Homo reciprocans: A Research Initiative on the Origins, Dimensions, and Policy Implications of Reciprocal Fairness

Samuel Bowles
Robert Boyd
Ernst Fehr
Herbert Gintis

June 7, 1997

Abstract

Experimental economists and other social scientists have discovered an important form of human behavior that has been inadequately analyzed by behavioral scientists. In public goods, ultimatum, and other games where players gain from cooperative behavior, agents have a predisposition to cooperate and to undertake costly punishment of defectors, even when this behavior cannot be justified in terms of traditional game-theoretic equilibrium and learning concepts. We call this 'reciprocal fairness.'

Our research has four goals. First, can the experiments on reciprocal fairness be replicated with diverse subject pools and various strategic settings? Second, how might such behavior have evolved, given that it is formally altruistic, and hence 'unfit' except under stringent circumstances? Third, how does the existence of reciprocal fairness influence our analysis of social policy in such areas as taxation, charity, redistributive expenditure, and criminal sentencing? Fourth, how to what extent does cultural variation induce differences in the strength of reciprocal behavior and conditions under which agents exhibit reciprocal fairness?

1 Explaining Cooperative Behavior

Despite its reputation for fostering competitive behavior, the contemporary market economy sustains important forms of cooperation. Employees cooperate with one another and with management in the enterprise, agents are more or less trustworthy in exchange situations, family members provide for one another, people give to charity, volunteer for public service, and support government redistributive expenditures.

As Bernard de Mandeville ("The Fable of the Bees") and Adam Smith ("The Invisible Hand") long ago observed, the proper institutional framework can induce self-interested agents to serve the interest of others. The most sophisticated modern statement of this principle is probably the Fundamental Theorem of Welfare Economics, based on the Walras-Arrow-Debreu general equilibrium model.[1]

This 'neoclassical' explanation of cooperation presumes that all economic transactions are fully contractible, and all contracts are costlessly enforceable. However cooperation in modern market economies is not limited to situations of complete and costlessly enforced contracts. When the assumptions of the general equilibrium model are appropriately weakened a different set of analytical tools are needed to explain why self-interested agents cooperate.

One such tool is the repeated game, which implies the famous 'Folk Theorem,' implying that when discount rates are sufficiently low, Pareto-optimal cooperation can be sustained if each agent responds to a non-cooperative action by refusing to cooperate for the remainder of the game. Under suitable conditions such threats can be shown to be credible, in the sense that it will indeed be in the self-interest of each agent to carry out the threat of retaliatory defection should the occasion to do so arise, given that every other individual does so as well.[2] While few instances of social cooperation actually use universal defection to sustain cooperation, other plausible sanctions can serve the job as well.[3]

The principal-agent model and its many variations have also been used widely to explain cooperation in the firm, and have been extended to provide an analytical basis for a theory of economic institutions based on treating individuals as rational and self-interested.[4]

However important forms of cooperative behavior are commonly ob-

[1] See, for instance, Debreu (1959), Arrow and Debreu (1954), Arrow and Hahn (1971), or Mas-Colell (1985).

[2] See Thompson and Faith (1981) and Fudenberg and Maskin (1986). For a more general treatment of the Folk Theorem and its variants, see Kreps (1990) or Fudenberg and Tirole (1991).

[3] For instance, a group of potential cooperators can offer incentives for individuals to monitor and punish defection (Weissing and Ostrom 1991, Bendor and Mookherjee 1987), or can ostracize defectors (Cremer 1986, Hirshleifer and Rasmusen 1989, Gintis 1989, Boyd and Richerson 1992). In addition, cooperators can differentially associate with other cooperators (Boorman and Levitt 1980, Grafen 1979, Hamilton 1963, Wilson 1980, Trivers 1971, Bergstrom and Stark 1993, Bergstrom 1995).

[4] See Shapiro and Stiglitz (1984), Bowles (1985), Hölmstrom (1979), Hölmstrom (1982), Bowles and Gintis (1993), Hölmstrom and Milgrom (1994), Grossman and Hart (1983, 1986) and Hart and Moore (1990).

served, and consistently reproduced in the laboratory, that are difficult to explain by modeling economic agents as self-interested actors in the tradition of the rational actor model. One is that people are 'irrationally' pro-social: in laboratory settings subjects consistently appear to contribute more than the rational actor model predicts in public goods games (Marwell and Ames 1979, Schneider and Pommerehne 1981, Dawes, de Kragt and Orbell 1988, Isaac and Walker 1988b, Isaac and Walker 1988a). For a review of the extensive literature on this subject, see Ledyard (1995). We may summarize this research as follows. Only a fraction of subjects consistently defect, contributing nothing to the public account. Rather, people begin by making contributions that average about midway between the perfectly cooperative and the perfectly noncooperative levels, although cooperation deteriorates if the game is repeated numerous times.

