ALTRUISTS, EGOISTS AND HOOLIGANS IN A LOCAL INTERACTION MODEL[1]

Illan Eshel
Department of Statistics
School of Mathematical Sciences
Tel Aviv University
Tel Aviv, Israel

Larry Samuelson
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, Wisconsin 53706 USA

Avner Shaked
Department of Economics
University of Bonn
Adenauerallee 24-26
D 53113 Bonn, Germany

May 14, 1996

[1]We thank Ed Green and Reinhard Selten for helpful comments. Financial support from the National Science Foundation and the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at the University of Bonn, is gratefully acknowledged. Part of this work was done while the first two authors were visiting the Department of Economics at the University of Bonn and while the latter two were visiting the Institute for Advanced Studies at the Hebrew University of Jerusalem. We are grateful to both for their hospitality and support.

ALTRUISTS, EGOISTS AND HOOLIGANS IN A LOCAL INTERACTION MODEL

by Illan Eshel, Larry Samuelson and Avner Shaked

1 Introduction

Despite the physical superiority of many animals, people are an evolutionary success. Darwin [9] traced the ascendancy of man to his "social qualities, which lead him to give and receive aid from his fellow-men." Other factors may also be important, but it is clear that people often engage in altruistic behavior.

An act is altruistic if it confers a benefit on someone else while imposing a cost on its perpetrator. From the standpoint of conventional economic theory, which assumes that people have well-defined, stable preferences and rationally choose utility-maximizing actions subject to resource constraints, altruism is hard to explain. How does costly altruistic behavior survive? Why doesn't the logic of utility maximization inexorably eliminate such behavior? This paper addresses these questions.

One answer is immediately available: the allegedly altruistic acts are not really altruistic. Instead, we have simply not identified preferences correctly, and what we have counted as a net cost is actually a net benefit.[1] If we push revealed preference theory to its logical limit, this conclusion becomes as inescapable as it is tautological. If someone commits an "altruistic" act, then this reveals that they prefer doing so, and hence they must derive net benefits rather than costs from the act.

We do not doubt that people often derive benefits from seemingly altruistic acts. However, we are unwilling to explain altruism by simply defining away the problem. Instead, we believe that if one begins with a substantive model of utility maximization, then one is often led to the conclusion that altruistic acts occur for which the model does not readily account. Embellishing the model to encompass such acts often leads to utility functions that are uncomfortably exotic. For example, people have given their lives to save strangers from danger.

[1] For example, charitable donations may be coupled with the provision of a benefit, such as public recognition, that overwhelms the cost of the donation. Alternatively, charitable giving is often explained by presuming that donors receive a "warm glow" from contributing (Andreoni [1]). Cooperation in the Prisoners' Dilemma is similarly explained in terms of a utility premium being placed on "being cooperative."

There are biological examples of altruism. Here, we have the advantage of an unambiguous measure of payoffs, namely reproductive success, and the issue of defining altruism away does not arise.[2] For example, there are several species of tropical butterflies that bear common markings and are preyed upon by birds (Benson [2], Eshel [12]). Some of these butterflies are endowed with genes that induce them to feed on herbs which make their flesh bitter. Others are genetically programmed to avoid such herbs and are "sweet." Predator birds cannot distinguish between bitter and sweet butterflies. However, if a bird eats enough bitter butterflies, it learns to stop preying upon these species. A bitter butterfly thus provides a public good, making the environment safer for his fellow butterflies by making it less likely that they will be eaten by his devourer.

This altruism comes at great personal cost to the bitter butterfly. His heroic death at the beak of the predator ensures that he does not enjoy the fruits of his own bitterness. More importantly, bitterness itself is costly. Bitter butterflies incur a biochemical cost that leads to lower expected rates of reproduction than sweet butterflies (Brower [5] and Brower and Brower [6]). Bitter butterflies thus bear an unambiguous cost for being altruistic. Our task is to explain how such altruism can survive.[3]

Behaving altruistically is analogous to cooperating in the Prisoners' Dilemma. The common response when studying altruism, and faced with the conclusion that rational behavior leads to defection in the Prisoners' Dilemma, is to abandon the assumption that the game is a one-shot game. If the interaction is repeated, then the folk theorem (Fudenberg and Maskin [17]) ensures that there are equilibria in which Altruists survive.[4]

[2] Some care must be taken, as reproductive success is not always straightforward to measure. Dawkins [10] argues that many apparently altruistic acts can be seen to be egoistic once payoffs are properly identified.

[3] Other biological examples include the commonly-observed inefficiency of weapons used in competition for mates, such as excessively branched or curved horns (Lorenz [25]). More extreme examples involve self-imposed limits on reproductive ability when a population is threatened by overpopulation. Wynne-Edwards [39] cites examples and evidence is provided by Christian [7] (rodents and small mammals), Fenner [15] (Mixoma viruses, which infect rabbits), and Stewart and Levin [33] (bacterial viruses). Finally, Cohen and Eshel [8] and Wood [37] discuss altruistic cooperation among bacteria.

[4] Kandori [22] and Okuno-Fujiwara and Postlewaite [31] show that cooperative outcomes can also be sustained when the one-shot Prisoners' Dilemma is repeatedly played by pairs of agents drawn from a large population of potential players. In biological contexts, explanations for altruism similar to those that arise in repeated games appear in the guise of reciprocal altruism.

We do not abandon the structure of the Prisoners' Dilemma. The keys to our explanation of altruistic behavior are two-fold. First, we abandon the assumption that people are rational agents choosing utility-maximizing actions. Instead, we believe that people must learn which actions work well, and that an important force in learning is imitation. Second, interactions between agents in our model are "local," meaning that altruistic acts are more likely to affect nearby agents than more distant neighbors and agents are more likely to imitate nearby than more distant neighbors.

To see how these forces can allow altruism to survive, suppose that there are two kinds of agents, Altruists, who provide a public good to their neighbors at a cost to themselves, and Egoists, who do not do so. Suppose further that Altruists tend to exist in concentrated groups. Altruists can then earn higher payoffs than Egoists, because Altruists are more likely to enjoy the public goods provided by other Altruists. The imitation-based learning process now prompts other agents to become Altruists. In addition, nearby agents are the ones most likely to imitate the Altruists. This preserves the tendency of Altruists to clump together in groups and hence preserves the conditions needed for altruism to survive.

This argument alone is unconvincing. Suppose that all agents are initially Altruists. Then imitation can never introduce any other type of behavior. At some point, however, we expect an Egoist to appear, perhaps because someone has analyzed the model and deduced that it is utility maximizing to be an Egoist, or someone has made a mistake, or someone has simply experimented with a new action. We group these possibilities together under the heading of "mutation." Because he enjoys the public goods provided by the Altruists, the mutant Egoist will thrive and will be imitated. More generally, an Egoist thrust into the midst of Altruists by a mutation will fare well and will be imitated, while an Altruist thrust in the midst of Egoists will fare poorly and will be ignored. Mutations thus apparently produces a force pushing toward egoism. We regard this as a crucial point, and consider essential that an explanation of altruistic behavior be able to demonstrate that Altruism can withstand such mutations.

The importance of mutations is reinforced by intuition drawn from biological models of spatial interaction. The potential importance of local interactions in biological models first appears in Wright's [38] concept of spatial isolation. More recently, similar ideas have appeared in models of group or kin selection, where agents are arranged in isolated groups created either by spatial separation or by kinship relationships.[5] However, group selection models not based on kinship relationships have been criticized as explanations for altruistic behavior (e.g. Dawkins [10], Williams [34]).The models were initially criticised because the mechanism that caused some groups to grow faster than others was not specified but subsequently because the combination of small groups, rare mutations and infrequent migration required to support altruism is thought to be implausible. In particular, the survival of altruism in the face of mutations appears to be problematic, as the success of altruism hinges upon a group of Altruists having adequate time to grow before being sullied by a mutant Egoist.[6]

Somewhat to our surprise, we find that introducing mutations can work in favor of Altruists. In particular, only states that are primarily composed of Altruists survive in the presence of rare mutations. What drives this outcome? We have noted that if a mutation introduces an Egoist in the midst of Altruists, then the Egoist will survive and spread. However, a group of Egoists in the midst of Altruists quickly confronts limits on its ability to expand, as each expansion causes the public goods supplied by neighboring Altruists to be shared among more and more Egoists and hence reduces Egoists' payoffs. Egoists are thus readily introduced but cannot expand beyond small, isolated groups. Isolated Altruists, in contrast, cannot even survive in the midst of Egoists. However, mutations will occasionally introduce a group of Altruists in the midst of Egoists. Such a group of Altruists can expand without bound. Mutations thus more readily lead to large groups of Altruists than Egoists, allowing the former to dominate in the presence of rare mutations.

The work most closely related to ours includes Blume [3], Ellison [11], and a series of papers by Nowak and May [28, 29, 30]. Our spatial structure matches the simplest case considered by Ellison, while Blume examines a spatial model in which agents are arranged in a plane rather than along a line. We differ from both in taking imitation, rather than some variant of best-reply dynamics, to be the driving force behind strategy selections. This is crucial, as Altruism has no hope in a world of best-responders.

---

[5] See, for example, Cohen and Eshel [8], Eshel [12], Eshel and Cavalli-Sforza [13], Hamilton [20, 21], Matessi and Jayakar [26], Maynard Smith [27], Williams [34], Wilson [35, 36], and Wynne-Edwards [39, 40].

[6] Boyd and Richerson [4] suggest that group selection arguments may be applicable in explaining the evolution of altruistic behavior among humans. They examine a model in which people have a taste for conformity, so that cooperating is a *strict* best response as long as sufficiently many other people cooperate.

Nowak and May simulate models in which agents are arrayed on the plane, playing the Prisoners' Dilemma with their neighbors, and changing strategies by imitating the neighbor with the highest payoff. They find outcomes in which Altruists and Egoists coexist. Our analysis differs in relying on analytical techniques rather than simulations, albeit for a very simple model, and especially in studying the effect of mutations.

