

Discussion Papers in Economics

Understanding Reciprocity

Rajiv Sethi and E.Somanathan

Discussion Paper 02-05

Revised 2002



Indian Statistical Institute, Delhi
Planning Unit
7 S.J.S. Sansanwal Marg, New Delhi 110 016, India

Understanding Reciprocity*

Rajiv Sethi[†]

E. Somanathan[‡]

Forthcoming: *Journal of Economic Behavior and Organization*

Abstract

This paper surveys the evolutionary game theoretic literature on reciprocity in human interactions, dealing both with long-term relationships and with sporadic interactions. Four basic themes, repetition, commitment, assortment, and parochialism, appear repeatedly throughout the literature. Repetition can give rise to the evolution of behavior that exhibits reciprocity-like features but a vast array of other behaviors are also stable. In sporadic interactions, reciprocity can be stable if the propensity to punish selfish actions can induce opportunists to cooperate, if reciprocators themselves behave opportunistically when they expect others to do so, or if matching is sufficiently assortative.

JEL Classification: C72, D62.

Keywords: Reciprocity, Evolution, Assortation, Commitment, Parochialism.

*We thank Duncan Foley, Dan Friedman, Herb Gintis, Joel Guttman, Paul Rubin, and an anonymous referee for their comments and suggestions, and gratefully acknowledge financial support from the National Science Foundation.

[†]Department of Economics, Barnard College, Columbia University, 3009 Broadway, New York, NY 10027.

[‡]Planning Unit, Indian Statistical Institute, New Delhi 110016, India.

1 Introduction

Reciprocity is a pervasive and economically significant phenomenon in human interactions. The propensity of individuals to reward generosity and punish opportunism is encountered frequently both in long-term relationships and in sporadic interactions, including brief and anonymous interactions in experimental environments.¹ Two widely replicated groups of experiments provide striking examples. In ultimatum and alternating-offer bargaining games, the costly rejection of unfair offers suggests a willingness to incur material losses to inflict harm on others who are perceived as being opportunistic (Güth et al., 1982, Ochs and Roth, 1989). In trust or gift-exchange games, the tendency to incur material losses to reward others who are perceived as being generous provides further evidence for reciprocity (Fehr et al., 1993, Berg et al., 1995). In many cases, such behavior is in conflict with the twin hypotheses of rationality and material self-interest which are the foundation of orthodox economic theory, and raises the question of how this behavior may have emerged and persisted in evolutionary competition with purely opportunistic or self-serving behavior.

There is now a significant body of literature in evolutionary game theory which sheds some light on this question. The hallmark of the evolutionary approach is the hypothesis that in a heterogeneous population, more successful traits will survive and spread at the expense of less successful ones. When the traits in question are transmitted genetically, the success of a trait refers simply to the reproductive fitness of its carrier. However, much of the literature on the evolution of reciprocity treats the biological model simply as a metaphor, interpreting the dynamics in terms of cultural transmission. Under this interpretation, traits are transmitted through processes of imitation and learning, and traits that bring their bearer higher material or monetary payoffs are replicated faster. Within this broad framework one finds a variety of models which differ along several dimensions. Interactions may be repeated or sporadic, matching may be random or assortative, and individuals may be endowed with programmed behavioral strategies or may choose in a calculated manner on the basis of well-defined preferences. It is the purpose of this survey to examine, within a unified framework, selected contributions to this literature.

The most favorable environment for the evolution of reciprocity is in repeated interactions, and the prototypical model of this kind is the infinitely repeated prisoners' dilemma. As is well known from the early work of Axelrod and Hamilton (1981), a 'tit-for-tat' strategy involving conditional cooperation can be a stable outcome of an evolutionary process in

¹For a recent survey of the vast and rapidly accumulating experimental literature on reciprocity, and a discussion of the economic implications of this work, see Fehr and Gächter (2000).

this model. However, it is also the case that very counter-intuitive and inefficient outcomes can also be stable outcomes of the same evolutionary process. The criterion of evolutionary stability alone is an insufficiently sharp predictor of outcomes in repeated games. Attempts to obtain sharper predictions have accordingly required some modification of the benchmark model. One such modification is based on the idea that complexity is costly: given any two strategies which are equally successful in the repeated interaction, the simpler one is assumed to have an evolutionary advantage. An alternative modification is to allow for the possibility that there may be errors in the implementation of a strategy, leading occasionally to an action being taken that differs from that which a strategy prescribes. Each of these approaches is successful in sharply narrowing the set of outcomes that are consistent with evolutionary stability. Strategies involving some form of conditional cooperation emerge as stable outcomes in these models, although in neither case does the simple and intuitively appealing ‘tit-for-tat’ strategy survive.

When interactions are sporadic or occur over a short horizon, the conditions under which reciprocity can be a stable outcome of an evolutionary process are more stringent. Reciprocity can survive if interactions are sufficiently assortative, so that opportunistic individuals cannot fully exploit those with a propensity for conditional cooperation. Such assortative interaction may itself arise endogenously through a process of conscious choice, with all individuals actively seeking to avoid contact with opportunists. Alternatively, reciprocity can survive if those with a propensity to punish opportunistic behavior can be recognized to possess this trait. This credible commitment to retaliate induces opportunists to behave cooperatively in interactions with reciprocators, though not in interactions with each other. Reciprocity can also survive if some individuals behave parochially, acting altruistically when encountering others like themselves but opportunistically or spitefully when interacting with opportunists. Each of these explanations requires that the propensities or preferences with which individuals are endowed are observable, at least in the probabilistic sense that truly opportunistic types are believed to be opportunistic with greater likelihood than are non-opportunistic types. The assumption of observability can be dropped in explanations of reciprocity relying on commitment only at the cost of weakening the concept of stability that is used. When observability is imperfect, the stable outcome often involves a mixture of types, some of whom are opportunistic. This is consistent with the experimental evidence which identifies considerable behavioral heterogeneity among subjects, and the presence in the population of some who appear to be maximizing their material self-interest.

Repetition, assortment, commitment and parochialism are four basic themes which oc-

cur, separately or in combination, in each of the contributions surveyed below. We begin in Section 2 with the case of reciprocity in long-term relationships. Sections 3 and 4 deal with reciprocity in sporadic interactions. The discussion in Section 5 deals with questions that remain open and speculates on fruitful directions for future research. In many cases the results surveyed are discussed within the context of specific examples, instead of being presented in their full generality. This enables us to survey a sometimes technical and specialized literature in a manner that should be accessible to a broad range of researchers across disciplinary boundaries. The survey is not meant to be exhaustive, but rather to identify basic methodological themes and important strands in the literature on understanding the origins and persistence of reciprocity in human interactions.

2 Reciprocity in Relationships

We begin with the case of long-term relationships in which the interaction does not have a fixed horizon. In such situations reciprocal behavior can be fully consistent with rationality and material self-interest. However, it is typically the case that a vast array of other behaviors are also consistent with rationality and material self-interest. This raises the question of whether one might expect reciprocity, rather than these other behaviors, to emerge from the process of evolutionary competition.

An example of an interaction which will be used repeatedly for illustrative purposes below is the following version of the symmetric two-player prisoners' dilemma. Each of two individuals chooses (without knowledge of the other's choice) one of two actions, 'cooperate' and 'defect'. Cooperation entails a private cost $\alpha > 0$ to the cooperator, and yields a benefit $\beta > \alpha$ to the *other* individual. Defection involves no cost and yields no benefit. The following matrix shows the material payoffs to the row player, given the actions taken by both players.

	C	D	
C	$\beta - \alpha$	$-\alpha$	(1)
D	β	0	

If two self-interested individuals are matched to play this game exactly once, they can both be predicted to defect since defection yields a higher payoff regardless of the action taken by one's opponent. This is clearly inefficient since the outcome of mutual cooperation is strictly preferred by both players.

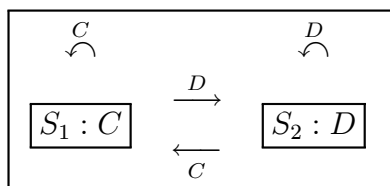
Suppose, instead, that the game is infinitely repeated (or alternatively, that there is no fixed bound on the number of repetitions, with a given probability of termination in each

period). A repeated game strategy for a given player specifies, for each history of past play, an action to be taken by the specified player following that history. Since there is no bound to the number of stages in the game, such strategies can depend in complex ways on arbitrarily long histories. We shall consider only strategies that are of finite complexity in a sense to be made precise below.

2.1 Finite Complexity

A strategy in a repeated game is said to be *finitely complex* if it can be implemented by an automaton with a finite number of states. An automaton is described by the following four components: a set of states, an output function which indicates which action is taken in each state, an initial state which the machine occupies at the start of the game, and a transition function which indicates which state is reached in the next period given the current state and the current actions of all other players. The simplest examples are the one-state automata ‘always cooperate’ and ‘always defect’ which adhere to the same action in each period regardless of history.

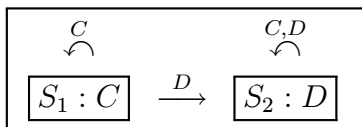
Slightly more complex is the ‘tit-for-tat’ strategy which begins by playing ‘cooperate’ and subsequently takes whichever action its opponent took in the previous period. Such a strategy entails ‘reciprocal altruism’ in the sense of Trivers (1971). ‘Tit-for-tat’ is clearly of finite complexity, and can be represented by the two-state automaton described in the following diagram:



The diagrams depicting finitely complex strategies are to be read as follows. S_1 denotes the initial state and S_i the i th state. The action to be taken in a state follows the colon after the state. For example, $S_1 : C$ indicates cooperation in state S_1 . The transition function is indicated by arrows from one state to another, each arrow being associated with the actions of the *other* player that cause that particular transition. Hence ‘tit-for-tat’ prescribes a move to state S_2 after a defection by one’s opponent and a return to state S_1 after cooperation by one’s opponent.