There have been some attempts to explain pro-social public good contribution on the basis of the rational actor model, most notably Joel Guttman (1986, 1987). The models we have seen, however, require implausible commitment assumptions, and do not explain the laboratory evidence on the public goods game. Others have attempted to reconcile this behavior with 'rationality' by noting that in repeated public goods games, cooperation decays over time, eventually approximating the unique subgame perfect equilibrium. This suggests that participants do not understand the game at first, but progressively learn the superiority of the free-riding strategy.

There is persuasive evidence that interpretation is incorrect, however. For instance, Andreoni (1988) and Fehr and Tyran (1996) find that when such games are repeated with the same subjects, the initial levels of cooperation are restored, but once again cooperation decays as the game progresses. These authors suggest that the decay of cooperation in repeated public goods games occurs because public-spirited contributors retaliate against free-riders in the only way available to them in the game: by defecting themselves.

Indeed, retaliation is a second form of anomalous behavior consistently found in laboratory studies: that people appear to be 'irrationally' vindictive. For instance, responders in ultimatum games reject positive offers apparently to impose costs on 'unfair' proposers (Forsythe, Horowitz, Savin and Sefton 1994, Hoffman, McCabe and Smith April, 1996, Cameron 1995, Abbink, Bolton, Sadrieh and Tang 1996, Eckel and Grossman 1996a, Eckel and Grossman 1996b, Eckel and Grossman 1997).[5]

Some have explained this anomaly simply as error on the part of re-

[5]For an overview, see Davis and Holt (1993) and Fehr and Tyran (1996).

sponders. The most sophisticated explanation of this type is perhaps Gale, Binmore and Samuelson (1995), who make the dynamic argument that the error rates of responders are likely to be high and negatively correlated with the size of the offer in a repeated ultimatum game. We find this explanation implausible, however, since human subjects are normally adept at distinguishing receiving a positive amount from receiving nothing (Cosmides and Tooby 1992).[6] Also Bolton and Zwick (1995) and Abbink et al. (1996) show that if the rules of the game are changed so that the proposer keeps the share he suggests whether or not the responder accepts the offer, the outcome quickly approximates the subgame perfect equilibrium predicted by the rational actor model: the proposer keeps as much as possible, and the respondent accepts whatever he is offered.

2 Homo reciprocans

A predisposition to cooperate and to undertake costly punishment are probably related phenomena. Our proposed research investigates the implications of combining the two forms of anomalous behavior sketched above. We refer to the combination as *reciprocal fairness*. Ostrom et al. (Gardner, Ostrom and Walker 1990, Ostrom, Walker and Gardner 1992) and Fehr et al. (Fehr and Gächter 1996, Fehr, Gächter and Kirchsteiger 1997, Fehr and Tyran 1996) have provided evidence for the existence of reciprocal fairness in a variety of social situations: a majority of individuals approach strategic interactions involving coordination problems with a propensity to cooperate, they respond to the cooperation of others by maintaining or increasing their level of cooperation, and they respond to defection on the part of others by retaliating against the offenders, even at a cost to themselves, and even when they cannot reasonably expect future personal gains from such retaliation.[7] In particular, when other forms of punishment are not available, individuals

[6] One of us (Herbert Gintis) has used artificial life simulations of the ultimatum game indicating that sufficiently high mutation rates and sufficiently low rates of migration among groups (or a sufficiently small size of the whole population) can reproduce the empirical finding of the ultimatum game. This result, however, is more plausibly interpreted as genetically-evolved vindictiveness rather than 'noise' or 'error' on the part of responders, since normally a considerable fraction of the population exhibits retaliatory behavior.

[7] In the work of Ostrom et al. (1992) the same group of subjects interacted for roughly 25 periods, and subjects could develop an individual reputation for punishing defectors. Their experimental design therefore permits an interpretation of costly retaliation in terms of strategically rational behavior: retaliation may increase cooperation in future periods. In Fehr and Gächter (1996), group composition is changed in every period and individual reputation formation is ruled out by the design. Therefore, costly retaliation does not

respond to defection with defection. We dub such agents *Homo reciprocans*, to highlight the contrast of this behavior with that of the more traditional *Homo economicus*. *Homo reciprocans* is thus neither the selfless altruist of utopian theory, nor the selfish hedonist of neoclassical economics. Rather, he is a conditional cooperator whose penchant for reciprocity can be elicited under the proper circumstances.[8]

The novel element in this research is not the recognition that reciprocal behavior exists. The importance of altruism within families has been stressed by William Hamilton in his seminal work on 'inclusive fitness' (1963, 1964), and Robert Trivers (1971) has shown that no concept of rationality is need to predict that even unrelated individuals can gain from reciprocal behavior in repeated interactions. The robustness of reciprocal behavior appears in computer simulations as well, as in the work of Hamilton and Robert Axelrod (1981, 1984) and others, reviewed in Axelrod and Dion (1988). Artificial life simulations of repeated prisoner's dilemma games using genetic algorithms also show the robustness of strategies that are 'nice' (never defect first), 'punishing' (always punish defection) and 'forgiving' (return to cooperation after a short period of punishing, if the other player is cooperating) (Bowles and Gintis 1998a).[9]

Rather, the novel element is the fact that retribution is 'altruistic' in the sense that retaliatory behavior benefits the group by fostering sustained cooperation, but at a strictly positive cost to the individuals who bear the trait.