Section 2 presents the model. Section 3 examines equilibria. We first examine the imitation process alone. There are many possible limiting outcomes of this learning process, depending upon the initial conditions of the system, but Altruists comprise a significant portion of the population in all but one of these. In addition, we establish conditions under which the probability of an initial condition leading to the elimination of altruism shrinks to zero as the population grows large.

We then incorporate mutations into the analysis. Here, the initial conditions become irrelevant and a "mutation-counting" argument allows us to select among the limiting outcomes of the imitation process. Only those limiting outcomes with a significant proportion of Altruists survive mutations.

Altruism is not the only type of externality that can arise between agents. The model can be easily adapted to consider Hooligans, or agents who benefit from imposing damages on their neighbors. The same forces that allow Altruists to survive in the presence of Egoists also allow Altruists to survive against Hooligans, or Egoists to survive in the midst of Hooligans, though Hooligans will typically not be eliminated entirely. We then find that the analysis can be extended to general 2×2 games. In particular, questions of payoff-dominance versus risk-dominance in games with two strict Nash equilibria can be considered. Imitation dynamics can have effects that are quite different from best-reply dynamics in such games.

Section 4 pursues generalizations of the model in which agents interact with more of their neighbors. In doing so, we find that it can be to Altruists' advantage to have a relatively high cost of altruism. A higher cost of altruism ensures that if Altruists survive, then they must do so in larger groups, because only then do they share enough of the public good to compensate for the high cost of being an Altruist. This in turn ensures that if there are any Altruists at all, then a higher proportion of the population is Altruists when costs are high. The argument is then completed by noting that once again, only those limiting outcomes with Altruists survive mutations.

Section 5 offers concluding remarks. Unless otherwise noted, proofs are contained in Section 6.

2 Altruists and Egoists

We consider a collection of N individuals, where N is finite. Each individual can be either an Altruist or an Egoist. An Altruist provides a public good that contributes one unit of utility to those who receive its benefits. The *net* cost to the Altruist of providing the public good is $C > 0$, so that the combination of enjoying the benefits of his own public good and bearing the costs of its provision reduces the Altruist's utility by C. Egoists provide no public goods and bear no costs. Instead they simply enjoy the benefits of the public goods provided by others.

Time is divided into discrete periods. At the end of each period, after consuming any public good that is available and bearing provision costs (if an Altruist), each agent observes the current environment and decides, according to a learning rule, whether to be an Altruist or Egoist in the next period.

The nature of this learning rule is important. One possibility is that the agents are fully rational and the learning rule leads them to adopt expected-utility-maximizing actions. In this case, they will realize they face a variant of the Prisoners' Dilemma and will play the strictly dominant strategy, namely Egoist. However, we assume that the players are not fully rational players. Instead of choosing best replies, our players imitate the strategies of others whom they observe to be earning high payoffs.

At one level, this imitation seems preposterous. How hard can it be to figure out that being an Egoist (or defecting in a Prisoners' Dilemma) is a strictly dominant strategy? This is indeed a trivial task for a game theorist facing the sterilized, 2×2 games with which we often work. However, these games are a simplified representation of a much more complicated reality, and it may be no easy task for the agents who must actually play the game to analyze this reality. These agents may not recognize that they are playing a game, may not know who their opponents are, may not know what strategies are available, and may not know what payoffs these strategies bring. It may then be impossible for the players to think like game theorists or like the agents in game-theoretic models. At the same time, we believe that people are generally able to form a good estimate of others' payoffs, whether these payoffs be measured in terms of money or other units such as social status or prestige,[7] and that people tend to imitate the behavior of those they observe

[7] For example, one need only observe how readily academics discuss either the incomes or the prestige of their peers.

earning high payoffs.

Even if people are unable to reproduce the reasoning processes of agents in game-theoretic models, it is not clear that they commonly make use of imitation. In particular, people are apt to give a host of reasons for their behavior, even to themselves, before resorting to "I saw Alice do it and she seems to be doing pretty well." We agree that imitation is not the only force shaping behavior. At the same time, we do not take the proclivity of people to cite reasons other than imitation for their actions as evidence that imitation is not important. Instead, we think that behavior is often established through imitation. Once it has become commonplace, however, a variety of other rationalizations are likely to appear to reinforce the behavior. For example, it is an altruistic act for people to bury their dead. However, when asked why the dead are buried, a host of reasons such as respect for the dead and religious imperatives are likely to be given. We thus suspect that imitation is a more powerful force than it may first appear.

Imitation alone appears to hold out no hope for the survival of altruism. Egoists will enjoy the same public goods as Altruists, while only the latter bear costs. As a result, all Egoists will earn higher payoffs than all Altruists and imitation can only lead players to become Egoists.

This argument is compelling only if the benefits of the public good provided by each Altruist extend to every agent in the population. The prospects for Altruists improve if the public good is a *local* public good. To make this precise, we introduce a neighborhood structure taken from Ellison [11]. Agents in the model are located around a circle.[8] Each agent interacts with his two immediate neighbors, i.e., with one agent to his right and one to his left. If an agent is an Altruist, then his immediate neighbors enjoy the benefit of his public good provision. The payoff of agent i is then given by $N_i^A - C$ if i is an Altruist and N_i^A if i is an Egoist, where $N_i^A \in \{0, 1, 2\}$ is the number of i's Altruist neighbors (excluding himself).

In each period, each individual assesses his strategy, or "learns."[9] He observes his own payoff and the average success of each strategy in his neighborhood. He then chooses to be an Egoist if the average payoff of the Egoists in his sample exceeds that of Altruists, and chooses to be an Altruist if the

[8]Though a spatial interpretation is convenient, local interaction structures may arise in other ways. In academia, for example, field of specialization is probably more important than location in determining patterns of interaction.

[9]This is a rigid rule, and we shall see that it can lead to volatile behavior. We discuss alternatives below.

average payoff of Altruists exceeds that of Egoists.[10] If an agent and his two neighbors all play the same strategy, be it Altruist or Egoist, then the agent will continue to play that strategy.

At the end of each period, and after imitation has occurred, each agent takes a draw from an independent, identically distributed Bernoulli random variable. With probability λ, this agent is a mutant and changes his type, either from Altruist to Egoist or from Egoist to Altruist. With probability $1-\lambda$, this agent experiences no mutation. We will be interested in the case in which λ is small, so that imitation is the primary force driving strategy revisions. We study this by examining the limiting case as the mutation probability λ goes to zero.[11]

A state is a specification of which agents are Altruists and which are Egoists. For states i and j, let p_{ij} be the probability that a single iteration of the imitation process changes the system to the state j given that the current state is i. Since the learning process is deterministic, p_{ij} is either 0 or 1. The collection $\{p_{ij}\}_{i,j\in\Theta}$ is a Markov process on the state space Θ. We refer to this Markov process as the "imitation dynamics." Let γ_{ij} be the probability that the combination of imitation and mutation changes to the state j given that the current state is i.[12] Then $\{\gamma_{ij}\}_{i,j\in\Theta}$ is again a Markov process on the state space Θ, which we refer to as the "imitation-and-mutation dynamics." Notice that $\gamma_{ij} > 0$ for all i and j, which is to say that for any two states i and j, there is some combination of mutations capable of changing the system from i to j.

[10] A tie would presumably prompt a random choice. We simplify the analysis by choosing C so that ties do not arise. There are many other plausible learning rules. For example, an agent may simply compare the best Egoist and best Altruist payoff among those payoffs he observes, or may compare the sum of the Egoist and Altruist payoffs, rather than considering averages. Gilboa and Schmeidler [19, 18], in the context of their case-based decision theory, examine the difference between considering the sums or the averages of payoffs. Agents may also learn from more (or fewer) agents then those with whom they interact, and it it may be that an agent pays more attention to some agents than others when choosing whose behavior to imitate. Our model employs a particularly simple "similarity" function that calls for the agent to consider all of those with whom he interacts, and to give equal weight to all those considered.

[11] In looking at a case of rare mutations, we are following the lead of Kandori, Mailath and Rob [23] and Young [41].

[12] If q_{ij} is the probability that mutations change the state to j, given that the current state is i, then $\gamma_{ij} = \sum_k p_{ik}q_{kj}$.

3 Equilibrium

3.1 Imitation

We first study the imitation dynamics in the absence of mutations. In terms of our model, we set $\lambda = 0$, leaving only the imitation process to shape behavior.

The imitation dynamics is a Markov process, and we are interested in the stationary distributions of this Markov process. We say that a set of states is *absorbing* if it is a minimal set of states with the property that the Markov process can lead into this set but not out of it. For each absorbing set of the Markov process, there is a unique stationary distribution whose support consists of that absorbing set. We can then learn much about the stationary distribution of the learning process by studying absorbing sets.

An absorbing set may contain only one state, say i, in which case $p_{ii} = 1$ and i is a stationary state of the Markov process. An absorbing set may contain more than one state, in which case $p_{ij} = 0$ if i is contained in the absorbing set and j is not, while the Markov process cycles between states in the absorbing set.

We begin our study of absorbing sets by compiling a simple description of the imitation dynamics. We assume $C < \frac{1}{2}$.[13] Under what circumstances will imitation cause an Egoist to become an Altruist? The "circumstances" in question here involve the strategies of the Egoist's four nearest neighbors, two on either side. The two agents closest to the Egoist are those whose strategies (as well as his own) he may imitate, depending upon their payoffs. These payoffs in turn depend on the strategies of the next two neighbors. The fate of an individual is then completely determined by the strategies of his four nearest neighbors.