When matched against itself or against the unconditionally cooperative strategy, both ‘tit-for-tat’ and its opponent earn $\beta - \alpha$ in each stage. When matched against the unconditionally defecting strategy, ‘tit-for-tat’ yields $-\alpha$ in the initial period (while ‘always

defect' yields β); in all subsequent periods both strategies yield 0. Another two-state strategy is 'grim', which begins by cooperating and responds to a single defection by defecting perpetually thereafter:



Note that 'grim' is behaviorally indistinguishable from 'tit-for-tat' when matched against any strategy that is never the first to defect; it is also indistinguishable from 'tit-for-tat' when matched against 'always defect'.

In the case of infinitely repeated games, any particular path of play gives rise to a potentially infinite sequence of payoffs and there are several criteria on the basis of which such payoff sequences can be compared. One of these is the limit-of-the-means or limit-average criterion, according to which each player receives the limit as $T \rightarrow \infty$ of the mean payoff obtained over the first T periods. When strategies are of finite complexity, such a limit always exists. We shall assume that this condition of finite complexity holds, and that the evolutionary success of a strategy depends on the limit-average payoff that it yields. Strategies such as 'tit-for-tat', 'grim' and 'always cooperate' yield the maximum possible combined limit-average payoff when matched against themselves. Such strategies are said to be *efficient*.

A central question of interest is whether or not efficient strategies such as 'tit-for-tat' can survive under evolutionary competition with other strategies and if so, whether such strategies are the only ones to survive. In an influential attempt to address this question, Axelrod (1984) invited a number of researchers to submit computer programs to implement behavioral strategies in the repeated prisoner's dilemma and allowed the programs to compete in a round-robin contest in which each program was matched pairwise with the other programs in the program population. The population composition of programs was itself modified over time, with the population share of more successful programs rising relative to that of less successful ones. The simple program which implements the 'tit-for-tat' strategy won the contest (by having the largest population share at the end of the tournament). A second such tournament resulted in the same winner. Partly in response to Axelrod's simulation results, several attempts have been made to provide an analytical basis for the claim that efficient equilibria involving reciprocity will be selected through some sort of evolutionary process. This literature is based on a particular notion of evolutionary stability, which is made precise below.

2.2 Evolutionary Stability

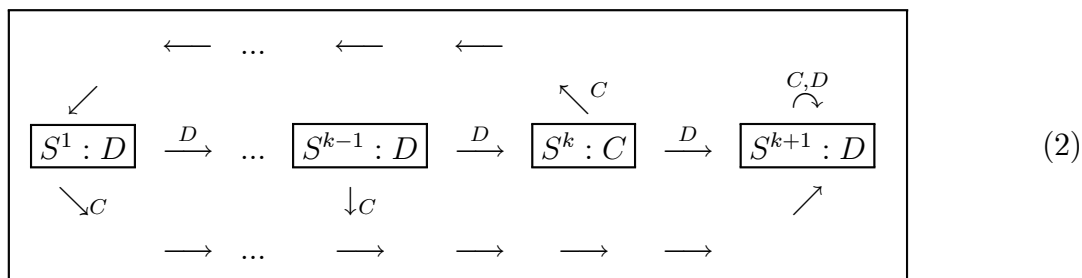
Consider a large population, the members of which are randomly matched in pairs to play the repeated prisoners' dilemma. In each such interaction, the expected (limit-average) payoff to each strategy present in the population is simply the weighted average of its payoffs against all other strategies, with the population proportions of each strategy determining the weights. Suppose that the population state consists largely of some 'incumbent' strategy together with a small proportion of some 'mutant' strategy. The incumbent is said to be an evolutionarily stable strategy (ESS) if, for all possible mutant strategies, the incumbent earns strictly greater payoffs than the mutant when the population share of the latter is sufficiently small. More precisely, an incumbent strategy is an ESS if, for every mutant strategy, either (i) the mutant earns a lower payoff against the incumbent than the incumbent does against itself, or (ii) the mutant earns the same payoff against the incumbent than the incumbent does against itself and the mutant earns a lower payoff against itself than the incumbent does against the mutant (Maynard Smith and Price, 1973). The interpretation of these conditions is that a population consisting largely of some incumbent strategy can expel any mutant strategy which appears in sufficiently small numbers relative to the size of the population as a whole.

It is easily seen that none of the strategies discussed above for the iterated prisoners' dilemma is an ESS, and that, in fact, no ESS exists. The non-existence of an ESS is typical in repeated games and more generally in games with a sequential structure.² This suggests the following weakening of the evolutionary stability concept. Given some incumbent strategy, we say that a mutant can invade if (i) either the mutant does better against the incumbent than the incumbent does against itself, or (ii) the mutant does as well against the incumbent as the incumbent does against itself, but better against itself than the incumbent does against it. A strategy that cannot be invaded is a neutrally stable strategy or NSS (Maynard Smith, 1982). Note that a population consisting of a neutrally stable strategy does not necessarily expel all rare mutants; in general there will be certain mutants that can survive (without spreading or shrinking) within the incumbent population.³

²This problem arises because any candidate ESS in a game with a sequential structure will typically be behaviorally indistinguishable from at least one mutant strategy in populations consisting only of these two strategies. For example, 'tit-for-tat' is behaviorally indistinguishable from 'always cooperate' in populations consisting only of these two strategies. In this case the mutant cannot be expelled since its payoff is identical to that of the incumbent.

³Note that all evolutionarily stable strategies are also neutrally stable, and that all neutrally stable strategies are Nash equilibria of the (stage) game. Hence the ESS and NSS criteria can be interpreted as

Axelrod and Hamilton (1981) showed that the strategy of perpetual defection could be invaded by a ‘tit-for-tat’ mutant under the limit average criterion and was therefore not neutrally stable; they also showed that ‘tit-for-tat’ is itself neutrally stable. This has been interpreted as theoretical support for the hypothesis that cooperation sustained by reciprocity is an inevitable outcome of the evolutionary process. It turns out, however, that there are a vast number of strategies that are also neutrally stable, including some which prescribe defection in almost every period. Fudenberg and Maskin (1990) give the following example of a strategy which is neutrally stable for any positive integer k : cooperate in any period which is a multiple of k and defect in all other periods, provided that past play has conformed to this pattern; otherwise defect in all periods. For $k = 1$ this strategy sustains complete cooperation when matched against itself. As k gets large, the strategy approaches one of perpetual defection. The strategy is finitely complex for any k and can be represented by the following $k + 1$ state automaton.



This example illustrates that while reciprocity can indeed be an outcome of an evolutionary selection process it is by no means an inevitable outcome. Outcomes which approximate the worst-case scenario of perpetual defection can also be stable, as can a vast array of outcomes between these two extremes.⁴ The criterion of neutral stability therefore rules out perpetual defection but very little else in repeated interactions. Coupled with either the presence of complexity costs or the possibility of mistakes in strategy implementation, however, neutral

refinements of the static concept of Nash equilibrium. They also have a dynamic interpretation. If the growth rate of each strategy’s population share is proportional to its payoff (the replicator dynamics), an NSS corresponds to a stable rest point of the dynamics, while an ESS is an asymptotically stable rest point. (Roughly speaking, stability of a rest point means that trajectories originating nearby do not move away, while asymptotic stability imposes the further condition that trajectories originating near the rest point converge to it.) There may, however, exist asymptotically stable rest points of the dynamics that are not evolutionarily stable strategies; an example based on Rosenthal (1997) is given below.

⁴This multiplicity of neutrally stable outcomes is a reflection of the ‘folk theorems’ establishing multiplicity of Nash equilibria in infinitely repeated games. Although neutral stability refines the set of Nash equilibria, it does not do so appreciably. As is well known, the same applies to other refinements such as subgame perfection (Fudenberg and Maskin, 1986).

stability yields surprisingly sharp predictions.

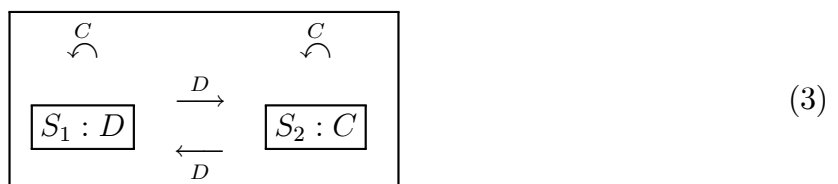
2.3 Costly Complexity

One approach to introducing complexity costs into the analysis is to suppose that when two strategies yield identical payoffs, the less complex strategy has a fitness advantage. Complexity costs become relevant only in the case of equal payoffs: when two strategies earn unequal payoffs, we continue to assume that the one with the higher payoff has a fitness advantage regardless of any differences in complexity. The level of complexity of a strategy is simply the number of distinct states in the corresponding automaton.

Binmore and Samuelson (1992) adopt this specification of costly complexity and look for population states which satisfy a modified version of neutral stability. As above, consider a population consisting largely of some incumbent strategy together with a small proportion of some mutant strategy. In the Binmore and Samuelson model a mutant can invade if either (i) the mutant does better against the incumbent than the incumbent does against itself, or (ii) the mutant does as well against the incumbent that the incumbent does against itself, but better against itself than the incumbent does against it, or (iii) the mutant does as well against the incumbent that the incumbent does against itself, as well against itself than the incumbent does against it, and is less complex. A strategy that cannot be invaded is a modified neutrally stable strategy.