3 Topic I: Experimental Research on Sustaining Cooperation Via Reciprocal Fairness

The Ostrom et al. and the Fehr and Gächter results show that free-riding is significantly alleviated if there is an opportunity for costly retaliation. The results, however, do not show that permitting costly retaliation leads to welfare gains for group members. In the presence of uncoordinated costly retaliation opportunities the total costs of retaliation tend to outweigh the gains that are achieved through higher cooperation rates. The Ostrom et al. results show that if costly retaliation opportunities are combined with communication opportunities almost no defection occurs and, therefore, no

confer any pecuniary benefit to those who punish. Nonetheless, punishment of free-riding was prevalent and gave rise to a large and sustainable increase in cooperation levels.

[8] Levine (1996) analyzes a utility function that combines cooperation and retribution.

[9] Non-reciprocating strategies also emerge from artificial life simulations. On the fitness of non-reciprocating strategies, see Nowak and Sigmund (1993).

resources are wasted for retaliation. This result raises, however, the question to what extent the uncontrolled effects of face to face communication alone can account for the increase in cooperation and welfare. One important task of the planned experiments is, therefore, to determine which coordination devices facilitate the achievement of welfare gains through the threat of costly retaliation.

Reciprocal fairness is not only characterized by conditional retaliation but also by conditional kindness. The Ostrom et al. and Fehr-Gächter experiments allow, however, only for retaliation. An interesting question is whether the combination of the opportunity to exchange "gifts" (i. e. the kindness feature of reciprocal fairness) with the opportunity to retaliate leads to improved outcomes in terms of rates of cooperation and the welfare of group members. Previous research indicates that in bilateral relations, positive reciprocity is generally associated with welfare gains. Whether this also holds in the context of n-person social dilemmas remains an open question.

In the Ostrom et al and the Fehr-Gächter experiments groups were relatively small and all group members could monitor and punish the cooperation behavior of all the other group members. This design assumes implicitly a rather dense social structure: every agent can monitor and punish every other agent in the group. In fact, people often interact with a relatively small number of neighbors, so the structure of social interaction is not as dense as assumed in the laboratory experiments. We will investigate how variations in the density of social interaction affect cooperation rates. Such experiments can help to detect those social structures that enhance cooperation and those that inhibit cooperation.

Further questions concern the effects of heterogeneity in endowments and preferences on cooperation. Does equality in endowments enhance cooperation? Does heterogeneity of preferences inhibit cooperation? To what extent do those with a strong preference for the public good force those with a weak preference to contribute? How are cooperation rates and welfare affected if heterogeneity in preferences is combined with mobility across groups? Virtually all of these questions are not yet rigorously examined although they are of fundamental importance if we want to better understand which social structures and institution will foster cooperative outcomes.

Another relevant issue is how social structure affects cooperation. We hypothesize that the more dense the social interaction among agents, the more effective is costly punishment in maintaining cooperation. Consider, for instance, a public goods game with 25 subjects. The return of one token on the public account is 0.08. So if each subject invests the whole

endowment, subjects can double their income relative to a situation where each puts the whole endowment in the private account. Treatment A: Each subject can punish all the other 24 players in the game. Treatment B: Each subject can only monitor and punish a subset of the other players; i.e. there exist "local interactions." The number of agents who can be monitored and punished by each person in the group is thus a proxy for social density or social capital. By varying the social density we can study its impact on cooperation. This may be a particularly important contribution to the current policy analysis of the role of 'social capital' in inducing cooperative behavior in communities. This same research design permits the study of the "spatial" diffusion of cooperation. Suppose, for example, that subjects are spatially ordered on a rectangular grid. Then those in the middle have the largest number of social contacts while those at the corners have a low number of contacts. In the corners it may therefore be easier to free-ride without being punished. Hence, in the corners free-riding may survive while in the center cooperation is maintained.

Finally, we intend to support the extension of experiments concerning reciprocal fairness to social interactions relevant to real world social policy issues, including charitable contributions, taxation, redistributive expenditures (the 'welfare state'), the treatment of social inequality, and criminal sentencing. Such experiments involve adding contextual and interpretive elements to the bare structure of game payoffs. While such additions can lead to a situation where experimental outcomes are subject to multiple interpretations, we believe carefully limiting and controlling the contextual material can avoid this problem.