An Egoist who learns by imitating his neighbors can become an Altruist only if at least one of his two nearest neighbors is an Altruist. If both of his immediate neighbors are Altruists, then the Egoist will not become an Altruist because the Egoist earns a payoff of two in this case, more than an Altruist can ever earn. An Egoist can therefore become an Altruist only if exactly one of his neighbors is an Altruist. In addition, the Egoist will convert to Altruism only if his Altruist neighbor obtains a higher payoff than

[13] We can always vanquish Altruistic behavior by making it too expensive. If $C > \frac{1}{2}$, we find that the only absorbing sets of the learning process are those consisting of single states in which either all agents are Altruists or all are Egoists. In the presence of mutations, only the latter absorbing state survives.

the average payoffs of the Egoist and his Egoist neighbor. This can occur only if the Altruist has an Altruist neighbor, since otherwise the Altruist receives the lowest possible payoff of $-C$, and if the other Egoist in the neighborhood faces a neighborhood containing only Egoists, so as to bring the average Egoist payoff below $1 - C$.[14] Hence, an Egoist can become an Altruist only if he faces either the following combination of strategies or its mirror image, where "a" represents an Altruist and "E" an Egoist,

$$\texttt{aa E EE} \tag{1}$$

and where it is the central Egoist who converts to an Altruist.[15] In all other cases, Egoists remain Egoists.

A similar calculation shows that an Altruist will remain an Altruist if and only if one of the following combinations of strategies (or their mirror images) occurs:

$$\begin{array}{c} \texttt{xa a ax} \\ \texttt{aa a EE} \end{array} \tag{2}$$

where it is the central Altruist whose fate is in question and where an "x" holds the place of an agent who may be either an Altruist or an Egoist. In all other cases, Altruists change to Egoists.

Conditions (1)–(2) provide a complete description of the individual imitation dynamics. Notice that the imitation dynamics is *boundary preserving* (see Eshel et al. [14]). In particular, the process cannot introduce new boundaries between groups of Altruists and Egoists, although a boundary may vanish as a group of Altruists or Egoists disappears.

To illustrate some absorbing sets, we represent the agents as being located on a line, where we think of the ends of the line as being joined to form a circle. From (1)–(2), we easily verify that the following are absorbing sets:

- The state in which all are Altruists

[14] This is where the requirement $C < 1/2$ is needed. If it is not satisfied then an Egoist can never become an Altruist.

[15] We use a lower case "a" to represent Altruists in order to make the displays easier to read, though we continue to use "A" to represent Altruists in the text. We will also often separate agents in whom we are interested by spaces, as in the case of the central Egoists here, though these spaces have no significance other than directing attention to particular agents.

- The state in which all are Egoists
- A state in which all are Altruists except two adjacent Egoists:

...aaaaaaaaEEaaaaaaaa...

- A set of two states, consisting of:

...aaaaaaaaaEaaaaaaaaa...

...aaaaaaaaEEEaaaaaaaa...

In this last case, the imitation dynamics cycle between the two states in the absorbing set. The lone Egoist initially earns the highest possible payoff of 2, inducing his two neighbors to become Egoists and leading to the second state in the cycle. Each of these new Egoists finds himself in the situation described by (1), where he has two Egoists on one side and two Altruists on the other. This causes the new Egoists to switch back to Altruism, beginning the cycle anew. We refer to such a cycle as a *blinker*.

The behavior described by a blinker may appear counter-intuitive. Why do these agents continually switch back and forth between A and E? Why can they not see that they are in a cycle, and simply adopt one behavior or the other?[16] On the one hand, we suspect that cycles in behavior do occur, as suggested by observations of the world of fashion, though our simple model captures these cycles in a particularly crude way. On the other hand, we could easily reformulate the model to eliminate blinkers. The presence of blinkers is a product of our rigid imitation rule that forces all agents to assess their strategies in every period, and blinkers would not arise if there were some inertia in learning. We examine an imitation scheme in Section 3.3 in which each agent takes a random draw from an independent, identically distributed Bernoulli trial in each period, causing the agent to "learn" with probability μ and to retain his strategy with probability $1-\mu$. This allows for the possibility that all agents will adjust their strategies in a given period, though typically only a subset of agents will learn. We refer to this as the "random imitation" model. Random imitation leaves the results in this section unchanged, except that blinkers are no longer absorbing sets.

[16] Equivalently, the two outside agents in the blinker face a coordination problem. It is an equilibrium for one but not for both to be an Egoist, and the learning scheme causes them to cycle around this equilibrium. Why don't they learn to coordinate?

Instead, all absorbing sets are singletons, containing groups of A's of length three or more separated by groups of E's of length two.[17]

These examples, and combinations constructed from them, include all of the possibilities for absorbing sets. Some terms will be useful in making this precise. If agents α and β play the same strategy, either Altruist or Egoist, and if all agents between α and β play this strategy, then we will refer to agents α, β, and the intermediate agents as an *interval* of either Altruists or Egoists. We call a maximal such interval a *string.* Notice that strings may be of any length from 1 to N, the length of the circle. We then have (the proof is in the Appendix):

Proposition 1 *Let* $0 < C < \frac{1}{2}$. *Then:*

(1.1) Absorbing sets consist of (i) *the state in which all agents are Egoists,* (ii) *the state in which all agents are Altruists, and* (iii) *sets containing states in each of which Altruist strings of length three or longer are separated by Egoist strings of length less than three. These sets are either singletons (in which case all Egoist strings are of length two) or contain two states (in which case any string of length one (three) in one of the states blinks to a string of length three (one) in the other.*

(1.2) Except for the state in which all agents are Egoists, the proportion of Altruists in an absorbing state, or the average proportion over the two states in an absorbing set, is at least 0.6.

Proposition 1 indicates that there are many absorbing sets, each of which is the support of a stationary distribution of the imitation process. In all but one of these absorbing sets, the majority of the population is Altruists. Hence, there is no possibility for moderation in Altruism. If Altruists survive at all, they must be the majority.

To see what lies behind this result, we first note that a single Altruist in the midst of Egoists will vanish, leaving no trace, and therefore cannot be a part of an absorbing set. Secondly, a string of Egoists in an absorbing set can never be longer than three. If the length of an Egoist string exceeds three, then it necessarily shrinks. The two Egoists at its edges will each

[17] To see why blinkers disappear, consider the state in which only the center member of the blinking trio is an Egoist. With positive probability, one and only one of his neighbors receives the learn draw and switches to being an Egoist. But now we have a pair of adjacent Egoists. Subsequent learn draws on the part of these agents will yield no further strategy revisions, so that the two Egoists will persist. Eventually, such events will eliminate all blinkers and hence, no absorbing set contains a blinker.

have two Egoists on one side and two Altruists on the other, and hence they will become Altruists (cf. (1)). Egoists can thus survive only in short strings, i.e, in strings of length two or strings of length one (where the latter alternate with strings of length three in a blinker). Altruist strings, in contrast, can expand, since doing so creates more and more high-payoff Altruists. The survival of an Altruist string in an absorbing set requires only that it be of length at least three, because then the interior Altruist earns a sufficiently high payoff to preclude the Altruists on the end from imitating their neighboring Egoist. This allows us to conclude that if there are any Altruists at all, then Altruists will occur in strings of length at least three while Egoist occur in strings of at most two (or in blinkers whose average length is two). Hence, if there are any Altruists at all, then there will be at least sixty percent of Altruists.

Because the state in which all agents are Egoists is absorbing, the system may drive Altruists to extinction. What is the likelihood of such an event? One way of making this question precise is to identify the initial conditions from which the system converges to an absorbing set containing Altruists. Such an initial condition need only contain at least one string of Altruists sufficiently long to preclude extinction. If the circle is sufficiently large and initial conditions are determined randomly, there will almost certainly be such an initial string:

Proposition 2 *Let* $0 < C < \frac{1}{2}$.

(2.1) The system converges to an absorbing state with at least sixty percent Altruists if the initial state contains either (i) *a string of at least five Altruists;* (ii) *a string of four Altruists bordered on at least one end by a string of at least two Egoists, or* (iii) *a string of three Altruists bordered on at least one end by at least three Egoists.*

(2.2) If agents' initial identities as Altruists or Egoists are determined by an independent, identically distributed random variable, then as N *gets large, the probability that these sufficient conditions hold approaches unity.*

3.2 Mutations

We now investigate how the possible survival of Altruists is affected by rare mutations. Hence, we let $\lambda > 0$, though we will be interested in the case in which λ is small. The first thing to note is that all of the transitions in the imitation-and-mutation dynamics Markov process are positive, so that for any two states i and j, $\gamma_{ij} > 0$. We then have the following familiar result from the theory of Markov processes:

Lemma 3 *For a fixed mutation rate, the imitation-and-mutation dynamics has a unique stationary distribution. The proportions of states reached along any sample path approach this distribution almost surely, and the distribution of states at time t approaches this distribution as t gets large.*

Proof Kemeny and Snell [24], Theorems 4.1.4, 4.1.6, and 4.2.1. □

For each mutation rate, we thus have a unique stationary distribution of the imitation-and-mutation dynamics. We study the limit of these stationary distributions as the probability of a mutation λ gets small, which we denote as the *limiting distribution.*

Proposition 4 *Let $0 < C < 1/2$. If N is sufficiently large, then the limiting distribution places positive probability only on states contained in absorbing sets of the imitation process in which the proportion of Altruists is at least* 0.6.

The techniques involved in establishing this result were developed by Freidlin and Wentzell [16] and were introduced into economics by Young [41] and Kandori, Mailath and Rob [23]. The argument begins by observing that when the mutation rate is small, the system spends virtually all of its time in absorbing sets of the imitation dynamics. Equivalently, the limiting distribution allocates all of its probability to such sets. Movements between absorbing sets of the imitation dynamics can be accomplished only by mutations. The system will allocate most of its probability to absorbing sets of the imitation process that are easy to reach, in the sense that it requires relatively few mutations to reach their basin of attraction from other absorbing sets. The proof involves showing that it is much easier for mutations to introduce Altruists into a world of Egoists than for mutations to eradicate Altruists from a mixed world.

One's initial impression might be that mutations should be inimical to Altruists, because a mutant Egoist will thrive and grow when introduced into a collection of Altruists while an Altruist will wither and die when introduced in a collection of Egoists. Then how can Proposition 4 hold?