Note that any strategy which does not require the use of all states when matched against itself cannot be a modified NSS, since it can be invaded by an otherwise identical mutant obtained by deleting the unused state and arbitrarily ‘rewiring’ all paths leading to that state. In particular, the class of strategies described in (2) above cannot be stable since state S^{k+1} is never entered when this strategy is matched against itself. Similarly, ‘tit-for-tat’ is not a modified NSS, because the less complex strategy which always cooperates can invade. In fact no strategy which is “nice” in the sense of never being the first to defect can be a modified NSS since it must be either the one-state cooperator, or it must have some redundant states which can be dropped without affecting its payoffs.

As an example of an efficient strategy that is a modified NSS (provided that $\beta > 2\alpha$), consider the following:



Binmore and Samuelson show that in any two person game, any modified NSS must be efficient in the sense that the combined payoff obtained when such a strategy is matched against itself is the maximum attainable combined payoff. Their result is based on the following argument. Suppose there exists an incumbent strategy which is a modified NSS but which is not efficient. Since complexity is costly, this strategy must use all of its states, so it must enter each state at least once. Consider a mutant which takes an initial action that differs from that taken by the incumbent. If its opponent also takes this action (thereby identifying itself as being a mutant), then the mutant switches to the initial state of some efficient strategy and mimics this efficient strategy thereafter. If, instead, its opponent takes an initial action that is identical to that which the incumbent takes, the mutant switches to the state which the incumbent would switch to when encountering the action taken by the mutant, and imitates the incumbent thereafter. Although this mutant will generally be more complex, it obtains the same payoff against the incumbent as the incumbent does against itself, and it earns a strictly higher payoff against itself than the incumbent does against it. It can, therefore, invade.⁵

The use of evolutionary stability as an equilibrium refinement in a model of costly complexity allows Binmore and Samuelson to obtain a sharp characterization of stable outcomes. While the analysis shows that evolutionary stability predicts efficiency in the presence of complexity costs, it also shows that a ‘nice’ strategy such as ‘tit-for-tat’ cannot be stable. Stable strategies such as (3) *defect* at first contact as a means of identifying each other, and then cooperate thereafter, conditional on cooperation by their opponent. This leaves an uncomfortable gap between the theory and the empirical results on reciprocal behavior, a point to which we return below.

An alternative approach to incorporating complexity costs into the analysis is to treat such costs as ongoing deductions from the stage game payoffs. Models of this kind have been studied by Hirshleifer and Martinez Coll (1988) and Rosenthal (1997). Rosenthal groups together all strategies that cooperate initially, maintain cooperation against unconditional cooperators, and defect after the first period against unconditional defectors under one umbrella, referring to them as strategies that ‘trust but verify’. Such strategies include ‘tit-for-tat’, ‘grim’ and a variety of others. In populations consisting only of unconditional cooperators, unconditional defectors, and ‘trust-but-verify’ strategies, all ‘trust-but-verify’ strategies are behaviorally indistinguishable from each other in all interactions and therefore

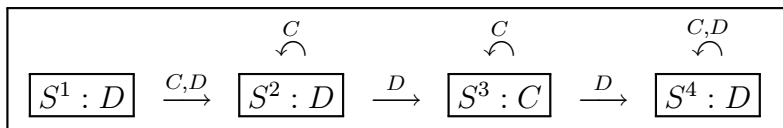
⁵Notice that the initial action of the mutant serves as a signal by which it identifies itself to others of its kind. The importance of such signals for achieving cooperation is discussed further below.

earn the same payoff. It is assumed the additional complexity of such strategies involves an ongoing cost κ which is deducted from their payoff in each period. For the prisoners' dilemma, the limit-average payoffs are therefore given by

	C	D	T
C	$\beta - \alpha$	$-\alpha$	$\beta - \alpha$
D	β	0	0
T	$\beta - \alpha - \kappa$	$-\kappa$	$\beta - \alpha - \kappa$

where T denotes a ‘trust-but-verify’ strategy.⁶ Starting with some arbitrarily given population composition in which all three types are present, suppose that the population composition evolves under pressure of differential payoffs such that the population share of more highly rewarded strategies grows more rapidly than that of less highly rewarded strategies. Under any such dynamics, it is easily seen that a population of unconditional defectors will be asymptotically stable. It can be shown, however, that under some such dynamics (including the commonly studied replicator dynamics) there may be an additional asymptotically stable population state in which all three types are present. Such a state exists if κ is sufficiently small. While the population share of ‘trust-but-verify’ strategies as a whole is determined by the condition that all three strategies earn the same payoff at the interior stable state, the specific ‘trust-but-verify’ strategies of which this group is composed remains indeterminate. Since all such strategies earn the same payoffs, however, the indeterminacy is not consequential.

This model seems more favorable to the evolution of reciprocity via repetition than those based on lexicographic complexity costs. However, the argument relies on the assumption that only the three specified classes of strategies are present in the population. Allowing for mutants that are outside this class undermines both the asymptotic stability of the more efficient state and the irrelevance of the specific ‘trust-but-verify’ strategies that are present in the population, provided that complexity costs remain small relative to total payoffs even for somewhat more complex strategies. For instance, if the only ‘trust-but-verify’ strategy in the population were the ‘tit-for-tat’ type, the following strategy could invade a population which was at the interior state considered above



⁶Rosenthal considers the discounted sum of payoffs rather than the limit average, but the difference is not consequential for the present discussion.

This strategy begins by defecting twice, then continues to defect as long as its opponent continues to cooperate. If the opponent defects, it switches to the ‘grim’ strategy of conditional cooperation. Such a strategy is strictly superior to ‘tit-for-tat’ in any population in which unconditional cooperators are present together with unconditional defectors and ‘tit-for-tat’ types. If its additional complexity does not come at too great a cost, it will outperform all three strategies at Rosenthal’s interior steady state.

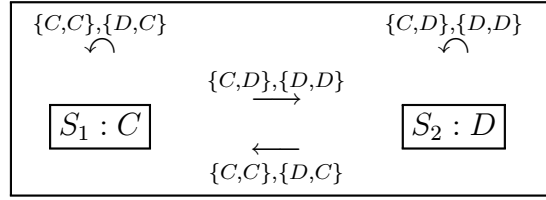
Including complexity costs explicitly in the payoff function causes a population of perpetual defectors to become asymptotically stable under payoff monotonic selection dynamics, so efficiency is no longer an inevitable outcome. If complexity is not too costly, there may exist other asymptotically stable states, for instance consisting of the strategy (3). The introduction of complexity costs directly into the payoff function, therefore, erodes the sharpness of the predictions obtained under the somewhat less appealing hypothesis of lexicographic costs.

2.4 Implementation Errors

Another method of selecting among stable outcomes in certain games (including the prisoners’ dilemma) has been found by Fudenberg and Maskin (1990), who disregard complexity costs but allow for the fact that there may be a small probability of errors in the implementation of a strategy, so that the action a player intends to take does not match the one that is actually taken. As in Binmore and Samuelson (1992), only strategies of finite complexity are considered. It is assumed that if two strategies yield the same payoff conditional on n errors, for all $n = 0, 1, \dots, m - 1$, then the one which yields a higher payoff conditional on m errors has the higher fitness. In particular if two strategies yield the same payoff conditional on no errors, then the one which yields a higher payoff conditional on a single error has the higher fitness. An incumbent strategy is stable in this framework if it is at least as fit as all other mutant strategies in a population consisting almost entirely of the incumbent strategy. This stability notion generalizes neutral stability to allow for the presence of rare mistakes which cause intended and realized actions to differ.

In the presence of implementation errors, it is necessary to specify not only how a strategy responds from a given state to the action taken by its opponent, but also how it responds in the case of an error on its own part. The ‘tit-for-tat’ strategy, which takes whichever action

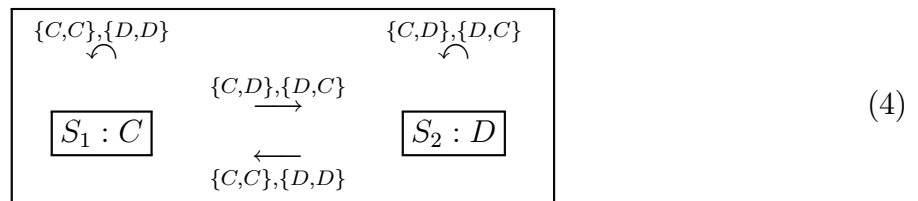
its opponent took in the previous period, can be represented as follows.⁷



Note that a single (unintended) defection can trigger an infinite sequence of alternating cooperation and defection which is clearly inefficient. Fudenberg and Maskin show that in a class of games which includes the prisoners' dilemma, strategies which satisfy their definition of evolutionary stability cannot have this property: they must result in efficient play after *every* history.

To obtain some intuition for their argument, consider any history of play in which both players have cooperated in all periods except the last period, in which exactly one player has defected. Suppose that the incumbent strategy results in inefficient play after this history (conditional on no subsequent errors). Then a mutant which is identical to the incumbent strategy *except after this particular history* can invade, provided that the mutant switches to efficient play after this history when matched against itself, and successfully mimics the incumbent when matched against the incumbent. Fudenberg and Maskin show that it is always possible to construct such a mutant for any such incumbent. This mutant is indistinguishable from the incumbent except after one particular history. If this history occurs, the mutant does as well as the incumbent when matched against the incumbent, but strictly better when matched against itself. Such a mutant strategy can invade, and hence no incumbent which does not attain efficiency after every history can be stable.

The following is an example of a strategy that is stable (provided that $\beta > 2\alpha$).



