defection, and may also give conformists a fitness advantage over payoff-dependent imitators. For these reasons, punishment, conformism, and cooperation coevolve in our model, and cooperation is high even in large groups.

Perhaps the most puzzling of our findings is the fact that cooperation increases with group size, instead of decreasing, as one might expect. Fig. 2 shows the frequencies of the three behavioral strategies in the baseline model, for different group sizes. As groups become larger, so does the share of punishers, until almost everyone is a punisher. This may be for the following reason. When groups are small, innovation and fissioning are likely to move groups out of the equilibrium favored by group selection: the one where everybody punishes. In addition to its impact on the number of punishers, such “noise” may also turn conformism into a drawback, since out of equilibrium the modal strategy of the group need not coincide with the strategy which is optimal for the group as a whole. In large groups, the law of large numbers dissipates the effects of random variation, and the mix of punishment and conformism displays its full potential.

FIGURE 2 ABOUT HERE

3.2 Sensitivity analysis

Fig. 2 shows how our model responds to a low conflict rate ($\mu = 0.0075$) and to a high migration rate ($m = 0.05$). As can be observed, the combination of conformism and altruistic punishment is able to sustain high levels of cooperation for all group sizes under these very adverse conditions. Note cooperation falls slightly at intermediate group sizes. This can be explained as follows. When groups are small, random variation keeps cooperation high, even though the variation weakens the effect of conformism and altruistic punishment. At intermediate group sizes, the law of large number dilutes random variation enough to dampen group selection, but not enough for conformism and altruistic punishment to fully counter the homogenizing force of migration. Finally, when groups are large, random variation vanishes completely, conformism and punishment thrive, and so does cooperation.

FIGURE 3 ABOUT HERE

4.0 DISCUSSION

Contrary to previous claims, we have shown that conformism can evolve when the only problem individuals face is a cooperative dilemma. We have also shown that conformism and altruistic punishment coevolve, allowing groups of greater size to sustain cooperation. This occurs because conformism preserves between group variation and stabilizes punishment, and because punishment protects groups from the spread of defection and gives conformists a fitness advantage over payoff-dependent imitators.

REFERENCES

Axelrod, R. & W. D. Hamilton (1981). The evolution of cooperation. *Science*, 211, 1390–1396.

Bowles, S., J. K. Choi & A. Hopfensitz (2003). The coevolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, 223, 135–147.

Boyd, R., H. Gintis, S. Bowles & P. Richerson (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Science*, 100, 3631–3635.

Boyd, R. & P. J. Richerson (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R. & P. J. Richerson (2005). *Solving the puzzle of human cooperation*. In S. Levinson (Ed.), Evolution and Culture. Cambridge, MA: MIT Press.

Cavalli-Sforza, L. L. & M. W. Feldman, (1981). *Cultural Transmission and Evolution*. Princeton, NJ: Princeton University Press.

Efferson, C., Lalive, R., P. J. Richerson, R. McElreath & M. Lubell (2006). Social Learning Experiments. Unpublished manuscript.

Fehr, E. & S. Gaechter (2000). Cooperation and punishment in public good games. *American Economic Review*, 90, 980–994.

Fehr, E. & S. Gaechter (2002). Altruistic punishment in humans. *Nature*, 415, 137–140.

Feldman, M., K. Aoki & J. Kumm (1996). Individual versus social learning: Evolutionary analysis in a fluctuating environment. *Anthropological Science*, 104, 209–232.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology,* 206, 169–179.

Gintis, H., S. Bowles, R. Boyd & E. Fehr (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24, 153–172.

Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology*, 37, 1--52.

Henrich, J. (2004). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior and Organization*, 53, 3–35.

Henrich, J. & R. Boyd (1998). The evolution of conformist transmission and between-group differences. *Evolution and Human Behavior*, 19, 215–242.

Henrich, J. & R. Boyd (2001). Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. *Journal of Theoretical Biology*, 208, 79–89.

Ostrom, E., J. Walker & R. Gardner (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review*, 86, 404–417.

Richerson, P. J. & R. Boyd (2001). Built for speed, not for comfort, Darwinian theory and human culture. *History and Philosophy of the Life Sciences*, 23, 425–465

Sober, E. & D. Wilson (1994). Reintroducing group selection to the human behavioral sciences. *Brain and Behavioral Sciences*, 17, 585–654.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*. 46, 35–57.

TABLE 1: Parameters of the baseline model.