To explain this seeming paradox, we note first that a single Altruist will not survive in the midst of Egoists. However, a small clump of Altruists will not only survive, but will grow. It takes only three adjacent Altruists to spread in a world that is completely Egoists, while five adjacent Altruists never vanish so the absorbing set reached will have at least 60% Altruists. [18] A minimum of three and a maximum of five mutations are then required to not only introduce Altruists into our world but to ensure that at least sixty percent of our agents are converted to Altruism.

In contrast, in takes only a single mutation to introduce an Egoist into a world of Altruists. However, the prospects for the growth of this newly created string of Egoists are quite limited. The payoffs of the members of this Egoist string drop as they grow (because they are interacting with more Egoists and fewer Altruists) and the string can grow no longer than three.

In light of this, consider an initial state that consists only of Altruists. A mutation creating an Egoist or even a clump of Egoists will ultimately, after imitation has led us to an absorbing state, yield no more than three Egoists. To get additional Egoists, additional mutations are required. These mutations can lead to states where there are many small clumps of Egoists. What happens as these clumps become more dense, and hence closer together, raising the proportion of Egoists? Why can't we continue in this fashion until all Altruists are eliminated? The answer is that as strings of Egoists become more densely packed in the sea of Altruists, additional mutations join together previously separated clumps of Egoists. But as long as there are still some strings of Altruists, these newly joined strings of Egoists will shrink, replacing two original strings with a new, shorter string (of length three or less) and ultimately *decreasing* the proportion of Egoists. In order to further increase the proportion of Egoists, mutations must simultaneously eliminate *all* strings of Altruists. But this requires a large number of mutations, at least if N is large, and hence is extraordinarily unlikely.

The advantage enjoyed by Altruists is then that Altruists can invade a world of Egoists with only a local burst of mutation that creates a small string of Altruists, who will then subsequently grow to a large number of Altruists. Mutations can create small pockets of Egoism, but these pockets destroy one another if they are placed too close together, placing an upper bound on the number of Egoists that can appear. The only possibility for surpassing this bound lies in a global mutation episode that simultaneously attacks all strings of Altruists. Mutations thus lead much more readily to absorbing sets with Altruists than absorbing sets without them, and the

[18] Notice that it may take four or five rather than three adjacent Altruists to spread if some nearby agents are Altruists rather than Egoists. This occurs because the nearby Altruists make the nearby Egoists better off, making it harder to convert the latter to Altruism and requiring a larger initial clump of Altruists.

limiting distribution concentrates all of its probability on the former.

Evolutionary models in which results are driven by mutation-counting arguments must be interpreted with some care. The results are established by computing a limit as the probability of a mutation becomes small. We do not believe that mutations are literally arbitrarily improbable. Instead, we view this limit as a convenient approximation of the case in which mutations occur but are unlikely. We also think that the most interesting case is that in which the population is large. This could also be approximated by taking a limit as the population size N gets arbitrarily large. In Proposition 4, we have examined one order in which the limits can be taken, with the probability of a mutation getting small before the population grows large. This is the type of inquiry most commonly found in the literature, but how do we know this is the right order, and does it matter?

Fortunately, the order of limits does not matter. The result of Proposition 4 could have been obtained by fixing the mutation rate and letting the size of the population grow. In particular, there are two key transitions behind Proposition 4. One involves moving from a state in which all agents are Egoists to a state in which Altruists survive, and requires three mutations, regardless of the population size. The other involves moving from an absorbing set in which Altruist survive to the state in which all agents are Egoists. This transition requires a number of mutations that grows linearly in the size of the population. The proof of Proposition 4 centers around the observation that for a fixed population size, the former transition becomes arbitrarily more likely as the probability of a mutation shrinks. Similarly, for a fixed mutation probability, the former transition becomes arbitrarily more likely as the population size grows, again leading to Proposition 4. We will be suspicious of any limiting result that does not have this robustness property.

3.3 Random Imitation

The imitation process in our model is rather rigid. All players assess their strategies in each period according to a deterministic imitation rule. When discussing *blinkers* in the previous section we raised the possibility of random imitation, in which each player assesses his strategy with probability $\mu > 0$ in each period. The resulting process is more complicated but has simpler absorbing sets. A blinker persists because the learning of the two outside Egoists is perfectly synchronized. If just one of these agents were to revise their strategy, the results would be pair of adjacent Egoists, neither of whom would then have an incentive to change strategies. But if learning is random, realizations of the learning-and-imitation process in which only one of the outside agents in a blinker revises strategies will eventually eliminate all blinkers. As a result, there are no absorbing sets containing blinkers under random imitation. Instead, the absorbing sets are singletons, consisting of states in which either all agents are Egoists, or all are Altruists, or strings of two Egoists are scattered among strings of Altruists of length three or longer. Except for the set with no Altruists, the proportion of Altruists in all other absorbing sets is at least 0.6. Propositions 1–4 therefore hold for this case.

What if once a player has received the impetus to learn, the imitation process itself is stochastic? Assume that a player continues to compare the average payoffs of the two strategies in his neighborhood, switching over to the one with the higher average payoff with a probability (but not changing his strategy if it is the one with the highest average payoff). This probability may depend both on the average payoffs and on the player. This "stochastic imitation" process is quite similar to the random learning model, since a player in the stochastic process who was given the opportunity to switch to a best reply and mistakenly does not do so is equivalent to a player in the random learning process who was not given the opportunity to revise strategies. Again, the results of Propositions 1–4 hold.

3.4 Hooligans

Altruism, conferring a benefit on someone else at a cost to oneself, is not the only way that one agent's actions may affect another. At the opposite extreme we have Hooligans, who benefit by imposing harm on others. Notice that hooliganism need not be limited to the psychopathic. Those who litter in order to avoid the cost of disposing of their refuse, those who pollute rather than take costly abatement measures, and those who shirk in group efforts are all Hooligans. Notice also that the prospects for the survival of hooliganism may be quite good. Because Hooligans impose harm on others, and either derive benefit from doing so or at least avoid costs that must be incurred to not do so, the average payoffs of Hooligans are likely to be relatively high. They may then be imitated and may well prosper.

Our model is easily generalized to accommodate Hooligans. The model is built on the assumption that Egoists neither make contributions to the welfare of others nor bear costs, while Altruists contribute a benefit one to each of their neighbors at a cost C to themselves. More generally, let there be two types of agents, denoted types 1 and 2. Let type 1 contribute K_1 to the payoff of each of his neighbors at a cost of C_1 to himself. Let type 2 contribute K_2 to each neighbor at a cost C_2 to himself. There is no loss of generality in assuming that $K_1 > K_2$.[19] It turns out that the behavior of the model depends only on a single parameter:

Proposition 5 *Let $K_1 > K_2$. Then any variation in the values of K_1, K_2, C_1, and C_2 that preserves*

$$\frac{C_1 - C_2}{K_1 - K_2} \tag{3}$$

gives rise to the same imitation dynamics.

Hence, any two specifications of the payoffs K_1, K_2, C_1 and C_2 that preserve $(C_1 - C_2)/(K_1 - K_2)$ give rise to the same behavior, including the same absorbing sets, basins of attraction, and dynamic paths for the imitation dynamics, and the same limiting distributions in the presence of mutations. This opens the door to a host of new interpretations and applications of the model. For the Altruist and Egoist model of the previous sections, the ratio (3) was C, which was interpreted as the cost of Altruism. Consider the following pairs of types of players. In each case, the first column identifies the effect an agent of type 1 has on his two neighbors and the cost to the agent of that effect, while the second column provides analogous information for an agent of type 2.

$$\begin{pmatrix} K_1 & K_2 \\ C_1 & C_2 \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 \\ C & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 0 & -C \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ C & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ C & -C \end{pmatrix}, \begin{pmatrix} -1 & -2 \\ -C & -2C \end{pmatrix}.$$

The first specification is the familiar Altruist and Egoist pair from previous sections. The second pair of agents consists of an Egoist and a Hooligan who enjoys (incurs a negative cost) causing damage of one unit to his neighbors. In case this Hooligan seems too malicious in his enjoyment of the harm he causes, the third pair rewrites this as an agent of type 1 who imposes no harm on others but incurs a cost of C to avoid doing so, with a type-2 agent who does not incur the cost and imposes damage of one unit on his neighbors.

[19] Our model of Altruists and Egoists is then the special case in which $K_1 = 1, C_1 = C$ and $K_2 = C_2 = 0$.

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{II} \\
 & & \multicolumn{1}{c}{X} & \multicolumn{1}{c}{Y} \\
\cline{3-4}
I & X & a, a & b, c \\
\cline{3-4}
 & Y & c, b & d, d \\
\cline{3-4}
\end{array}
$$

Figure 1: 2×2 Game

The third pair includes an Altruist and a Hooligan. The last pair has two Hooligans, one of whom causes twice the damage and doubly benefits from doing so. In each of these specifications, the ratio $(C_1 - C_2)/(K_1 - K_2)$ is given by C, and hence these are equivalent models. As long as $C < \frac{1}{2}$, it is always the first type in each pair whose survival is guaranteed. Hooligans then do not fare well against Egoists or Altruists, but Hooligans survive if paired with even worse Hooligans.

Similar insights can be used to extend the analysis to general 2×2 game. Consider the game shown in Figure 1.

Without sacrificing generality, we can assume $a > d$. We will then further concentrate on the case in which $a > b$. This latter assumption excludes some games but retains all of the common examples of 2×2 games. Then an argument analogous to the proof of Proposition 5 shows that the resulting imitation dynamics depends only upon the two numbers:

$$\alpha = \frac{c - b}{a - b}, \qquad \beta = \frac{d - b}{a - b}. \tag{4}$$

In particular, two specification of payoffs that give the same value for α and the same value for β thus give rise to identical behavior. In light of this, we can transform the payoffs in Figure 1 by subtracting b from each payoff and dividing by $a - b$ to obtain the equivalent representation of the game given in Figure 2.