4 Topic II: Explaining Reciprocal Fairness

Reciprocal fairness as we have defined it is formally an example of the much studied phenomenon of altruism: a behavior costly to the agent that confers benefits on others. However there is a major difference: much of the evidence for reciprocity concerns people's willingness to inflict costly punishment upon others who have wronged them, their loved ones, or 'society.' Revenge is an example of reciprocal fairness, as are the violent responses to personal insult that characterize what Richard Nisbet and Dov Cohen call the *culture of honor* that is common in herding societies and among whites in the United States South. The benefits conferred upon others are thus indirect rather than direct: retaliatory behavior hurts the object of the agent's ire, but helps the group by inducing a reduction in the behavior inviting retaliation.

Evidence that reciprocal fairness is a ubiquitous behavioral pattern, should the research of topic I support this hypothesis, then poses a puzzle. How could costly punishment and other forms of reciprocal fairness evolve? The problem is not new; Hume rhetorically asked:

> Who sees not that vengeance, from the force alone of passion, may be so eagerly pursued as to make us knowingly neglect every consideration of ease, interest, or safety?

Formally, the puzzle is this. Costly punishment, and the various other forms of reciprocal fairness need not be, and generally are not in the interest of the actor at the time the action is to be taken. If actors have outcome-based preferences that are entirely self-regarding, the threat to retaliate is not generally credible, and an equilibrium in which behaviors characterized by reciprocal fairness are present may not be subgame perfect. An obvious resolution of the puzzle is to posit that people take pleasure in inflicting harm on those who have done harm to them or to others or are driven to do so by a sense of obligation, honor, or dignity. But this simply displaces the puzzle. How could preferences or compulsions of this type have evolved?

We will attempt to provide an answer using models of genetic evolution, cultural evolution, and gene-culture co-evolution. Thus we will support the development of evolutionary models of preference formation that determine conditions under which reciprocal fairness might emerge. Such research involves relatively uncharted territory, although there has been some research into the evolutionary emergence of the rate of time preference, degree of risk aversion, and altruistic behavior, using group selection arguments (Hansson and Stuart 1990, Rogers 1994, Mailath, Samuelson and Shaked 1995, Robson 1995, Robson, Bergstrom and Prichard 1996) and kin selection (Bergstrom and Stark 1993, Bergstrom 1995).[10]

An obvious candidate for a mechanism explaining the emergence of reciprocal preferences is group selection operation on either cultural or genetic transmission mechanisms. Why might some genetic basis be contemplated for so complex a social behavior as reciprocal fairness? The answer is that

[10]The most notable attempt to explain retaliation is doubtless Sethi and Somanathan (1996), who use neither group selection, nor local interactions, nor other forms of heterogeneity to prove that under the appropriate conditions costly retaliation against defectors can sustain a cooperative equilibrium in a common pool resource game. However their result depends on the absence of mutation or what we consider to be implausible patterns of mutation—in particular, that cooperators mutate into retaliators at a sufficiently high rate to squelch the emergence of defectors. More important, their model implies a very low level of defection and retaliation, whereas we believe human societies exhibit high levels of defection and very high levels of reciprocal fairness.

formal institutions involved in the transmission of culture in Western societies do not 'teach' reciprocal fairness. The norm of 'returning kindness with kindness' is widely shared and promulgated, but the norm of 'returning evil for evil' is held to betray a low level of moral reasoning and behavior. Indeed, few forms of retaliatory behavior are considered praiseworthy, although some are widely understood and tolerated (for instance, we sympathize with the efforts of the victim of a crime to increase the severity of punishment imposed upon the perpetrator of the crime). By contrast, of course, victims of crimes are praised for their pro-social efforts to reduce the *general level* of a type of criminal activity. This however is not a case of reciprocal fairness, since it could easily be understood in terms of a reputation effect—people are rewarded for being 'good citizens' in opposing anti-social activity.

While biologists have expressed considerable skepticism concerning the ability of group selection arguments to explain altruism in most species,[11] there may be unusual characteristics of *Homo sapiens* and our close ancestors allowing biological group selection to work with greater force. Among these distinct characteristics is the superior ability of *Homo sapiens* to maintain group membership boundaries and practice exclusion thus reducing the level of inter group mobility and enhancing the force of group selection.

Similarly, while recent empirical work on cultural group selection (Soltis, Boyd and Richerson 1995) suggests that the process may work very slowly, we are not persuaded that this counts as an argument against the force of cultural group selection for traits of the kind we are studying. Further, as the proposed research of Robert Boyd suggests, the assertion that cultural group selection is slow-moving may be a model-specific result stemming from particular assumptions concerning intergroup migration. To address these possibilities, we will extend the group selection work of Boyd and his coauthors, combining it with other research (Bowles and Gintis 1997, Bowles and Gintis 1998b) concerning endogenous group formation, boundary maintenance, and the evolution of cooperation within groups.