	Parameter	Value
Number of groups	G	128
Group size	N	64
Cost of cooperation	c	0.2
Cost of punishing	k	0.2
Cost of being punished	p	0.8
Probability of erroneous defection	e	0.02
Migration rate	m	0.01
Innovation rate (behavioral strategies)	μ	0.01
Conflict rate	ε	0.015
Death rate	q	0.1
Mutation rate (learning rules)	v	0.001

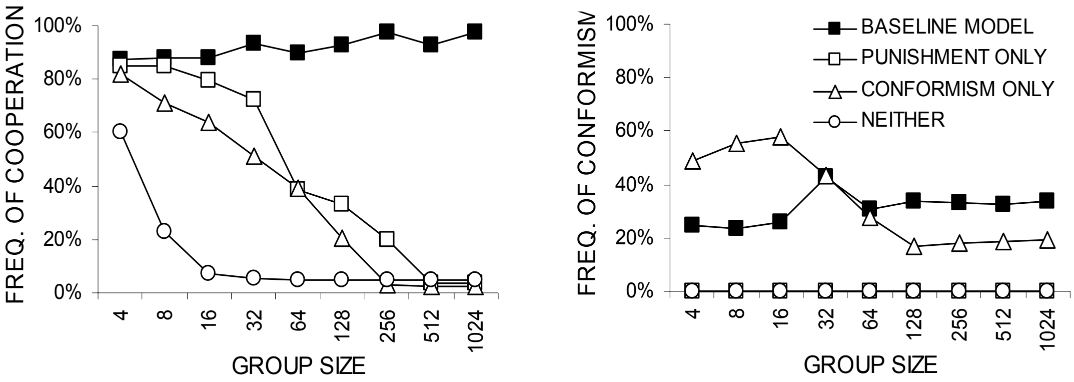

Figure 1: Cooperation and Conformism in Alternative Models

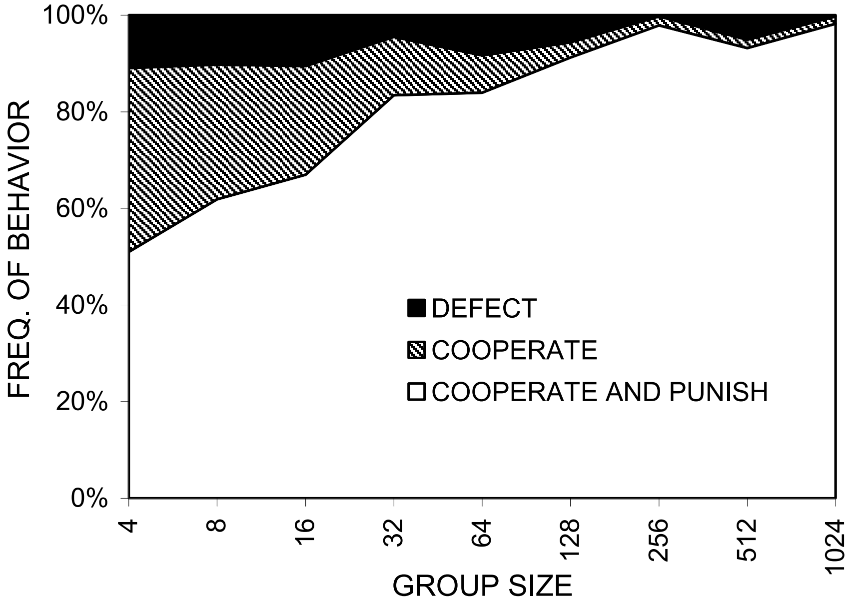

Figure 2: Distribution of Behavioral Strategies for the Baseline Model

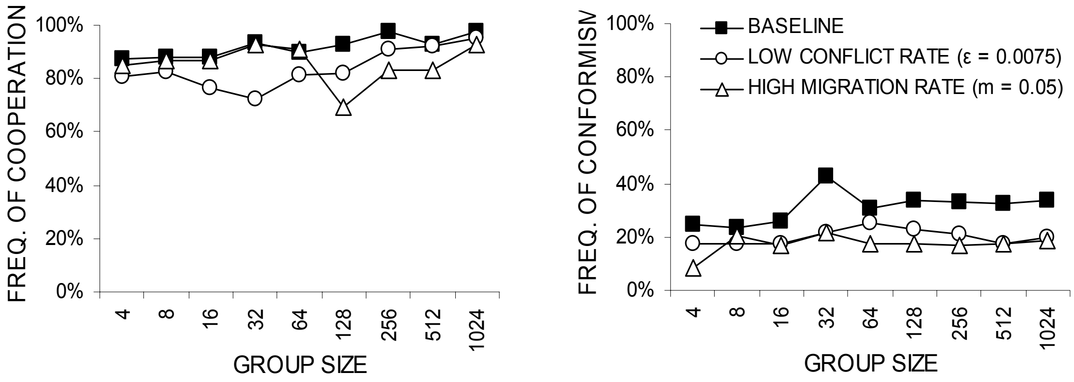

Figure 3: How Conflict and Migration Affect Cooperation and Conformism