We can now classify games according to the values of α and β, where $\beta < 1$ (because we have assumed $a > d$). We have:

- Prisoners' Dilemma: $0 < \beta < 1, \quad 1 < \alpha$.
- Coordination Game: $0 < \beta < 1, \quad \alpha < 1$.
- Chicken: $\beta < 0, \quad 1 < \alpha$.
- Efficient Dominant Strategy: $\beta < 0, \quad \alpha < 1$.

$$
\begin{array}{cc|c|c|}
\multicolumn{4}{c}{\text{II}} \\
 & \multicolumn{1}{c}{} & \multicolumn{1}{c}{X} & \multicolumn{1}{c}{Y} \\
\cline{3-4}
I & X & 1,1 & 0,\alpha \\
\cline{3-4}
 & Y & \alpha,0 & \beta,\beta \\
\cline{3-4}
\end{array}
$$

Figure 2: Transformation of game in Figure 1

Classification of Games

This classification is illustrated in Figure 3. An "efficient dominant strategy" game is one in which X is a strictly dominant strategy and the outcome (X, X) is efficient, unlike the Prisoners' Dilemma. In the case of a coordination game, the payoff-dominant equilibrium (X, X) is also risk dominant if $\alpha + \beta < 1$, while the equilibrium (Y, Y) is risk dominant in $\alpha + \beta > 1$. The short interval $\alpha = 1 + C,\ \ \beta = C,\ \ C < \frac{1}{2}$ (Figure 3) describes the range of Altruists & Egoists games that was analysed in the previous sections. However, the methods that were developed for the Altruists & Egoists games can be directly applied to any other game in this classification. The absorbing sets of the learning process can be found and the limit of the learning and mutation process as the mutation rate decreases. The rest of this section is devoted to a discussion of Coordination Games.

A common target for recent evolutionary analyses has been the tension between risk-dominance and payoff-dominance as equilibrium selection principles in 2×2 coordination games. The structure of our model of spatial interaction is taken from Ellison [11], who shows that best-reply learning in such a model leads to the selection of the risk-dominant equilibrium. It is illuminating to compare the differences between best-reply learning and imitation. Let $\alpha + \beta > 1$ so that (X, X) is the payoff-dominant equilibrium but (Y, Y) is the risk dominance equilibrium. Now consider a boundary between a group of agents playing strategy X (denoted by x in the following figure) and a group playing strategy Y, or

...xxxxxxxxxxxYYYYYYYYYYYY...

The only agents at risk of changing their strategies are the two agents, one playing X and one play Y, at the ends of their respective strings. Each faces a neighborhood with one X and one Y agent, in addition to themselves. In Ellison's best-reply learning, each chooses a best response to their two neighbors. By assumption, Y is risk-dominant and hence is best reply when one neighbor plays X and one plays Y. Hence, the agent playing Y retains his strategy while the agent playing X switches to Y. The string of Y's thus grows while the string of X's shrinks.

In our imitation model, the X player on the boundary earns a payoff of 1, while the adjacent X player earns 2. The Y player on the boundary earns $\alpha + \beta$ while the adjacent Y player earns 2β. Computing and comparing the average payoffs, we find that the boundary player Y retains his strategy if $\alpha + 3\beta > 2$, while a boundary X will turn into Y if and only if $\alpha + \beta > \frac{3}{2}$. It is therefore possible to find three regions of (α, β), in all regions (Y, Y) is the risk dominant equilibrium while (X, X) is the payoff dominant one. In the first region both boundary players will play Y in the following period, in the second region both will play (X, X) while in the third both retain their strategy. In the first region the string playing the risk dominant action will grow, in the second it will shrink while the payoff dominant strategy will win and in the third region each string maintains its length.

When strategy adjustments are driven by imitation, then strategy Y being risk-dominant does not suffice for the string playing Y to expand. If a string of agents playing the risk-dominant action Y is to expand, its payoffs must provide a premium over that required for risk-dominance. This is necessary because an agent at the end of a string of X agents compares not whether X or Y is a best reply, but whether the X or Y players in his neighborhood are earning higher average payoffs. One of the X players in his neighborhood is bordered by two X players and hence receives an exceptionally high payoff. Risk dominance alone is not enough to overcome this payoff.

It is straightforward to see that if a string of agents playing the payoff-dominant action X is to expand (while keeping Y risk-dominant), then $\alpha > \beta$ must hold. Hence, the payoff to playing strategy Y must be greatest if the opponent plays X, even though (Y, Y) is an equilibrium. This occurs because the neighborhood of a Y playing on the end of a string of Y contains, in addition to himself, a Y player who faces two Y opponents and hence earns a relatively high payoff. The average payoff to X can be highest only if the payoff in equilibrium (Y, Y) is relatively small. This is in turn compatible with risk dominance only if $\alpha > \beta$.

Given that strategy Y must receive a premium over risk-dominance in order to expand, and an additional condition is required on the relative magnitudes of the payoffs to Y if X is to expand, it is then no surprise that there are some cases in which neither string will expand. Imitation can then yield peaceful coexistence of the two strategies in some cases where best-response behavior would banish one strategy. Imitation allows these coordination failures to occur because agents on the boundary of a string, and hence experiencing coordination failures, are most likely to observe other agents of their own strategies who are not facing coordination failures but agents with the other strategy who are plagued by such failures. This introduces a force against changing strategies, and builds sufficient inertia into the system to support coexistence.

4 Larger Neighborhoods

We have assumed that agents interact only with their immediate neighbors. This has the technical advantage that learning by imitation, while shifting existing borders between regions of Altruists and Egoists, creates no new borders. If an individual can learn from others who are not his immediate neighbors, then an Altruist currently sitting inside a string of Altruists may become an Egoist. This may create a new string of Egoists and increase the number of borders.

It is clearly important to investigate the how Altruists fare in more complicated models. In this section we consider the case where each Altruist contributes one unit of the public good to each of his *four* closest neighbors. Each agent observes his own payoff and that of his four closest neighbors, and then chooses the strategy from those played by this group with the highest average payoff. We say that neighborhoods are of “radius” two in this case.

As in the previous case, the cost of Altruism plays a crucial role in shaping the results. We must pay particular attention to "critical" cost levels, or cost levels that create cases in which the average payoff to the Altruists and Egoists in a given agent's neighborhood are equal. Values of C that lie on different sides of a critical cost level often give rise to different behavior, while a critical cost level itself creates special difficulties which we avoid by (generically) restricting C to take on noncritical values.

We study two intervals for the parameter C, namely $(3/4, 5/6)$ and $(5/6, 1)$.[20] Changing the value of C within such an interval does not affect the outcome, while we shall see that the two intervals give different behavior.[21]

We again find that there is a lower bound on the proportion of Altruists in any absorbing set that has Altruists at all. We were initially surprised to find that this lower bound decreases with C. Hence, there may be fewer Altruists when the cost of being an Altruist is lower $(3/4 < C < 5/6)$ than is possible when cost is high $(5/6 < C < 1)$. Higher cost ensures that if Altruists do survive, then they must do so in relatively long strings, since only then can the public goods provided by the Altruists overwhelm the relatively high cost of altruism. Altruists can survive in smaller groups than when the cost of being an Altruist is lower, allowing the minimum proportion of Altruists (given that there are Altruists at all) to fall. In both cases, a sufficiently long initial cluster of Altruists will survive and grow under any circumstances, ensuring that Altruists will prosper under the imitation process when the initial identity of the individuals is determined by a random variable and when N is sufficiently large (a result analogous to Proposition 2).

The effects of mutations are quite different for the two cases. When costs are high $(5/6 < C < 1)$, the limiting distribution will always include Altruists. This is not the case when costs are low $(C \in (3/4, 5/6))$, where the limiting distribution is concentrated on the single absorbing set with only Egoists.[22] The mutation-counting argument for the case of $5/6 < C < 1$

[20] When agents interacted only with their immediate neighbors, the only relevant cost consideration was whether C was larger or smaller than $\frac{1}{2}$.

[21] If $C > 5/4$, then Altruism is so costly that only Egoists survive. The results for $1 < C < 5/4$ are qualitatively similar to those for $5/6 < C < 1$, with one quantitative difference noted below. Cost levels $1/2 < C < 3/4$ similarly give results similar to those of $3/4 < C < 5/6$. Costs $C < 1/2$ give noticeably different and more complicated behavior that we do not investigate here.

[22] This is another version of the result that the easier it is for an individual to be an Altruist, the fewer Altruists will survive.

satisfies our criterion of continuing to hold if the population size is allowed to grow before the probability of a mutation shrinks to zero. However, this is not the case for cost levels $3/4 < C < 5/6$. This case accordingly calls for a cautious interpretation of the results and further investigation.

4.1 High costs ($5/6 < C < 1$).

The investigation of the model with neighborhoods of radius two and costs $C \in (5/6, 1)$ begins with a calculation of transition rules. The conditions under which an Altruist will remain an Altruist are the following cases and their mirror images, where the A in the center is the agent in question and x stands for a strategy that could be either A or E:

$$\begin{array}{l} \texttt{Eaaa a xEEE} \\ \texttt{aaaa a EEEx} \\ \texttt{aaaa a AEEE} \end{array} \tag{5}$$

The cases in which an Egoist will become an Altruist are the following, as well as their mirror images:

$$\begin{array}{l} \texttt{aaaE E EEEx} \\ \texttt{Eaaa E EEEE} \\ \texttt{aaaa E EEEx} \end{array} \tag{6}$$

These allow us to prove:

Proposition 6 *Let neighborhoods be of radius two and let $5/6 < C < 1$. Then absorbing sets generically consist of:*

- *The state in which all agents are Egoists and the state in which all agents are Altruists.*
- *Sets containing states in which strings of Altruists of length five or more are separated by either strings of three E's or blinkers, where blinkers consist alternately of one E and five E's or consist alternately of two E's and six E's.*

With the exception of the state in which all agents are Egoists, the proportion of Altruists is at least $5/9$. If the initial condition is obtained by *a collection of draws from independent, identically distributed random variables, then as N grows the probability of an initial condition from which the system reaches an absorbing set with at least* $5/9$ *Altruists approaches unity.*[23]

The proof, which mimics those of Propositions (1) and (2) and which is omitted, proceeds by first observing that any strings of A of length one or two will be immediately eliminated without creating any new strings of A's. Strings of A's of length three of four will either be eliminated or will give rise to one or two strings of A's of length one or two, each of which will then be eliminated. Hence, we must reach a state in which A's occur in strings of length five or longer, and these strings must be separated by E strings of length three or blinkers which alternate between string lengths of one and five or two and six. To obtain the minimal proportion of Altruists in the stable sets that contain Altruists, we note that we can pack blinkers that alternate between two and six E next to each other with five A's between them, in the following way:

$$\begin{array}{c}\texttt{...aaaaaEEaaaaaEEEEEEaaaaa...}\\ \texttt{...aaaEEEEEEaaaaaEEaaaaaaa...}\end{array} \tag{7}$$

This guaranties a maximum of Egoists, and here we have a proportion of $10/18 = 5/9$ Altruists. Hence, if A's exist in an absorbing set, at least $5/9$ of the agents must be A's. Finally, we note that a string of nine A's in the initial state (among other sufficient conditions) suffices to ensure that the system reaches an absorbing set including Altruists.