This strategy cooperates if and only if both players took the same action in the previous period. An accidental defection leads to a single period of mutual defection after which there is a return to mutual cooperation. Such a strategy is efficient after every history,

⁷An ordered pair of actions associated with an arrow now represents the vector of actions that cause that transition from one state to another, with the first component being the player's own action and the second the other player's action.

which requires it to be only mildly punitive after a defection. As in the case of (3), however, the strategy (4) does not easily fit an intuitive notion of reciprocal behavior. The criterion of evolutionary stability in the face of implementation errors therefore not only eliminates inefficient strategies, it also eliminates a number of intuitively appealing efficient ones.

3 Reciprocity in Sporadic Interactions

When interactions are sporadic the possibilities for punishment are limited and the conditions for the evolution of reciprocity are accordingly more stringent. Consider for instance the case of a large population the members of which are repeatedly matched pairwise at random to play the one-shot prisoners' dilemma with payoffs as given above. If s is the population share of cooperators, then the expected payoff to a defector is simply βs while that to a cooperator is $\beta s - \alpha$. From any initial state in which both types are present, the population evolves to the state in which only defectors are present. There is no scope for reciprocal behavior and the only stable outcome is the least efficient one.

There are at least two distinct ways in which reciprocity in sporadic interactions can be modeled. First, one could permit explicit punishment possibilities as part of the interaction, for instance by adding a stage to the game in which punishment in the form of costly sanctions may be imposed on others. Second, one could allow for the possibility that individuals interact selectively, rather than randomly, with others in the population, or condition their behavior on some potentially observable property that identifies a group to which their opponent belongs. We begin with the second case.

3.1 Assortation and Parochialism

Consider players from a large population matched to play the one-shot prisoners' dilemma. Players are programmed to either cooperate or defect. A cooperator meets another cooperator with probability r and a random member of the population with probability $1 - r$. The same interaction probabilities apply to defectors. After players have been matched, they receive payoffs and reproduce in proportion to those payoffs. The next generation of players is then matched again according to the interaction rule above. If the degree of assortative interaction is sufficiently high, the fraction of cooperators in the population will grow. To see this, observe that if the population share of cooperators is s , then the expected payoff to a cooperator is $\beta(r + (1 - r)s) - \alpha$ while the expected payoff to a defector is $\beta(1 - r)s$. The share of cooperators will increase over time if and only if $r > \alpha/\beta$, that is, if the rate of

assortative interaction is above the cost-benefit ratio.⁸

As the example above makes clear, all that is required for cooperative behavior to evolve (even in one-shot interactions) is that interaction be sufficiently assortative. Assortation may also arise endogenously, if individuals with a propensity to cooperate deliberately seek out other cooperators with whom to interact. This can be achieved if those with a propensity to cooperate are able to communicate this to each other by some pre-game signal.⁹ Assortation is conditioned on the predicted behavior of others and, if signals are sufficiently sharp, cooperators could interact only or mainly with other cooperators (Frank, 1987, 1988).

In Frank's model, individuals can choose whether or not to play prisoner's dilemmas with others from a large population consisting of two types. One type cooperates in all interactions, and the other always defects. The two types inherit the realization of a random variable or signal, X , from distributions $f(X|C)$ and $f(X|D)$ respectively. The support of $f(X|C)$ lies partly to the right of the support of $f(X|D)$, ensuring that at least a fraction of cooperators are recognizable as such with certainty. High values of X signal that a player is more likely to be a cooperator so such players are in demand as partners. The resulting sorting process leads to partial assortation: cooperators are more likely to be matched with other cooperators than with defectors. The information contained in the signals allows players to partly control for their opponent's type in choosing whether or not to interact with them. Frank shows that under certain conditions on the payoffs, there is an equilibrium proportion of cooperators which evolves. The intuition for this is that when most members of the population cooperate, then everyone is less discriminating in whom they choose to interact with because they can afford to be. This allows the share of defectors to grow until cooperators become sufficiently wary to prevent further growth in the defector share.

The assumption that at least a fraction of cooperators are recognizable with certainty, regardless of the population composition, is critical in allowing cooperators to invade a population of defectors. Without this assumption, the posterior probability that a player is a cooperator conditional on the value of X goes to zero as the population share of cooperators approaches zero.¹⁰ Without recognition, therefore, cooperators will fail to match with each

⁸This condition, known as Hamilton's rule, has a biological interpretation in terms of kin selection. Hamilton's rule states that an altruistic gene will spread in a population if individuals share at least a proportion r of their genes on average with those with whom they interact. In this context r is said to be the coefficient of relatedness (Hamilton, 1964).

⁹There is some experimental evidence to support the notion that individuals can predict when others will cooperate in prisoners' dilemmas after spending half an hour with them (Frank, Gilovich, and Regan, 1993).

¹⁰This would be less of an issue if the population were finite. In that case, even a single mutant has a non-zero population share, so that if the signal that players get about a player's type is sufficiently sharp,

other and cannot do better than defectors. Another potential objection to this account of the evolution of conditional cooperators is that a mutant who fakes the signal successfully could receive the benefits of cooperation without paying any of its costs. However, the ease of falsifying a signal of altruism should not be exaggerated. For such a mutation to be selected for, there must already be conditional cooperators present in the population. It is not at all clear that a mutation that perfectly copies the signal will arrive in any reasonable evolutionary time. What is more likely is the arrival of a mutant with an imperfect copy of the signal, which is precisely what is modeled by Frank.

While the discussion of signals in the formal models is entirely abstract, Hirshleifer (1987) and Frank (1988) have argued that behavioral propensities are at least in part determined by emotional states. Since emotions are mostly not under conscious control, and since their outward expressions are also only partly so, it is difficult to fake emotional states associated with particular behaviors. This line of argument suggests that while particular behavioral propensities and signals may evolve culturally, their credibility is at least partly a consequence of biological evolution. It also goes some way towards explaining why face-to-face interactions appear to generate more trust and cooperative behavior than those conducted electronically. Another implication of this observation is that reciprocity between large institutions such as firms or nation-states may have a basis that is quite different from that observed between individuals or small groups.

To continue with the discussion of models with signals, note that even if individuals cannot select their partners, they may be able to condition their *behavior* on a signal. We refer to this dependence of actions on opponents' types as parochialism. Robson (1990) considers the case of a mutant which emits a signal, and cooperates if and only if the signal is detected in the other player. If one adds the mutant strategy M to the payoff matrix (1), the following describes the payoffs to the three players:

	C	D	M
C	$\beta - \alpha$	$-\alpha$	$-\alpha$
D	β	0	0
M	β	0	$\beta - \alpha$

Even under purely random (nonassortative) matching, the such a mutant drives cooperators and defectors to extinction since the expected payoffs to M are strictly greater than the payoffs to the other types whenever the population share of M is strictly positive. Such behavior represents a parochial form of reciprocity, where cooperation is induced by means

 even though never perfectly revealing, it will suffice to enable cooperators to attain higher expected payoffs.

of a “secret handshake”. As Robson recognizes, however, a population consisting of such reciprocators is itself vulnerable to invasion by a second mutant which perfectly mimics the signal but always defects; such defecting mimics can drive the reciprocators to extinction. This suggests an endless evolutionary cycle involving the successful spread of parochial reciprocators with new signals, who are subsequently eliminated by defectors who develop the ability to perfectly mimic the signal. Alternatively, if the ability to mimic is imperfect, it is possible for mimics to coexist with reciprocators in the long run for reasons similar to those discussed in the context of Frank’s model above.

Yet another way of attaining conditional cooperation on the basis of signals is to use a prior cooperative act itself as a signal, thus rendering it difficult to falsify. Nowak and Sigmund (1998) take this route. In their model, players are randomly matched in pairs, one as a potential donor (helper) of a benefit β which costs α and one as a potential recipient. A player’s “image score” in each encounter is 1 if on his last encounter as a donor, he helped the recipient, 0 if he did not. Consider two possible strategies, ‘discriminate’, that is, help if (and only if) the recipient’s image score is 1, and ‘defect’, that is, never help.¹¹ If the probability q that a potential recipient’s image score is known to the potential donor is greater than α/β , then both strategies are asymptotically stable under payoff monotonic dynamics. However, with the introduction of a third strategy, ‘always help’, ‘discriminate’ is no longer asymptotically stable (though it remains weakly stable). The problem is that when there are no defectors in the population, the other two strategies get equal payoffs. With no selection pressure against the helpers, they can come to constitute the bulk of the population through random drift, after which defectors can invade. Moreover, the defector equilibrium remains asymptotically stable. This problem of reciprocity being undermined by the gradual encroachment of unconditional cooperation is pervasive in the literature.

3.2 Costly Sanctions

In sporadic interactions in which one cannot condition behavior on a pre-game signal, or interact selectively with some subset of the population, reciprocity can still evolve if there is the possibility of punishment within the interaction itself. Consider, for instance, the case of the prisoner’s dilemma augmented by a second ‘sanctioning’ stage. Specifically, suppose that after initial actions have been observed, each of the two players has the option of imposing a punishment on the other which costs γ to the punisher and inflicts damage δ on the victim.

¹¹Kahneman, Knetsch, and Thaler (1986) provide some experimental support for indirect reciprocal behavior of the kind modeled by Nowak and Sigmund.

The second stage payoffs can be represented by the following matrix, where P represents the action ‘punish’ and R the action ‘refrain from punishing’.

	P	R
P	$-\gamma - \delta$	$-\gamma$
R	$-\delta$	0

There are eight possible strategies in this two stage game: two first period choices (cooperation or defection) coupled with one of the following four second period actions: no punishment, punishment only of defectors, punishment only of cooperators, and punishment of both cooperators and defectors.