We will also explore other mechanisms, including variants of nonrandom pairing of agents other than group selection (e.g., local interaction—see Mailath et al. (1995)) and Herbert Simon's 'pleiotropic' model, whereby

[11] Mathematical biologists have shown that the conditions favoring group selection of altruistic behavior are extremely restrictive, and the behavior of most life forms can be explained without recourse to dynamics based upon altruism or group selection (Williams 1966, Dawkins 1989, Maynard Smith 1976). Nevertheless eusocial nonhuman species have emerged and do very well, and group selection is part of the account of their existence. See Alcock (1993), Ch. 16, for a recent review of the literature. For the case of *Homo sapiens* see Caporael (1987), Simon (1993) and the discussion in Wilson and Sober (1994).

costly but socially valuable traits may evolve by dint of being"tied" either genetically or culturally to other individually beneficial traits.

5 Topic III: Cross-Cultural Research in Reciprocal Fairness

In an effort to explore how humans respond to bargaining situations, economists have administered a number of different experiments. Among the simplest and most widely used of these experiments, the ultimatum game, seems to provide a number of robust and important insights into human economic reasoning that strongly contradict the predictions of standard game theory. Because the ultimatum game yields similar results in many different places (including Taiwan, Israel, Tokyo, Pittsburgh, Slovenia and even in Java), many economists have come to think of this bargaining behavior as a product of some innate human-universal economic reasoning process. For example, Roth (1995) proposed that humans possess an evolved cognitive process that balances a drive for immediate self-interest with a drive to directly punish associates for inequitable transactions. Yet, recent data from experiments performed among the Machiguenga, an indigenous group living in a remote region of the Peruvian Amazon, substantially deviate from the typical ultimatum game responses: modal offers were much smaller, and even very small offers were typically accepted. We interpret these results as indicating that game performance may not result simply from the basic functioning of some innate optimizing psychological machinery, but that, in addition, it may depend on socially transmitted behavioral rules that vary from cultural group to cultural group.

However, in order to administer the game in a small Amazonian village, the usual ultimatum game protocols had to be modified, and it is possible that the very different results were caused by these modifications. We propose to test for this possibility by running a controlled experiment with UCLA graduate students which duplicates the procedures used in Peru. If, as we expect, these graduate students behave much like other American and European subjects in the ultimatum game, and quite different from the Machiguenga, then the hypothesis that cultural differences can affect ultimatum game performance will be strengthened.

If these results hold up, we are left with an interesting and important research question: specifically how do culturally transmitted norms, rules and behaviors interact with innate cognitive economic calculations to produce ultimatum game performance? To address this question, we propose to assemble a group of between six and eight experienced, economically-oriented, researchers that specialize in very different cultural groups, train

them in the theory and methodology of the ultimatum game and other related experimental games, devise a set of experimental variations to test specific hypotheses, and send them off to their field sites across the globe with a common research design intended to acquire cross-cultural, comparative data on ultimatum game performance. When all the researchers have completed their investigation, we will reconvene, share our data, and discuss our analyses. The results and conclusions of this conference, which will be published in an integrated form in an edited volume, should allow us to begin differentiating and delineating the aspects or components of ultimatum game behavior that result from innate pan-human cognitive processors from those aspects rely on culturally evolved behavioral norms or rules.

In the ultimatum game, two players are allotted a sum of money. The first player, often called the proposer, offers a portion of the total sum to a second person, often called the 'responder.' The responder, then, has the opportunity to accept or reject the proposer's offer. If the responder accepts, she receives the amount of the offer and the proposer receives the remainder. If the responder rejects, then nobody gets anything; both the responder and proposer receive zero.

The ultimatum game first sparked the interest of economists because its highly replicable results radically departed from the standard behavior of self-interested income maximizers. Game theory unambiguously predicts that proposers will offer the smallest non-zero amount possible, and responders will always accept. Instead, offers made by proposers typically average between 30 and 40 percent of total, with the modal offer at 50 percent. Responders usually accept the average offers, and will often reject offers lower than 20 percent (Camerer and Thaler 1995).

Experimental economists have systematically studied the influence of various factors on the game's results, including stake size, degree of anonymity, context and 'culture.' Lisa Cameron's (1995) analysis of data from Indonesia, where she was able to provide sums equivalent to approximately three month salary for test subjects, strongly rejects the hypothesis that higher stakes move individuals closer to game-theoretic behavior. In fact, her data suggest that proposers generally move away from game-theoretical predictions and toward a 50-50 split; responders, consequently, accept these proportionately higher offers more frequently. Similarly, Hoffman, McCabe and Smith (1994) tested the effect of raising the stakes from $10 to $100 dollars, and found they could not reject the hypothesis that the offers are identical with $10 stakes and with $100 stakes.

Researchers have also suggested that the experimenter's knowledge of the proposer's behavior may contribute to non-equilibrium (non-rational-

actor-model) results. In several tightly controlled experiments designed to test this hypothesis, Bolton and Zwick (1995) and Bolton, Katok and Zwick (forthcoming) concluded that subject-experimenter anonymity makes little difference. Subjects behaved similarly regardless of the experimenter's knowledge of their behavior.