Next, we add mutations and obtain the following proposition, whose proof is a straightforward variation on the proof of Proposition 4.

Proposition 7 *Let neighborhoods be of radius two and let* $5/6 < C < 1$*. Then generically the limiting distribution attaches zero probability to the state in which all agents are Egoists.*

4.2 Low costs $(3/4 < C < 5/6)$.

For costs in the interval $(3/4, 5/6)$, we have the following:

[23] For $1 < C < 1.25$ ($C \neq 7/6$), we have the same characterization of absorbing sets, except that blinkers must be separated by at least six Altruists, making the minimal percentage of Altruists 0.6, higher than 5/9.

Proposition 8 *Let neighborhoods be of radius* 2 *and let* $3/4 < C < 5/6$. *Then generically, absorbing sets include:*

(8.1) The state in which all agents are Egoists and all are Altruists.

(8.2) A collection of sets containing states in which strings of three or more Altruists are separated by strings of exactly three Egoists, or by blinkers which alternate between one and five or between two and six Egoists.[24]

Except for the state in which all agents are Egoists, the proportion of Altruists in an absorbing set is a least $1/2$. *If the initial condition is generated by a collection of independent, identically distributed random variables, then as* N *grows the probability of an initial condition from which an absorbing set is reached where Altruists survive approaches unity.*

The proof again mimics that of Propositions (1) and (2). In this case, the relevant transitions are the following. The cases in which an A remains an A (in addition to the obvious case in which an A is surrounded entirely by A's) consist of the following and their mirror images:

```
aEaa a xEEE
Eaaa a EEEx        (8)
aaaa a EEEx
xaaa a aEEE
```

The cases in which an E becomes an A are the following and their images:

```
Eaaa E EEEx
aaaa E EEaE        (9)
aaax E EEEx
```

The argument then proceeds as previously. To obtain the lower bound on the number of Altruists, note that it is possible to pack blinkers in the following way:

```
...aaaEaaaEEEEEaaaEaaaEEEEEaaaEaaa...
...aEEEEEaaaEaaaEEEEEaaaEaaaEEEEEa...        (10)
```

[24]If two one/five blinkers are separated by a string of only three Altruists, then the blinkers must be out of phase, so that the state in which one of the blinkers has five Egoists is the state in which the other blinker has only one Egoist. A two/six blinker requires at least five Altruists on each side.

This has $\frac{1}{2}$ of the population as Altruists. There is no denser way to pack blinkers. For the second part of the proposition it suffices to show that a cluster of nine Altruists will survive under any circumstances.

Once again we have reason to believe that in the long run, Altruists are likely to flourish in this model. The lower bound on Altruists in this case is lower than in the case of higher costs, being 1/2 rather than 5/9. In this sense, it can be disadvantageous for Altruists to have their altruism come too cheaply. The forces behind this result are revealed by comparing (7) and (10). Example (7) reflects the fact that when costs are relatively high, strings of Altruists must be at least five Altruists long in order to survive. The result is then small islands of Egoists separated by at least five Altruists. Example (10) reflects the fact that for lower costs, strings of three Altruists can survive. The imitation dynamics can then lead to outcomes in which islands of Egoists are separated by strings of only three Altruists, and hence a smaller proportion of Altruists.

When introducing mutations, we have the following:

Proposition 9 *Let neighborhoods be of radius two and let* $3/4 < C < 5/6$. *Then the limiting distribution attaches unitary probability to the state in which all agents are Egoists.*

The proof is contained in the Appendix. The heart of the proof is constructive, showing that we can find a sequence of absorbing sets, beginning with the state in which all agents are Altruists and ending with the state in which all agents are Egoists, with each absorbing set being reached from its predecessor by a single mutation. In contrast, if we begin only with Egoists, four mutations are required to introduce a cluster of Altruists that will survive. Using the techniques of Freidlin and Wentzell [16], it is straightforward to use these two facts to show that only Egoists survive in the limit as the probability of a mutation becomes very small.

The initial steps in the sequence of absorbing sets that leads from all Altruists to all Egoists are straightforward. Beginning with the state in which all are Altruists, a sequence of single mutations can lead to an absorbing set containing densely packed blinkers, as in (10). The next step is to show that adding another Egoist mutant, in precisely the right position, will destroy all Altruists.

Unlike previous cases, this result does not survive reversing the order of limits and first allowing the size of the population to grow. As the proof of Proposition 9 shows, the path from absorbing sets with Altruists to the state of all Egoists requires only single-mutation links, but only a very special sequence of mutations will work. As the population grows, the likelihood of such a sequence shrinks to zero, reversing the result. It thus remains an open question as to which absorbing sets, those with or without Altruists, provide the best approximation in this case. It is clear, however, that an analysis based entirely on characterizing the stationary distribution as the probability of a mutation becomes small is inadequate.

5 Conclusion

We have shown that if players choose their strategies in games by imitating successful players, and if there is a local or neighborhood structure to both the interaction between agents and their learning, then altruistic behavior can survive. The key to the survival of Altruists is that they tend to occur huddled together in concentrated groups. The benefits of the public goods supplied by Altruists are then enjoyed primarily by Altruists. This allows Altruists to earn higher payoffs than Egoists, who tend to be surrounded by other Egoists. The imitation process then induces agents who are close to Altruists to become Altruists, causing the groups of Altruists to expand.

A group of Altruists is always a ripe target for invasion by a mutant Egoist, who will thrive on the public goods provided by the Altruists. For this reason Altruists can survive, but they generally cannot conquer. Instead, the Altruists will be riddled with pockets of Egoists. One might then expect the Egoists introduced by mutations to eventually eliminate Altruists. The pockets of Egoists cannot expand, however, because as they get larger, more and more Egoists are trying to free-ride on nearby Altruists, decreasing the payoffs of the Egoists and inducing imitators to become Altruists. The result is the preservation of altruism in coexistence with egoism.

Our model of agents occupying locations around a circle is very simple. What happens if they are placed in a plane, or in a higher-dimension structure? To gain some insight into these cases, recall that Altruists fare poorly when exposed to too many Egoists, while Egoists fare well when exposed to many Altruists. Taking agents to be arranged along a circle ensures that any group of A's cannot have too many Altruists who are on the boundary and hence are exposed to Egoists, and ensures that any group of Egoists cannot have too many members exposed to Altruists. This in turn produces conditions under which Altruists are likely to thrive. Moving to the plane or to richer spaces raises the possibility that groups of Altruists will appear that are irregularly shaped and that expose virtually all of their members to Egoists. These Altruists may then not survive, and the persistence of Altruism appears to be less certain.

We have been unable to obtain analytical results for such cases, and must appeal to simulations. The extensive simulations of Nowak and May [28, 29, 30] suggest that in the absence of mutations, there are many initial conditions from which a significant proportion of Altruists persist. Once again, Egoists in their model do well in the midst of Altruists while Altruists do poorly in the midst of Egoists, and concentrated groups of Altruists can then expand. The dynamics are much more complicated than in our simple model, but it appears as if Altruistic behavior fares well.

Finally, we have restricted attention to cases in which there are only two strategies. A great deal of work remains to be done in extending the analysis to larger games as well as more complicated spatial structure. It is clear, however, that dynamics driven by imitation can differ significantly from the familiar best-reply dynamics and that imitation coupled with local interactions opens the possibility that altruistic behavior can thrive.

6 Appendix: Proofs

Proof of Proposition 1. It is immediate that the states in which all agents are Altruists or all agents are Egoists are absorbing states because imitation cannot introduce Egoists into a world in which there are only Altruists, or vice versa.

To find the remaining absorbing sets, consider what happens to a string of A's as the imitation dynamics proceed. From (2), any A string of length one immediately disappears. Similarly, if we have an A string of length two, the two A's in this string immediately become E's. In the process, however, the adjacent E's may switch to A's. What happens to these adjacent E's? There are four possibilities. The following transitions describe the fate of the E (the center agent in each case) that initially sits just to the left of the string of two A's. A similar analysis holds for the E on the right. An "x" holds the place of an agent whose type we do not have sufficient information to ascertain.

```
EE E aa      aE E aa      Ea E aa      aa E aa
xE a EE      xE E EE      EE E EE      xE E EE
EE E EE      xx E EE      xE E EE      xx E EE
```

Moreover, the x's in the final line can be A's only if there existed a string of three of more A's to the left of our segment, to which these agents have now become attached. Hence, any A string of length two disappears after two periods without creating any new A strings.

What of A strings that are of length three or longer? From (1)–(2), the A's at the end of such string are the only potential candidates for becoming E's, and the only way that such a string can increase in length is for a single adjacent E at an end to change to A. Hence, such a string may undergo a change in length of $\{-2, -1, 0, 1, 2\}$. Because the string can increase in length only if it borders a segment of three E's (from (1)), the string cannot merge with any other A strings of length three or more. There are then only two possible fates for such a string. It can persist forever as a distinct string, perhaps varying in length, or its length can fall below three at some point, causing it to be eliminated within the next two periods without giving birth to new strings. We thus have that strings of A's can be destroyed but cannot be created.