Regardless of the first stage outcome, and regardless of one’s expectation of the other player’s behavior, a self interested individual will always refrain from punishment since there is to be no further interaction between the players. As a result, the threat of punishment by a player who is known to be self-interested will not be credible, and when two such individuals are matched to play the two stage game, the predicted outcome is defection by both at the first stage and no punishment at the second. It is easily verified that a population in which only this strategy is present is neutrally stable. However, the strategy which prescribes cooperation and punishes defection is also neutrally stable.

The possibility that reciprocity in sporadic interactions can be stable even when the underlying threats of punishment are not ‘credible’ from the perspective of material self-interest is a special case of a very general phenomenon in extensive form games.¹² As in the case of Nowak and Sigmund’s model of indirect reciprocity, however, this stability is of a rather tenuous nature. Since unconditional cooperators do as well as reciprocators in the absence of defectors, the population share of the former can grow through random drift until defectors are able to invade. A neutrally stable population of unconditional defectors is not vulnerable in this sense. In a population of defectors there may be growth by random drift of the rather perverse strategy ‘defect and punish cooperation’, but any population consisting of these two strategies is itself uninvadable. The stability of defection is more robust than that of reciprocity.¹³ Notice, though, that this result is obtained *without* assuming that the

¹²See Somanathan (1997) for more on this point. Gale, Binmore and Samuelson (1995) and Sethi and Somanathan (1996) apply this insight to ultimatum bargaining and common pool resource games respectively; see also Axelrod (1986) for a simulation study of a prisoner’s dilemma model with costly sanctioning that has similar properties. Boyd and Richerson (1992) consider targeted retribution in large groups in the case of infinitely repeated interactions, again finding cooperation sustained by threats of punishment to be one of numerous stable outcomes.

¹³One way to strengthen the stability of cooperation enforced by threats of punishment is to include ongoing

commitment to punish defection is observable.

Observability of the commitment to punish defection gives rise to additional possibilities. Sethi (1996) considers the stability properties of various population states in the model discussed above, under the additional hypothesis that there exists a self-interested type who, instead of being programmed to play any particular strategy, plays an optimal response to its opponent's strategy at each interaction.¹⁴ When two such optimizers interact, they play according to the standard game theoretic prediction, which entails defection and no sanctioning. A population of reciprocators remains neutrally stable in the presence of such optimizers. Moreover, a population of optimizers is itself unstable and vulnerable to invasion by the reciprocator strategy provided that the damage δ from punishment exceeds the cost α of cooperating. When δ exceeds α , optimizers cooperate with reciprocators to avoid being punished. Reciprocators can eliminate optimizers in a population consisting only of these two types, since they obtain a payoff of $\beta - \alpha$ in each interaction, while opportunists obtain $\beta - \alpha$ only when matched with reciprocators and 0 when matched with each other.¹⁵ When one allows for all possible mutants, however, there can be other stable population states. A population of defectors who never punish is also neutrally stable, as is a population consisting of a combination of optimizers and 'bullies', where the latter defect at the first stage but punish defection by their opponent. When matched against bullies, optimizers cooperate to avoid punishment, and when matched against each other they defect since threats of punishment are not credible. In evolutionary equilibrium, the advantage that bullies have when facing opportunists is exactly outweighed by the substantial costs they incur when facing each other. Hence, while reciprocity is a stable outcome, it is just one of several such outcomes.

The introduction of self-interested optimizers into an otherwise standard evolutionary random drift explicitly in the model along the lines of Binmore and Samuelson (1999). If punishments are sufficiently severe and all types of mutations are equally likely, it is possible for there to be an asymptotically stable population state in which multiple types (including reciprocators) are present, and in which almost all individuals cooperate.

¹⁴This device of introducing rational players into an otherwise standard evolutionary game theoretic model had previously been employed by Banerjee and Weibull (1994), who explore conditions under which 'non-rational' types can survive.

¹⁵A population of optimizers would be *stable* if optimizers could not tell different opponents apart and were therefore forced to take the same action in each interaction, where this action is chosen optimally on the basis of the overall population composition. The ability to recognize reciprocators and adjust their behavior raises the payoffs of optimizers, but not as much as it raises the payoffs of reciprocators; the latter are therefore able to invade.

game theoretic framework produces a hybrid model in which some players are programmed to play a particular behavioral strategy while others (the optimizers) are flexible enough alter their behavior depending on their beliefs about their opponent. Gintis (2000) takes this feature further in a model of the evolution of reciprocal behavior in a public goods setting in which there are two types of players, reciprocators and defectors, both of whom are flexible. In this model, groups of a fixed size are formed at random every period. Defectors contribute to the public good only if they find themselves in a group in which the proportion of reciprocators is high enough that a failure to contribute results in a punishment that outweighs the cost of contributing. Reciprocators contribute to the public good only if they find themselves in a group in which they are in a high enough proportion that all players contribute. In such groups, they are known to be willing to punish players who do not contribute, although doing so is costly. Reciprocators neither contribute nor punish when they find themselves in a group in which their numbers would be insufficient to deter defection even if they were to punish it. Thus, although groups may contain any proportion of reciprocators, the parochial behavior of reciprocators, whose commitment to punish defection depends on the proportion of their own type in their group, has a homogenizing effect on behavior: there are only two kinds of groups; those in which all players contribute, and those in which none do. The evolutionary advantage of reciprocators arises from their being more likely than defectors to be in groups in which the public good is provided.

The model as described thus far would clearly result in the extinction of defectors, a result that is belied by data which show that a significant fraction of individuals behave like defectors (Fehr and Gächter, 2000). Gintis, however, makes the realistic assumption that monitoring costs incurred by reciprocators in groups in which they predominate lower their payoffs below those of defectors in these groups. If defectors are rare, most groups containing defectors will be of this kind. This enables defectors to invade a population of reciprocators. Gintis shows that, provided the monitoring cost is not too high, reciprocators will get higher average payoffs than defectors when the latter predominate in the population. Thus, the model predicts the evolution of a mixed population.

4 Reciprocity as a Preference

In the last model discussed, all players were flexible without necessarily being self-interested. The literature reviewed in this section takes this flexibility a step further by modeling reciprocity as an attribute of individual *preferences* rather than of a behavioral strategy. In this

framework, individuals are endowed with psychological payoff functions or utility functions, which may or may not be materially self-interested, and are assumed to take actions consistent with a Nash equilibrium of the game defined by their psychological payoff functions. This naturally allows individual behavior to be flexible in the sense that a given individual may take different actions against opponents with different preferences. As before, traits which are more successful with respect to material rewards are passed on to future generations with greater frequency. The traits in question are preferences, however, and not strategies. Whether or not preferences for reciprocity are stable in this setting is the main question of interest.

The basic idea underlying much of the literature on the evolution of non-opportunistic preferences is that a *recognized* adherence to such preferences can provide a commitment device that allows non-opportunists to outperform opportunists in certain strategic environments. The idea goes back at least as far as Schelling (1960) who observed that self-interested individuals would seek to pretend other motivations but that such pretences were liable to be imperfect. This theme has recurred in the work of Hirshleifer (1987) and Frank (1987, 1988), where emotional or otherwise “irrational” motivations can be profitable to those who possess them provided that they are at least partially observable and induce adjustments in the behavior of others. Building on this work, models of preference evolution have been developed in which all players are assumed to be maximizers of utility but to differ with respect to the relationship between material payoffs and subjective utilities (Güth and Yaari, 1992, Güth and Kliemt, 1992, Güth, 1995). Utilities determine equilibria which in turn determine material payoffs and hence evolutionary stability. This is the so-called *indirect evolutionary approach*.

4.1 Game-Specific Preferences

An essential component of models adopting the indirect evolutionary approach is the distinction between material and psychological payoff functions. Self-interested individuals or opportunists have psychological payoffs that are identical to their material payoffs, but not all players are self-interested. Given a specific game, preferences for reciprocity can be represented in terms of a particular transformation of the material payoff function to obtain a psychological payoff function. In Güth and Yaari’s (1992) analysis of the ultimatum bargaining game, for instance, a reciprocator is defined as an individual who gains subjective pleasure from rejecting unfair offers (but who does not gain subjective pleasure from *proposing* fair divisions). If this propensity is observable all proposers will make fair offers to reciprocators

and unfair ones to opportunists, resulting in greater material rewards to former. Even when the preferences of one's opponent are only partially observable, for instance through a signal that is correlated with preferences, reciprocators can survive and spread in a population of opportunists if the signal is sufficiently informative.¹⁶

Applying this approach to the one-shot prisoners' dilemma described in (1) above, suppose that reciprocators differ from opportunists in that they get some subjective disutility when they defect against someone who cooperates. If their willingness to pay to avoid this is $\rho > \alpha$, the payoffs to a reciprocator are given by

	C	D
C	$\beta - \alpha$	$-\alpha$
D	$\beta - \rho$	0

If $\rho > \alpha$ then mutual cooperation emerges as a second equilibrium outcome in this game when two reciprocators interact (the prisoners' dilemma is transformed into a coordination game). Consider a large population with two types, reciprocators and opportunists, in which players are randomly matched to play the one-shot prisoners' dilemma. Provided that reciprocators can be recognized as such, and interactions involving two reciprocators result in the selection of the efficient equilibrium, material payoffs to reciprocators will be strictly positive, unlike the material payoffs to opportunists. Reciprocators will drive opportunists to extinction (Guttman, 2000).