Experimental context, instructions and game terminology have some minor effects on game performance. Forsythe et al. have shown that using buyer/seller terminology drops the median offer slightly. Similarly, Hoffman et al. (1991) show that when participants believe that some individuals have 'earned the right' to be proposers, offers tend to be lower, and responders are more likely to accept low offers, than when roles are perceived as randomly assigned.

As we have seen, Bolton and Zwick (1995) provide substantial evidence for the importance of punishment opportunities in creating substantial deviations from equilibrium behavior. Bolton and Zwick hypothesized that an important determinant of player 2's behavior is a desire to punish player 1 for an inequitable or unfair division. To test this, they compared their ultimatum results to an identically scripted and structured game called the 'impunity game,' which has the same form as the ultimatum game, except that player 2's rejection does not affect player 1's take; that is, player 1 gets her portion no matter what player 2 does, and player 2 receives either the offered amount (acceptance) or zero (rejection). The impunity game rapidly achieves perfect equilibrium (regardless of experimenter anonymity). They conclude that the absence of punishment opportunities strongly affects game performance.

5.1 The Effect of Environmental Differences on Behaviors

Economists have attempted to investigate the effect of cultural differences on ultimatum game performance. Taken together (Cameron 1995, Kachelmaier and Shehata 1992a, Kachelmaier and Shehata 1992b, Thaler 1988, Roth, Prasnikar, Okuno-Fujiwara and Zamir 1991, Hoffman et al. 1994), this collection of cross-cultural studies shows that 'culture' produces only minor deviations in the game's results. In the largest of these experiments, Roth et al. (1991) conducted a carefully controlled comparative ultimatum game study in Jerusalem, Ljubljana, Pittsburgh and Tokyo. As expected, the results from all four locations differed greatly from game theoretical predictions, but comparatively, they differed only slightly from one another. However, some small, but interesting, differences did appear. For example, Israeli proposers tended to make somewhat lower offers (with a mode of

40%) than proposers in Pennsylvania (who provided a mode of 50%). And, Israeli responders were, on average, willing to accept somewhat lower offers than students at the University of Pittsburgh. Even Cameron's extensive data from students and faculty working at Gadjah Mada University in Yogyakarta (Indonesia), perhaps the best cross-cultural test of the ultimatum game, revealed no significant differences in comparison to data from Roth et al. (1991) or Hoffman et al. (1994). The mean proposer demands, for example, from Indonesia and the U.S. were 0.5734 and 0.5625, respectively.

In total, many factors generate small deviations in ultimatum game performance, but no alterations in experimental variables have produced substantial deviations from the usual results (except for fundamental changes like removing punishment possibilities from the game). After his review of these studies in the Handbook of Experimental Economics (1995), Alvin Roth concludes the following:

> Thus we see here a series of experiments whose results seem to be that even initially very skeptical investigators are becoming persuaded that the experimental results observed in ultimatum games are not easily displaced artifacts of the experimental methods, but rather represent a very robust phenomenon.

Interestingly, in his cross-cultural analysis of the small differences between American and Israeli performance, Roth (1995) suggests that these apparent differences indicate, not a difference in aggressiveness or toughness, but rather a difference in what is perceived as fair, or what is expected under the circumstances. Some economists might suggest that these minor differences in what is 'expected' or 'fair' between cultural groups result from a sort of window-dressing effect that cultural beliefs can have on more fundamental, innate economic reasoning processes. Some new research supports Roth's conclusions, and indicates that culture may be an important variable in understanding game performance.

Recently anthropologist Joe Henrich, working with Robert Boyd, conducted an experiment among the Machiguenga, an Arawakan-speaking indigenous group inhabiting the southeastern Peruvian Amazon, which suggests that cultural differences can have a substantial effect on behavior in the ultimatum game. Traditionally, the Machiguenga live in mobile, single family units or small family hamlets and subsist on a combination of swidden (sometimes called *slash and burn*) agriculture, hunting, gathering and fishing. Within the last 30 years, missionaries, government-sponsored bilingual schools, and markets have sedentized and centralized the Machiguenga in a gradual process of increasing market integration. Currently,

most Machiguenga live in small semi-permanent agricultural communities (of between 250-350 people), grow some cash crops, and subsist primarily on manioc, plantains and some fish (Henrich forthcoming).

In the summer of 1996, during Henrich's third visit to the Machiguenga community of Camisea, he performed a modified version of the ultimatum game experiment. First, he gathered twelve men together between the ages of 18 and 30 under the auspices of "playing a fun game for money." He explained the game to the group in Spanish using a set script written with simple terminology like 'first person' to reference the proposer and 'second person' for the responder. After this he had a bilingual school teacher (an educated Machiguenga) re-explain the game in the Machiguenga language (translating from his script), and display the money that he would be using to make payments. After this, each participant entered Henrich's house (the guest hut) individually, he and the teacher explained the game a third time, and Henrich asked a number of practice hypothetical questions intended to test the participants' comprehension of the game. They re-explained parts of the game as necessary. After the individual confidently answered at least two hypothetical questions correctly, he would submit the actual question with a pile of soles (Peruvian money) in view. The following day, after having successfully gotten 12 responses and paid out some money, he began seeking randomly selected individuals to play the game. Most people had already heard of the game and were eager to play. He privately explained the game to each individual (usually in their house) and ran through the same testing procedure as the previous day.