Together, these results give: *There exists a time τ such that the number of A strings at time τ is less than or equal to the number of A strings of length three or more at time zero; the number of A strings in any subsequent period is equal to the number at time τ; and all A strings in subsequent periods are length three or longer.*

What can we say about E strings? First, notice that the number of A and E strings must be equal. Next, suppose that time τ has been reached, so that all A strings have length at least three. Then from (1), any E string whose length is more than two declines in length by two, a string of length two retains its length, and a string of length one increases in length by two. Hence, we will eventually have Egoist strings of length two or blinkers, alternating between lengths one and three, but no longer strings, giving: *There exists a time τ' after which the number of E strings is less than the number of E strings in the initial state and is constant, and E strings either remain at length two or alternate between lengths one and three.* This gives Proposition 1.1.

It is now an easy calculation to check that since A strings occur in lengths at least three, and since E strings occur in either length two or alternations between length one and three, that the proportion of A's, if there are to be any A's at all, must be at least 0.6. This gives Proposition 1.2. $\square$

Proof of Proposition 2. We examine the case of five adjacent A's. The remaining two cases in (2.1) are straightforward variations.

We show that a string of A's, whose length is at least five, cannot disappear. In particular, we show that if there exists a string of five A's at time t, then either all five of these agents must also be Altruists at time $t+1$ or they must all be Altruists at time $t+2$. This holds regardless of the strategies played by other agents in the system.

Suppose we have a string of five or more A's bordered on each end by an E. Each of these two E's must have either an A's or E on its other side. This gives us four possibilities to consider. First, suppose each E has an A on its other side. Then from (1)–(2), the system proceed as follows:

```
....aE aaaaa Ea....
....EE EaaaE EE....
....xE aaaaa Ex...
```

As usual, an x holds the place of an agent who may be either an Altruist or an Egoist. For convenience, the original string of five A's is separated by spaces. A similar result clearly holds if the original string contains more than five A's.

Alternatively, one of the E's on the end of the string of A's may have an E on its other side while the other may have an A on its other side. This gives us the following case and its mirror image:

```
....aE aaaaa EE....
....EE Eaaaa xE....
....xE aaaaa xx....
```

Finally, the E's on both ends of the string of A's may be bordered by E's. Then we have:

```
....EE aaaaa EE...
....xx aaaaa xx...
```

In each case, the result is that any string of at least five Altruists must persist, and hence any state with such a string lies in the basin of attraction of an absorbing set in which at least sixty percent of the agents are Altruists. $\square$

Proof of Proposition 4. Let $\mathcal{E}$ denote the state in which all agents are Egoists. Let $\mathcal{A}$ be the state in which all agents are Altruists. Let $\mathcal{X}(n, m)$ denote the collection of absorbing sets with the property that in any state contained in such an absorbing set, at least some agents are Altruists and all agents are Altruists except n strings of Egoists of length two and m blinkers, where $n \geq 0$ and $m \geq 0$. We define $\mathcal{X}(n, m)$ only for values of (n, m) for which $\mathcal{X}(n, m)$ is nonempty. Then $\mathcal{A}$ is the unique element of element in $\mathcal{X}(0, 0)$ and every absorbing set other the $\mathcal{E}$ is contained in some $\mathcal{X}(n, m)$.

It suffices to show that $D(\mathcal{A}) < D(\mathcal{E})$, where D is defined in Lemma 3 of Samuelson [32]. For this, it suffices to show that:

- Three mutations suffice to transform $\mathcal{E}$ into a state in the basin of attraction, under the imitation process, of a state in $\mathcal{X}(n, m)$ for some (n, m).
- Given any absorbing set in $\mathcal{X}(n, m)$ with $(n, m) \neq (0, 0)$, there exists a state in the absorbing set which a single mutation can transform into a state in the basin of attraction, under the imitation process, of an absorbing set in $\mathcal{X}(n', m')$ with $n' + m' < n + m$ or with $n' < n$ and $m' \leq m + 1$.
- Given any state in any absorbing set in $\mathcal{X}(n, m)$ for any (n, m), it takes at least $N/10$ mutations to reach a state in the basin of attraction, under the imitation process, of $\mathcal{E}$.

Recall that N is the size of the population. A state is in the basin of attraction, under the imitation process, of an absorbing set, if the deterministic imitation process (without mutations) leads from the state to the absorbing set.

To establish the first condition, we need only note, from (1)–(2), that if three mutations introduce three adjacent Altruists into state $\mathcal{E}$, then the imitation process will induce the two bordering Egoists to switch to Altruists, yielding a string of five Altruists. Proposition 2 ensures that we then have a state in the basin of attraction $\mathcal{X}(n, m)$ for some (n, m). To establish the second condition, consider an absorbing set S in $\mathcal{X}(n, m)$. If $m > 0$, then we need only choose a state in S which at least one blinker has only one Egoist. A mutation switching this Egoist to an Altruist then produces a state in absorbing set in $\mathcal{X}(n, m - 1)$. Hence, consider an absorbing set in $\mathcal{X}(n, 0)$. Now let a mutation switch an Egoist to an Altruist. The result is an isolated Egoist (that was adjacent to the Egoist affected by the mutation). The next iteration of the imitation process will produce a string of three Egoists. If all Altruist strings are still of length at least three, then we have a blinker and a state in an absorbing set contained in $\mathcal{X}(n-1,1)$. If instead at least one Altruist string is now of length only two, then (from the proof of Proposition 1) that string of Altruists will disappear, while no new string can appear, yielding a state in an absorbing set in $\mathcal{X}(n', m')$ with $n' + m' < n$.

Finally, we calculate a lower bound on the number of mutations required to convert a state in an absorbing set in $\mathcal{X}(n, m)$ into a state in the basin of attraction of $\mathcal{E}$. The mutations must eliminate all of the strings of Altruists in the original state. We first notice that in order to eliminate a string of A's of length k, we must have at least $[k/5]$ — the integral value of $k/5$ — mutations.[25] A lower bound on the number of mutations needed to eliminate all string of A's is then $N/10$, which arises in the case in which there are strings of A's of length 9 (which are the longest that can still be eliminated by a single mutation) with blinkers at the end of the string. For sufficiently large N this number exceeds three, giving the result. □

Proof of Proposition 5. Let the types of players be denoted by 1 and 2. Let each player i have a set of players whom he potentially imitates, called his learning neighborhood, and a set of players with whom he interacts, called his interaction neighborhood. In particular, player i's imitation rule is to adopt the strategy that receives the highest average payoff of the strategies represented in his learning neighborhood.[26] In our model of Altruists and Egoists, the interaction neighborhood of player i included his two nearest neighbors, while his learning neighborhood included these two nearest neighbors and himself. Let N_i^{L1} and N_i^{L2} be the sets of type-1 and

[25] This number is calculated by observing that if an Egoist is placed in the midst of a string of Altruists, the result is a blinker, with three Egoists in the next period. In order to eliminate a string of A's, enough Egoists must be inserted so that after a period has passed and each Egoist given rise to a string of three Egoists, with blinkers possibly also converting the A's at each end of the string into E's, all remaining strings of A of the original string must be at most of length 2. This requires at least $[k/5]$ mutations.

[26] Player i is excluded from his interaction neighborhood and included in his learning neighborhood. In the case of Altruists and Egoists, this simply reflects our choice to measure the costs of Altruism as the net costs, after any benefits of one's own public good provision have been realized.

type-2 players in agent i's learning neighborhood. Let N_i^{I1} and N_i^{I2}) be the sets of type-1 and type-2 players in agent i's interaction neighborhood. Let n_i^{L1}, n_i^{L2}, n_i^{I1}, and n_i^{I2}) be the numbers of players in the sets N_i^{L1}, (N_i^{L2}, N_i^{I1}, and N_i^{I2}). Let P_i be player i's payoff, and let $n_i^{I1} + n_i^{I2} = n^I$ be the size of the interaction neighborhood. According to the imitation rule, player i will become type 1 iff:

$$\frac{1}{n_i^{L1}} \sum_{j \in N_i^{L1}} P_j > \frac{1}{n_i^{L2}} \sum_{j \in N_i^{L2}} P_j \tag{11}$$

When player j is of type s, then his payoff is given by

$$P_j = K_1 n_j^{I1} + K_2 n_j^{I2} - C_s = (K_1 - K_2)\, n_j^{I1} + K_2 n^I - C_s,$$

and the imitation rule (11) becomes:

$$\frac{K_1 - K_2}{n_i^{L1}} \sum_{j \in N_i^{L1}} n_j^{I1} > (C_1 - C_{2)} + \frac{K_1 - K_2}{n_i^{L2}} \sum_{j \in N_i^{L2}} n_j^{I2}$$

It is now obvious that when $K_1 - K_2 > 0$ the dynamics will be identical for all pairs of types (K_s, C_s) for which $\frac{C_1 - C_2}{K_1 - K_2}$ is the same. $\square$

Proof of Proposition 9. Let $\mathcal{A}$ be the absorbing state in which all agents are Altruists, $\mathcal{E}$ the absorbing state in which all are Egoists, and $\mathcal{X}(n, m, p)$ the collection of absorbing sets with the property that each state in an absorbing set in $\mathcal{X}(n, m, p)$ contains states in which at least some agents are Altruists, contains n strings of Egoists of length three, m one/five blinkers, and p two/six blinkers. We define $\mathcal{X}(n, m, p)$ only for cases where it is nonempty.

We shall show that $D(\mathcal{E}) < D(S)$ for any absorbing set S in which Altruists survive. Proceeding analogously to the proof of Proposition 4, it is easy to first show that for any absorbing set in $\mathcal{X}(n, m, p)$, there exists a state and a single mutation that *(i)* yields a state in the basin of attraction of $\mathcal{X}(n, m+1, p-1)$ if $p > 1$, *(ii)* yields a state in the basin of attraction of $\mathcal{X}(n, m-1, 0)$ if $m > 0$ and $p = 0$, and *(iii)* yields a state in the basin of attraction of either $\mathcal{X}(n', m', p)$ with $n' + m' < n$ or with $n' < n$ and $m' = 1$ if $n = m = 0$.