Note the crucial role played by the hypothesis that preferences are observable. With no information about types, a reciprocator population will be invaded by opportunists provided that reciprocators continue to attempt coordination on the efficient equilibrium when they believe with sufficiently high probability that their opponent is a reciprocator. This occurs because reciprocators will cooperate with all opponents (including opportunists, whom they cannot identify) when the population share of reciprocators is close to 1. Since opportunists defect in all interactions, their payoffs will exceed that obtained by reciprocators and their population share will rise. This process continues until the population share of reciprocators drops to a level at which reciprocators find the probability of being matched with each other

¹⁶Huck and Oechssler (1999) show that preferences for rejecting unfair offers can survive even when preferences are unobservable, provided that individuals interact in small groups and the preference distribution in the group is known to all. In the role of a proposer, reciprocators have an advantage over materialists since their opponent is less likely to be a reciprocator, so that unfair offers are less likely to be rejected. In the role of a responder, reciprocators have a disadvantage relative to materialists since they reject unfair offers, but this disadvantage is negligible since unfair offers are worth almost nothing to the responder.

too small to make cooperation worthwhile. At this point all individuals defect and one is left with a population of behaviorally indistinguishable reciprocators and opportunists.

If we introduce imperfect information about types in the form of a signal as described in the previous section, and some values of the signal perfectly identify reciprocators, then they will be able to invade (because those reciprocators with the appropriate signal values, when matched with each other, will achieve the efficient equilibrium, while the remaining reciprocators will be behaviorally indistinguishable from opportunists). This finding is driven by essentially the same logic as that of Frank (1987) described in the previous section. Allowing for a player's strategy to be conditioned on his opponent's type (parochialism) rather than simply allowing a player to choose whether or not to play (which results in assortment) yields essentially the same result.

These results rely on sufficiently informative signals. This is a strong assumption since, as noted earlier, evolution would favor mimics who fake the signal. What if, instead, the quality of the signal is arbitrarily poor?¹⁷ In this case reciprocators cannot invade a population of opportunists because when the latter's population share is close to 1, reciprocators will consider it extremely likely that they are matched with an opportunist even if they observe the reciprocator signal. This induces them to ignore the signal and defect against all opponents, so that the two types get equal payoffs. However, there will exist some threshold value of the reciprocator population share such that, when two reciprocators are matched and both (are commonly known to) observe a signal, each player's posterior probability that he is facing a reciprocator becomes high enough to make mutual cooperation an equilibrium. Suppose that reciprocators always cooperate in this situation, that is, when both players are commonly known to have received signals, and defect otherwise. In this case the only difference between reciprocator and opportunist payoffs arises due to the fact that reciprocators cooperate when both players are commonly known to have received signals, while opportunists defect. If signal quality is sufficiently low, opportunists will outperform reciprocators: they will be approximately as likely as reciprocators to find themselves in a situation where both players observe signals, and will get a strictly greater payoff than reciprocators do in this situation. What works against the reciprocators here is that the opportunists' payoff advantage from defection outweighs their loss from being slightly less likely to be thought to be reciprocators.

There is a simple way of making the payoff advantage from defection small: by allowing

¹⁷Suppose the signal X takes two values: zero and one. We mean that $P(X = 1|R)$ can be made arbitrarily close to $P(X = 1|O)$, where R denotes a reciprocator and O denotes an opportunist. Suppose $P(X = 1|R) > P(X = 1|O)$. We shall refer to a signal value of 1 as a 'reciprocator signal'.

players to play repeatedly in a single generation and letting the number of periods of play be sufficiently large. Guttman (1999) considers the possibility that there are a large but finite and commonly known number of periods of interaction in each generation. Each player in a generation inherits a particular realization of a signal X and players are randomly matched with others in each period. The entire history of a player’s play is known to all potential opponents, and, contingent on the observed history and signal, players may choose not to interact with their assigned opponent and be randomly matched with someone else instead. If the population share r of reciprocators is sufficiently low, reciprocators are behaviorally indistinguishable from opportunists and defect in each period regardless of the observed signal. Guttman shows that there exists an intermediate range of values of r for which only players with the “good” value of the signal cooperate with each other. Of these, the reciprocators cooperate for all periods while the opportunists cooperate for all periods except the last.¹⁸ Thus players with the “good” value of the signal get much higher payoffs than those with the “bad” value, and among these, opportunists do slightly better than reciprocators. By making the number of periods sufficiently large, this advantage of opportunists is shrunk until it is overwhelmed by the greater probability that reciprocators will have the “good” value of the signal, even though the probability of having the good value of the signal is only slightly greater for the reciprocators than for opportunists. Finally, for r sufficiently large, the probability of a player being a reciprocator is so high that reciprocators always cooperate, thus giving the advantage to opportunists. In consequence, there is an interior value of r , say r^* , with a non-trivial basin of attraction under the evolutionary dynamics. For lower values of r , payoffs are equal so it is possible for r to enter the basin of attraction of r^* through random drift after which it will be fixed near that value by selection. Hence, the model predicts the coexistence of opportunists and reciprocators in the long run.

This section has so far emphasized the importance of informative signals regarding player types for the evolution of reciprocity via parochialism or assortment. In contrast, recent papers by Friedman and Singh (1999) and Bowles and Gintis (1999) deal with the evolution of reciprocity when individual preferences are unobservable. These two papers differ in a number of details but share in common with the models discussed in Section 3.2 the feature that players may punish noncooperative behavior by others at some cost to themselves. Reciprocators are defined as individuals who have a taste for, and therefore are committed

¹⁸Materialists cooperate in all periods except the last because a failure to do so would expose them as materialists and lead to defection by their opponents in all subsequent periods. The idea that even a small probability of being mistaken for a type that is not self-interested can lead to dramatic effects on equilibrium behavior in the finitely repeated prisoners’ dilemma was developed in a seminal paper by Kreps et al. (1982).

to, punishing noncooperative behavior. Bowles and Gintis assume that society is segmented into distinct groups within which interaction occurs. If noncooperative behavior is detected and punished, individuals are ostracized from their group. Although reciprocators may also behave noncooperatively in equilibrium (and hence face punishment and expulsion), opportunists behave noncooperatively with greater frequency and are therefore more likely to be expelled. This results in assortative interaction: reciprocators are more likely than opportunists to be in a group with a large proportion of reciprocators. This compensates for the losses incurred by costly sanctioning of noncooperative behavior and both types can coexist in the long run. As in many other models, however, this conclusion depends critically on the assumption that there does not exist a type of player who cooperates but refuses to sanction noncooperative behavior. Such pure altruists would escape the costs of inflicting punishment while continuing to enjoy the advantages of assortative interaction thus destabilizing any population state in which reciprocators are present.

Friedman and Singh consider a population divided into a number of distinct groups and consider pairwise interactions involving members of *different* groups. As before, reciprocators are defined as individuals who have a taste for punishing noncooperative behavior. Although such individual tastes are unobservable, each group has a “vengeance norm” which prescribes some level of punishment. The group to which an individual belongs is itself observable, and this provides information regarding that individual’s propensity for punishment. Members of groups which prescribe a vengeance norm benefit from this reputation since the threat of punishment induces their opponents to behave cooperatively. Moreover, individuals face incentives to adhere to their group norms since they are punished for deviations by their fellow group members. Specifically, the punishment is imposed by a group on one of its members who fails to be sufficiently vengeful in an interaction with an outsider, and consists of a loss of status for that individual, which has material consequences. Friedman and Singh assume that this second-order punishment is costless to impose, and on this basis obtain asymptotic stability of reciprocal preferences. This result depends critically on the assumption of costless (second-order) punishment because in the presence of such punishment costs, individuals who do not punish deviators would enjoy an evolutionary advantage over those who do, which in turn would lead to an erosion of the vengeance norm and ultimately to the decay of cooperation itself.¹⁹

¹⁹One way to recover the result would be to assume, as in Akerlof’s (1976) model of ‘caste equilibrium’, that those who fail to ostracize are themselves ostracized and so on, *ad infinitum*. Then it becomes necessary to assume not only that the original offence is observable, but that all higher-order offences are also observable.

4.2 Universal Preferences

Within the context of particular games, one may represent preferences for reciprocity by some transformation of the material payoff functions that makes intuitive sense. It is possible, however, to model preferences for reciprocity in an entirely different way, one which is independent of any particular game or strategic environment. Recent research, motivated by the search for a unified explanation of the results from different experimental games, has proposed specifications of preferences which are *context-free*, that is, which are applicable to any game. While the game-specific preferences considered in the previous section can be interpreted as reduced forms of some underlying (though unspecified) universal preferences, the literature reviewed in this section is based on an explicit specification of such preferences.

Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) have recently provided specifications of preference interdependence in which individuals are assumed to care not only about their own material payoffs but about the entire distribution of payoffs. Although they differ with respect to a number of details, both papers require that individuals experience some disutility from being at either extreme of the payoff distribution, that is, they have some aversion to inequality. These papers are able to explain much more of the data from laboratory experiments than can simpler specifications of preference interdependence. A second group of papers explicitly adopts the approach of *psychological games* in which player utilities depend not just on action profiles, but also on their initial beliefs (Geanakoplos et al., 1989, Rabin, 1993, Dufwenberg and Kirchsteiger, 1998, Falk and Fischbacher, 1998). In equilibrium, all beliefs (including higher-order beliefs) are correct, and individuals take optimal actions conditional on these beliefs and the actions of others. Different beliefs (corresponding to different equilibria) imply possibly different utility profiles at any given action profile. This endogeneity of utility profiles represents a considerable departure from standard game theoretic methodology. Papers using the apparatus of psychological games to explain data from experiments are based on the hypothesis that beliefs about the kindness or unkindness of opponent strategies will give rise to the desire to reciprocate, where the kindness or unkindness of an individual's strategy is assessed in terms of the (material) payoff implications of other strategies available to them.²⁰

²⁰Charness and Rabin (2000) have recently proposed an alternative model which incorporates a concern for efficiency, distributional justice, and reciprocity. Individuals place some weight on a "disinterested social ideal" which itself reflects a concern for efficiency (maximizing the sum of all payoffs) and a Rawlsian notion of justice (maximizing the payoff of the least well-endowed) but are prepared to abandon this ideal and behave selfishly or maliciously if they believe that others are behaving in too selfish a manner.