After three days of doing this he accumulated 32 responses. During this process several people were rejected because they, after 30+ minutes of explanation, could not understand the game (at least they could not answer the hypothetical questions). While typical U.S. results produce a mean offer of 40%, a mode of 50% and few offers below 20%, the Machiguenga proposed a mean offer of 27.5%, with a mode of 25%, and many offers of 15%. Similarly, Machiguenga responders, with one exception, always accepted; many offers of 15% were accepted, whereas, Americans frequently reject offers below 20%. These results seem to be very different from what has been observed elsewhere.

We hypothesize that the Machiguenga behave differently from subjects in other experiments because they are culturally different from those subjects; that is, because they have socially learned different values, beliefs and behaviors from preceding generations. We believe that cultural effects were not detected in previous experiments because the subjects in prior experiments were in fact culturally very similar; all were urban university

students living in a sedentary, literate, market society. Jerusalem, Ljubljana, Pittsburgh and Tokyo represent only a tiny fraction of the range of human cultural variation in any observable dimension. Focusing on kinship systems, religious systems, marriage rules, or property rights tells the same story: industrial societies represent only a small subset of the global cultural diversity.

Why should we expect that norms governing economic transactions to be any different? If one accepts that many aspects of culture are adaptive responses to local conditions, this makes sense. The urban, literate, market-dominated world shared by previous subjects represents only a small fraction of the spectrum of human life ways. For example, until very recently, the Machiguenga were nomadic subsistence horticulturalists without a cash economy, written language, or established government. It is quite plausible that Machiguenga behave differently in the ultimatum game because the culturally transmitted beliefs and values that evolve in such an environment are very different from the beliefs and values that characterize urban, literate, industrialized societies.

5.2 Research Method in a Small Scale Society

A preliminary analysis of the ultimatum game data collected among the Machiguenga Indians of the southeastern Peruvian Amazon suggests that culture may have a much greater effect on behavior than was previously supposed. This result suggest two further lines of research: a control experiment to confirm that the Machiguenga behavior was not an artifact of the nonstandard experimental procedures, and, a expanded program of cross-cultural experimental economics research aimed at determining which cultural factors are important in affecting ultimatum game performance.

In order to administer the experiment in the small scale Machiguenga society, usual experimental protocol had to be modified, and it could be that these results are different for this reason. In order to control for the effects of such experimental factors which may be contributing or creating these behavioral differences, we propose to repeat, as accurately as possible, the research performed among the Machiguenga with students at the University of California, Los Angeles. Repeating the same experiment, with the same protocol, the same experimenter, similar stakes and identical procedures will allow us to eliminate potentially confounding influences, and, depending on the results, demonstrate the importance of cultural differences in understanding ultimatum game performance It is important to see that redoing the experiment with the Machiguenga using more typical stakes

would not serve the same purpose as the proposed control experiment. The proposed experiment is meant to control for as much as possible, including experimenter effects and the specifics of a cross-cultural protocol, not the effect of stakes. Moreover, given the cost of Amazonian research, redoing the Machiguenga study would also be much more expensive.

In comparison to most previous ultimatum game procedures, the Machiguenga experiment used relatively high stakes. The base sum was 20 soles ($8.40), which represents about 2.3 days pay for a Machiguenga male doing wage labor. We propose offering graduate students approximately $170, or roughly 2.3 days pay (readers, for example, make about $9.50/hour). This sets the stakes high (in comparison to the usual amount), and approximately equal to the Machiguenga experiment. If the Machiguenga were only responding to high stakes, then UCLA graduate students should perform similarly.

The Machiguenga seemed confident of anonymity among participants (in fact, they did not seem to care about anonymity at all), but each participant knew she or he was playing against someone else in a small community of approximately 70 adults, so it was a guarantee they knew the other player, they just did not know specifically who it was. Again, by using only UCLA anthropology graduate students, which number approximately 80, and by making this known to all participants, we hope to create a comparable social situation. Participants will be assured of anonymity, but they will know they are playing against someone else they know.

The 32 student participants will be randomly selected from a stratified list of all anthropology graduate students to match gender and age differences reflected in the Machiguenga sample. That is, the total list of anthropology graduate students will be subdivided by gender and age. The appropriate number of participants will be randomly selected from each age/gender subdivision. This should mitigate any argument that differences in performance between Machiguenga and UCLA students arise from variations in the gender/age composition of the samples.