It then suffices for the result to show that *(i)* the number of mutations required to transform $\mathcal{E}$ into a state in the basin of attraction of an absorbing set in which Altruists survive is four, and *(ii)* we can construct a sequence of absorbing sets $(\mathcal{A}, S_1, \ldots, S_k, \mathcal{E})$ such that each absorbing set contains a state that can be transformed into a state in the basin of attraction of the succeeding absorbing set with a single mutation. The first of these statements follows directly from (8)–(9). Mutations that introduce four adjacent Altruists create a single string of Altruists that the imitation process causes to grow until an absorbing set is reached in which all agents are Altruists except possibly a string of three Egoists or a single blinker of Egoists. To verify the second statement, we note that a sequence of single mutations, each of which creates a one/five blinker, can lead through a sequence of absorbing sets from $\mathcal{A}$ to an absorbing set in which blinkers, which alternate between strings of one and five Egoists, are packed out of phase next to each other (so that if a blinker has five Egoists in a given state, its adjacent blinkers have one Egoists each), with exactly three Altruists between. An example of such an absorbing set is shown in (10). As (10) shows, the length of two out-of-phase blinkers and their Altruists buffers is twelve. The circle will then be composed of units of twelve agents, each containing two blinkers. Depending upon the length of the circle, there will also be a string of leftover Altruists that is too short to contain two out-of-phase blinkers. The length of this string, denoted r, satisfies $0 \leq r \leq 11$, since a length longer than eleven would allow mutations to create two additional blinkers.

The proof can then be completed by showing that for each value of r, $0 \leq r \leq 11$, a single mutation placing an Egoist in the midst of the string of leftover Altruists yields a state in the basin of attraction of $\mathcal{E}$, i.e., leads the imitation process to eliminate all Altruists. We will show that such a result holds, regardless of the length of the circle, though the number of steps required to reach the state of all Egoists depends on N.

There are 12 cases to consider. The proof now proceeds by brute force, identifying for each of the twelve cases the position in which a mutant Egoists must appear in order to prompt the imitation process to eliminate Altruists. Here, we show one case. We let $N = 16$. We begin with an absorbing set containing two blinkers and a string of four extra Altruists. This absorbing set is pictured in the first three lines of (12) below. In the fourth line, the imitation process calls for the final agent to remain an Altruist, but a mutation switches this agent to an E (depicted below as a bold letter). The imitation dynamics then destroys all Altruists in three steps.

$$
\begin{array}{l}
\texttt{aaaEaaaEEEEEaaaa} \\
\texttt{aEEEEEaaaEaaaaaa} \\
\texttt{aaaEaaaEEEEEaaaa} \\
\texttt{aEEEEEaaaEaaaaa}\mathbf{E} \\
\texttt{EEEEaaaEEEEEaEEE} \\
\texttt{EEaaEEEaaEEEEEEE} \\
\texttt{EEEEEEEEEEEEEEEE}
\end{array}
\tag{12}
$$

The proof consists of showing that this can be done for each of the 12 cases.

To conclude that $D(\mathcal{E}) < D(S)$ [27] for any absorbing set in which Altruists survive, note that $D(S)$ must be at least as large as the number of absorbing sets plus three. In particular, every absorbing set contributes at least one to $D(S)$, since only with a mutation can one leave an absorbing set, and the absorbing set E contributes four, since only with four mutations can one move from $\mathcal{E}$ to a state outside the basin of attraction of $\mathcal{E}$. We then note that $D(\mathcal{E})$ equals the number of absorbing sets, since we have shown that a path from $\mathcal{A}$ to $\mathcal{E}$ can be constructed consisting of absorbing sets, with only a single mutation required to make each transition, and that we can move form all other absorbing sets other than $\mathcal{E}$ to $\mathcal{A}$ through a sequence of absorbing sets connected by single-mutation links.

Notice, however, that moving from $\mathcal{E}$ to a state outside the basin of attraction of $\mathcal{E}$ requires four mutations that are adjacent, but can occur anywhere in the population. In contrast, the sequence that leads from $\mathcal{A}$ to $\mathcal{E}$ requires a quite special combination of mutations. If we fix the mutation rate and let the size of the population grow, then this special combination of mutations becomes increasingly unlikely, and eventually becomes less likely that the mutations required to escape the basin of attraction of $\mathcal{E}$. Hence, the limiting result does not survive reversing the order of limits. $\square$

References

[1] James Andreoni. Impure altruism and donations to public goods: A theory of warm- glow giving. *Economics Journal*, 100:464–477, 1990.

[27] see lemma 3 of Samuelson [32]

[2] W.W. Benson. Evidence for the evolution of the unpalatability by kin selection in the heliconiinae (lepidoptera). *American Naturalist*, 105:213–226, 1971.

[3] Lawrence E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5:387–424, 1993.

[4] Robert Boyd and Peter J. Richerson. *Culture and the Evolutionary Process*. University of Chicago Press, Chicago, 1985.

[5] L.P. Brower. Ecological chemistry. *Scientific American*, 220:22–29, 1969.

[6] L.P. Brower and J.V. Brower. Birds, butterflies and plant poisons: A study in ecological chemistry. *Zoologica*, 49:137–159, 1964.

[7] J.J. Christian. Social subordination, density and mammalian evolution. *Science*, 168:84–90, 1970.

[8] D. Cohen and I. Eshel. On the founder effect and the evolution of altruistic traits. *Theoretical Population Biology*, 10:276–302, 1976.

[9] Charles Darwin. *The Descent of Man and Selection in Relation to Sex*. Murray, London, 1871.

[10] R. Dawkins. *The Selfish Gene*. Oxford University Press, Oxford, 1976.

[11] Glenn Ellison. Learning, local interaction, and coordination. *Econometrica*, 61:1047–1072, 1992.

[12] I. Eshel. On the neighbor effect and the evolution of altruistic traits. *Theoretical Population Biology*, 3:258–277, 1972.

[13] I. Eshel and L.L. Cavalli-Sforza. Assortment of encounters and evolution of cooperativeness. *Proc. Nat. Acad. Sci. USA*, 79:1331–1335, 1982.

[14] I. Eshel, E. Sansone, and A. Shaked. Evolutionary dynamics of populations with a local interaction structure. Sfb discussion papers, Bonn University, Germany, 1995.

[15] F. Fenner. Myxoma virus and cryctologus cuniculus: Two colonizing species. In H.G. Buker and G.I. Stebbins, editors, *The Genetics of Colonizing Species*, pages 485–499. Academic Press, New York, 1965.

[16] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems.* Springer–Verlag, New York, 1984.

[17] Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting and incomplete information. *Econometrica*, 54:533–554, 1986.

[18] Itzhak Gilboa and David Schmeidler. Case–based decision theory. Mimeo, Northwestern University, 1992.

[19] Itzhak Gilboa and David Schmeidler. Case–based optimization. Mimeo, Northwestern University, 1993.

[20] W.D. Hamilton. The genetic evolution of social behavior. *Journal of Theoretical Biology*, 7:1–52, 1964.

[21] W.D. Hamilton. Altruism and related phenomena. *Annual Review of Ecological Systems*, 3:193–232, 1972.

[22] Michihiro Kandori. Social norms and community enforcement. *Review of Economic Studies*, 59:61–80, 1991.

[23] Michihiro Kandori, George J. Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61:29–56, 1993.

[24] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains.* D. Van Nostrand Company, Inc., Princeton, 1960.

[25] K. Lorenz. *On Aggression.* Harcourt, Brace and World, 1963.

[26] C. Matessi and S.D. Jayakar. Condition for the altruism under darwinian selection. *Theor. Pop. Biol.*, 9:360–387, 1975.

[27] J. Maynard Smith. Group selection. *The Quarterly Review of Biology*, 51:277–283, 1976.

[28] Martin A. Nowak, Sebastian Bonhoeffer, and Robert M. May. More spatial games. *International Journal of Bifurcation and Chaos*, 4:33–56, 1994.

[29] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

[30] Martin A. Nowak and Robert M. May. The spatial dilemmas of evolution. *International Journal of Bifurcation and Chaos*, 3:35–78, 1993.

[31] Masahiro Okuno-Fujiwara and Andrew Postlewaite. Social norms in random matching games. Mimeo, University of Tokyo and University of Pennsyvlania, 1989.

[32] Larry Samuelson. Stochastic stability in games with alternative best replies. *Journal of Economic Theory*, 64:35–95, 1994.

[33] F. M. Stewart and B.R. Levin. The population biology of bacterial viruses: Why be temperate? *Theoretical Population Biology*, 26:93–117, 1984.

[34] G. C. Williams. *Adaptation and Natural Selection.* Princeton University Press, Princeton, 1966.

[35] D.S. Wilson. A theory of group selection. *Proc. Nat. Acad. Sc. USA*, 72:143–146, 1975.

[36] D.S. Wilson. Altruism in Mendelian populations derived from sibling groups: The haystack model revised. *Evolution*, 41:1059–1070, 1987.

[37] I. Wood. Altruism among bacteria. In I.C. Gunzalus and R.Y. Stainer, editors, *The Bacteria: Vol. II*, pages 83–93. Academic Press, New York, 1961.

[38] S. Wright. Population structure in evolution. *Proceedings of the American Philosophical Society*, 93:471–478, 1949.

[39] V. C. Wynne-Edwards. *Animal Dispersion in Relation to Social Behavior.* Oliver and Boyd, Edinburgh, 1962.

[40] V. C. Wynne-Edwards. *Evolution Through Group Selection.* Blackwell Scientific Publications, Oxford, 1986.

[41] Peyton Young. The evolution of conventions. *Econometrica*, 61:57–84, 1993.