A third approach is based on the hypothesis that an individual's utility depends not only on the distribution of material payoffs but also on that individual's beliefs about the *preferences* of others (Levine, 1998). Consider a group of n individuals and any material payoff distribution (π_1, \dots, π_n) . The interpretation of these payoffs is that they arise from the choice of actions by each individual in some strategic interaction. The psychological payoff or utility achieved by each individual is assumed by Levine to depend on the material payoff distribution as follows

$$u_i = \pi_i + \sum_{j \neq i} w_{ij} \pi_j, \quad (5)$$

where w_{ij} is the weight that i places on j 's material payoff. This weight itself depends on properties of the preferences of the two players. Each player's preferences are represented by two parameters a_i and λ_i , where a_i may be interpreted as a measure of an individual's pure altruism, and λ_i a measure of the degree to which the weight w_{ij} placed by individual i on the material payoffs of individual j is sensitive to the altruism of the latter. Specifically, it is assumed that

$$w_{ij} = \frac{a_i + \lambda_i a_j}{1 + \lambda_i}$$

Levine argues that a suitably chosen, stable distribution of preferences belonging to this class can simultaneously account for results from a wide variety of experimental games. This specification does not, however, pass a natural test for evolutionary stability. Note that an individual i with $a_i > 0$ can never place a negative weight on the payoffs of an individual j who is purely self-interested ($a_i > 0$ and $a_j = 0$ implies $w_{ij} > 0$). Such behavior, being altruistic, can survive and spread in competition with opportunist preferences only under restrictive conditions. This problem does not arise in the following modification of Levine's preferences:

$$w_{ij} = \frac{a_i + \lambda_i (a_j - a_i)}{1 + \lambda_i}. \quad (6)$$

If both a_i and λ_i are positive, and $\lambda_i > 1$, then the possessor of such a utility function will be altruistic toward those with a large enough value of a and *spiteful* toward opportunists.

Consider a large population consisting of two types of preferences, opportunists and reciprocators, both belonging to the class (6). Opportunists are defined by the parameters $a = \lambda = 0$ and reciprocators by $a > 0$ and $\lambda > 1$. Opportunists place no weight on the payoffs of others. Reciprocators place weights k_1 and k_2 on the payoffs of reciprocators and

opportunists respectively, where

$$k_1 = \frac{a}{1 + \lambda} > 0; \quad k_2 = -\frac{a(\lambda - 1)}{1 + \lambda} < 0.$$

Applying this to the example of the prisoners' dilemma, when two reciprocators interact, the payoff matrix is given by

	<i>C</i>	<i>D</i>
<i>C</i>	$(\beta - \alpha)(1 + k_1)$	$-\alpha + \beta k_1$
<i>D</i>	$\beta - \alpha k_1$	0

Provided that $k_1 > \alpha/\beta$, cooperation is a dominant strategy for each player. It is easily verified that when a reciprocator interacts with an opportunist, defection is a dominant strategy for both players (opportunists defect out of self-interest, reciprocators out of spite). If preferences are perfectly observable the expected payoff to reciprocators exceeds that to opportunists regardless of the population composition and the latter are therefore driven to extinction. Under unobservable preferences, an argument similar to that of Guttman (1999) can be made to predict a stable population composition in which both types of preference are present.

While the prisoners' dilemma example illustrates some of the evolutionary advantages that such preferences enjoy, it obscures the fact that the potential for spite implicit in these preferences can confer other advantages. Building on earlier work by Koçkesen et al. (2000), Sethi and Somanathan (2001) identify a class of games, including multiperson common pool resource and public goods games, in which such preferences can survive. In this model, individuals are matched in subgroups of size k drawn randomly from a large population. The reciprocator advantage stems from two factors: their spiteful behavior when they find themselves in a subgroup consisting mainly of opportunists, and the opportunists' recognition of their spite. This recognition causes opportunists to adjust their actions in such a manner as to increase the material payoffs of reciprocators. Provided that reciprocators are not too spiteful towards opportunists, this results in reciprocators having higher expected payoffs than opportunists when the population share of the latter is sufficiently high. As evolutionary pressure raises the population share of reciprocators, their actions become increasingly altruistic (since they are more and more likely to find themselves in groups with many reciprocators). This tends to favor opportunists and may therefore limit the extent to which the population share of reciprocators can rise. For some parameter values the only stable population states consist of both types. Note that this argument depends crucially on

players having sufficient information about the preferences of others in their group so that behavior towards members of the subgroup can be conditioned on its composition.²¹

If the sorting into subgroups is assortative rather than random, then reciprocators may have an advantage for a different reason. Their altruism towards other reciprocators results in groups composed exclusively of reciprocators getting higher payoffs than groups consisting exclusively of opportunists. It is the flexibility of reciprocal preferences that gives them survival advantages in both these very different matching environments. Neither purely altruistic nor purely spiteful preferences have this flexibility. Hence preferences for reciprocity are able to survive and spread in a broader range of environments than purely altruistic or spiteful preferences.

5 Discussion

Given the pervasiveness and economic importance of reciprocity in human interactions, it is not surprising that there have been a variety of approaches to understanding how such behavior can survive and spread in evolutionary competition with opportunistic behavior. Each of the contributions reviewed above shed some light on the phenomenon. Four basic themes, repetition, commitment, assortment, and parochialism, have appeared repeatedly in various guises throughout the literature. Infinite repetition can give rise to the evolution of behavior that exhibits reciprocity-like features but the stable strategies that emerge in this environment are less intuitively recognizable as representations of actual human behavior than strategies such as tit-for-tat, which are unstable. This gulf between the predictions of theoretical models and the behavior of human subjects in repeated interactions remains to be bridged. In sporadic interactions, reciprocity can be stable if the propensity to punish selfish actions can induce opportunists to cooperate, or if reciprocators themselves behave opportunistically when they expect others to do so. Reciprocity can also be stable in sporadic interactions if matching is sufficiently assortative. In each case, the propensities or preferences of others must be at least partially observable, the only exception being that commitment can induce a weak form of stability even when preferences are unobservable.

An implicit assumption made throughout the literature is that traits that are more highly rewarded in material terms will increase in populations at the expense of those that are less highly rewarded. This hypothesis is uncontroversial when material payoffs are interpreted

²¹If players in a subgroup receive no information about the preferences of others in their subgroup, then materialists cannot receive lower expected material payoffs than players with other utility functions (Ok and Vega-Redondo, 2001).

as Darwinian fitness and the transmission of traits is genetic. But as far as the transmission of cultural traits is concerned, the hypothesis requires greater scrutiny and alternative dynamic processes based on learning, imitation and socialization need to be examined (see, for instance, Bisin and Verdier, 2001, for an explicit model of parental socialization). Cultural traits that are most easily transmitted to children by parents or peers need not be the ones that yield the highest material rewards. Complex and varied motives govern the actions of parents who attempt to instill values, habits or propensities in their children, and the desire to see their children become materially prosperous need not be a dominant parental concern. Furthermore, when learning about payoffs is costly, there are advantages to imitating frequent behaviors so the transmission of traits may be partly conformist (Boyd and Richerson, 1985). Conformist transmission can overwhelm small payoff disadvantages and, provided that the costs of sanctioning are low, virtually any behavior enforced by punishments can be asymptotically stable.

A related question which has received insufficient attention is the time scale over which meaningful changes in the population composition may be expected to occur. Even in the case of cultural transmission, traits transmitted horizontally within peer groups can spread rapidly within a generation, while traits that are transmitted vertically from one generation to the next, for instance within families, may require several generations to diffuse through the population. Closer attention to the precise manner in which preferences are acquired in human populations is clearly warranted.

Alternative assumptions regarding population structure also warrant exploration. The random matching hypothesis is analytically convenient but relatively unrealistic and the structure of human interaction more closely resembles a network with a high degree of clustering. Interactions are predominantly with a relatively small set of social ‘neighbors’ and one’s neighbors are quite likely also interact with each other. Recent studies of the evolution of behavior on simple symmetric networks suggest that cooperation and reciprocity can be sustained in such environments (Nowak and May 1992, Eshel et al., 1998, Boyd and Richerson, 2000, Albin and Foley, 2001). The evolution of behavior and preferences in more complex environments such as ‘small-world’ networks (Watts, 1999) remains to be systematically explored.

The literature reviewed here includes models in which individuals are boundedly rational in an extreme sense, mechanically choosing actions, receiving payoffs, and myopically adjusting their behavior. It also includes models based on the equally extreme assumption that individuals exhibit an unlimited capacity to anticipate the actions of others and coor-

dinate instantaneously on an equilibrium action profile. The vast ground between these two extremes clearly merits serious exploration, through the explicit modeling of cognitive constraints (see Bednar and Page, 2001, and Samuelson, 2001, for recent steps in this direction). It has recently been argued that even allowing for a richer set of preferences, the behavior of experimental subjects commonly violates the hypothesis of sequential rationality (Binmore et al., 1999). This suggests that modeling bounded rationality and/or directly letting evolution act on behavior may be more fruitful than modeling non-opportunistic utility functions at least in some contexts.

Finally, future research within the preference-based approach would benefit from using universal (context-free) specifications of preferences rather than the game-specific payoff transformations that are currently prevalent. The main source of such specifications is clearly the experimental literature, but evolutionary reasoning can help refine and select among various hypotheses. Experimental economists are seeking the most parsimonious specification that will fit a wide variety of data, and evolutionary models can be used to guide and discipline this search. This symbiotic evolution of experimental and theoretical work is one of the most promising and exciting areas of future research on the nature and origins of reciprocity.

References

- [1] Akerlof, George, 1976. The economics of caste and of the rat race and other woeful tales. *Quarterly Journal of Economics*, 90, 599-617.
- [2] Albin, Peter S. and Duncan K. Foley, 2001. The co-evolution of cooperation and complexity in a multi-player local-interaction prisoners' dilemma. *Complexity*, forthcoming.
- [3] Axelrod, Robert and William D. Hamilton, 1981. The evolution of cooperation. *Science*, 211, 1390-1396.
- [4] Axelrod, Robert, 1984. *The evolution of cooperation* (Basic Books, New York).
- [5] Axelrod, Robert, 1986. An evolutionary approach to norms. *American Political Science Review*, 80, 1095-111.
- [6] Banerjee, Abhijit V. and Jörgen Weibull, 1994. Evolutionary selection and rational behavior, In: Alan Kirman and Mark Salmon, (Eds.), *Rationality and Learning in Economics*. Basil Blackwell, Oxford.
- [7] Bednar, Jenna and Scott Page, 2001. Explaining behavioral diversity through strategic choice in game ensembles: A formal analysis of culture. Mimeo, University of Michigan.
- [8] Berg, Joyce, John W. Dickhaut and Kevin A. McCabe, 1995. Trust, reciprocity, and social history. *Games and Economic Behavior*, 10, 122-42.
- [9] Binmore, Kenneth G. and Larry Samuelson, 1992. Evolutionary stability in repeated games played by finite automata. *Journal of Economic Theory*, 57, 278-305.
- [10] Binmore, Kenneth G. and Larry Samuelson, 1999. Evolutionary drift and equilibrium selection. *Review of Economic Studies*, 66, 363-93.
- [11] Binmore, Kenneth G., John McCarthy, Giovanni Ponti, Larry Samuelson, and Avner Shaked, 1999. A backward induction experiment. Mimeo, University College London.
- [12] Bisin, Alberto and Thierry Verdier, 2001. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97, 298-319.
- [13] Bolton, Gary E. and Axel Ockenfels, 2000. ERC: A theory of equity, reciprocity and competition. *American Economic Review*, 90, 166-193.

- [14] Bowles, Samuel and Herbert Gintis, 1998. The evolution of strong reciprocity. Santa Fe Institute Working Paper 98-08-073E.
- [15] Boyd, Robert and Peter J. Richerson, 1985. Culture and the evolutionary process (University of Chicago Press, Chicago).
- [16] Boyd, Robert and Peter J. Richerson, 1992. Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13, 171–195.
- [17] Boyd, Robert and Peter J. Richerson, 2000. Group beneficial norms can spread rapidly in a structured population. Mimeo, UCLA.
- [18] Charness, Gary and Matthew Rabin, 2000. Social preferences: Some simple tests and a new model. Mimeo, University of California at Berkeley.
- [19] Dufwenberg, Martin and Georg Kirchsteiger, 1998. A theory of sequential reciprocity. CentER Discussion paper 9837, Tilburg University.
- [20] Eshel, Ilan, Larry Samuelson and Avner Shaked, 1998. Altruists, egoists, and hooligans in a local interaction model. *American Economic Review*, 88, 157-179.
- [21] Falk, Armin and Urs Fischbacher, 1998. A theory of reciprocity. Mimeo, University of Zurich.
- [22] Fehr, Ernst and Simon Gächter, 2000. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14, 159-181.
- [23] Fehr, Ernst, Georg Kirchsteiger, and Arno Reidl, 1993. Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics*, 108, 437-460.
- [24] Fehr, Ernst, and Klaus M. Schmidt, 1999. A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 817-68.
- [25] Frank, Robert H., 1987. If Homo Economicus could choose his own utility function, would he want one with a conscience? *American Economic Review*, 77, 593-604.
- [26] Frank, Robert H., 1988. Passions within reason: The strategic role of the emotions (W.W. Norton, New York).
- [27] Frank, Robert H., Thomas Gilovich and Dennis T. Regan, 1993, The evolution of one-shot cooperation: An experiment. *Ethology and Sociobiology*, 14, 247-256.

- [28] Friedman, Daniel and Nirvikar Singh, 1999. On the viability of vengeance. Mimeo, UC Santa Cruz.
- [29] Fudenberg, Drew and Eric S. Maskin, 1986. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54, 533-554.
- [30] Fudenberg, Drew. and Eric S. Maskin, 1990. Evolution and cooperation in noisy repeated games. *American Economic Review Papers and Proceedings*, 80, 274-279.
- [31] Gale, John, Kenneth G. Binmore and Larry Samuelson, 1995. Learning to be imperfect: The ultimatum game. *Games and Economic Behavior*, 8, 56-90.
- [32] Geanakoplos, John, David Pearce and Ennio Stacchetti, 1989. Psychological games and sequential rationality. *Games and Economic Behavior*, 1, 60-79.
- [33] Gintis, Herbert, 2000. Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206, 169-179.
- [34] Güth, Werner, 1995. An evolutionary approach to explaining cooperative behavior by reciprocal incentives. *International Journal of Game Theory*, 24, 323-44.
- [35] Güth, Werner and Harmut Kliemt, 1994. Competition or cooperation: On the evolutionary economics of trust, exploitation, and moral attitudes, *Metroeconomica*, 45, 155-187.
- [36] Güth, Werner, Rolf Schmittberger and Bernd Schwarze, 1982. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3, 367-388.
- [37] Güth, Werner, and Menahem Yaari, 1992. Explaining reciprocal behavior in simple strategic games: An evolutionary approach, In: U. Witt, (Ed.), *Explaining process and change: Approaches to evolutionary economics*. University of Michigan Press, Ann Arbor, pp. 23-34.
- [38] Guttman, Joel M., 1999. Self-Enforcing agreements and the evolution of preferences for reciprocity. Mimeo, Bar-Ilan University.
- [39] Guttman, Joel M., 2000. On the evolutionary stability of preferences for reciprocity. *European Journal of Political Economy*, 16, 31-50.
- [40] Hamilton, William D., 1964. The genetical evolution of social behavior. *Journal of Theoretical Biology*, 7, 1-16.

- [41] Hirshleifer, Jack, 1987. On the emotions as guarantors of threats and promises, In: John Dupré, (ed.), *The latest on the best: Essays in evolution and optimality*, MIT Press, Cambridge.
- [42] Hirshleifer, Jack and Juan Carlos Martinez Coll, 1988. What strategies can support the evolutionary emergence of cooperation? *Journal of Conflict Resolution*, 32, 367-398.
- [43] Huck, Steffen and Jörg Oechssler, 1999. The indirect evolutionary approach to explaining fair allocations. *Games and Economic Behavior*, 28, 13–24.
- [44] Koçkesen, Levent, Efe A. Ok and Rajiv Sethi, 2000. The strategic advantage of negatively interdependent preferences. *Journal of Economic Theory*, 92, 274-299.
- [45] Kreps, David, Paul Milgrom, John Roberts and Robert Wilson, 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27, 245-52.
- [46] Levine, David K., 1998. Modeling altruism and spitefulness in experiments. *Review of Economic Dynamics*, 1, 593-622.
- [47] Maynard Smith, John, 1982. *Evolution and the theory of games* (Cambridge University Press, Cambridge).
- [48] Maynard Smith, John and George R. Price, 1973. The logic of animal conflicts. *Nature*, 246, 15-18.
- [49] Nowak, Martin A. and Robert M. May, 1992. Evolutionary games and spacial chaos. *Nature*, 359, 826-829.
- [50] Nowak, Martin A. and Karl Sigmund, 1998. Evolution of indirect reciprocity by image scoring. *Nature*, 393, 573-577.
- [51] Ochs, Jack and Alvin E. Roth, 1989. An experimental study of sequential bargaining. *American Economic Review*, 79, 355-384.
- [52] Ok, Efe A. and Fernando Vega-Redondo, 2001. On the evolution of individualistic preferences: An incomplete information scenario. *Journal of Economic Theory*, 97, 231-254.
- [53] Rabin, Matthew, 1993. Incorporating fairness into game theory and economics. *American Economic Review*, 83, 1281-302.

- [54] Robson, Arthur, 1990. Efficiency in evolutionary games: Darwin, Nash, and the secret handshake. *Journal of Theoretical Biology*, 144, 379-396.
- [55] Rosenthal, Robert W., 1997. Trust and social efficiencies. Mimeo, Boston University.
- [56] Samuelson, Larry, 2001. Analogies, adaptation, and anomalies. *Journal of Economic Theory*, 97, 320–366.
- [57] Schelling, Thomas C., 1960. *The strategy of conflict* (Harvard University Press, Cambridge).
- [58] Sethi, Rajiv, 1996. Evolutionary stability and social norms. *Journal of Economic Behavior and Organization*, 29, 113-140.
- [59] Sethi, Rajiv, and E. Somanathan, 1996. The evolution of social norms in common property resource use. *American Economic Review*, 86, 766-88.
- [60] Sethi, Rajiv and E. Somanathan, 2001. Preference evolution and reciprocity. *Journal of Economic Theory*, 97, 273–297.
- [61] Somanathan, E., 1997. Evolutionary stability of pure-strategy equilibria in finite games. *Games and Economic Behavior*, 21, 253-65.
- [62] Trivers, Robert L., 1971. The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46, 35-57.
- [63] Watts, Duncan, 1999. *Small worlds: The dynamics of networks between order and randomness* (Princeton University Press, Princeton).