All other conditions of the Machiguenga experiment will be approximated as closely as possible. We will use the identical game description, terminology and explanatory examples (except in English, rather than Spanish or Machiguenga). In designing the Machiguenga experiment Henrich intentionally avoided any complex or suggestive terminology like 'buyer' and 'seller' or 'proposer' and 'responder,' and instead opted for labels like '1st person' and '2nd person.' Because he initially avoided words imbued with implicit cultural assumptions, translations should proceed smoothly and with little distortion. As he did with the Machiguenga, he plans to

meet with each participant privately, explain the game individually, make certain they understand the game with a few hypothetical practice questions, and present them with the real situation while the sum of money is in view. If graduate students, when faced with experimental conditions and circumstances nearly identical to those encountered by the Machiguenga, perform similar to all other groups of American university students, and the not like the Machiguenga, then researchers will have to entirely reconsider the relative importance of cultural differences in making economic decisions.

Such a result would have important implications for economics. It would suggest that cultural differences, perhaps related to expectations about fairness and punishment, influence behavior in the ultimatum game more strongly than all other variables combined. Thus, to the extent that behavior in experiments is relevant to other behavior, this result would suggest that economic behavior is more than simply a product of innate processes. Rather, economic behavior could only be understood as influenced by culturally transmitted beliefs and preferences that set expectations, define fairness and promote punishment. To account for the role of cultural differences in economic models, economists would need to ask where these culturally transmitted beliefs and preferences come from, why some beliefs and preferences proliferate under certain conditions, and why some are maintained through time.

5.3 Expanded Cross-Cultural Research

If the control experiment confirms our suspicions about the Machiguenga data (that it is an effect of culture), we propose to assemble a group of between six and eight experienced, economically-oriented, researchers that specialize in very different cultural groups, train them in both ultimatum game theory and methodology, devise a set of experimental variations to test specific hypotheses, and send them off their field sites cross the globe with a common research design intended to acquire cross-cultural, comparative data on ultimatum game performance. When all the researchers have completed their investigation, we will reconvene, share our data, and discuss our analyses. The results and conclusions of this conference, which will be published in an integrated form in an edited volume, should allow us to begin differentiating and delineating the aspects or components of ultimatum game behavior result from innate pan-human cognitive processors from the aspects rely on culturally evolved behavioral norms or rules.

The research group will design ultimatum experiments to distinguish between alternative hypotheses about how culture affects game performance,

and which recurrent aspects of game performance result from innate reasoning. These experiments will be designed to address questions such as: Does every culture punish 'unfair' offers in at least some contexts, or are there some groups that never punish? How much cultural variation is there in what constitutes 'fair' and 'unfair?' How does what is fair vary with context? Are there cross-cultural recurrent patterns of 'fairness'; that is, does 'fairness' follow any predictable variation along some dimension as we move from one society to another. Do some cultural groups maintain standards are fairness without any enforcement by punishment. Or, in game terms, do most proposers provide amounts near 50/50 even though proposers who propose less are rarely (or never) rejected. The following discussion proposes several research possibilities that may be tested with an appropriate experimental design.

If economics is like other components of human behavior, then certain context-specific elements may strongly influence behavioral responses. Analytically, context may be subdivided into three components: the medium of exchange, the sphere of exchange, and the situational characteristics of the exchangers. In addition, cultural evolution may have generated norms about fairness and norms that evoke punishment for different contexts. What is fair when trading jaguar skins for steel tools may not be what is fair when exchanging meat for sorghum. New circumstances or non-traditional items may not be governed by the same rules as traditional items or recurrent situations.

Medium of exchange. Economic theory typically assumes that behavioral responses should be independent of the medium of exchange. It should not matter very much whether players in an ultimatum game are dividing a pot of ten dollars or ten candy bars that can be exchanged for a dollar each. Daily experience, however, contrasts with this assumption. Suppose you ask a friend to drive you to the airport, and he says that he cannot because he wants to watch his favorite TV game show, and instead, he offers you $30, the cost of a cab ride to the airport. Is this socially acceptable? Does this friend suffer any loss in your eyes? Will you offer this friend cab fare when he has to go to the airport? To test the importance of the medium of exchange in influencing ultimatum behavior, the game can be played with different mediums, including cash, food, service time and symbolic items. For example, among a group of Amazonian Indians called the Machiguenga, Joe Henrich could perform the identical ultimatum experiments with cash and meat. Cash is relatively new to the Machiguenga (appearing only within the last thirty years), while it is very likely they have been sharing meat for many millennia. Thus even though the Machiguenga definitely understand

the value of money, in terms of what it can buy, the social rules for meat distribution may be quite different from the rules for cash distribution.

Sphere of exchange. The sphere of exchange may influence game performance. Although the game is usually played anonymously, people may have implicit assumptions about the pool of potential participants (the sphere) with which they are re-playing. Most experiments are performed at universities, so student participants may accurately assume they are paired with another student, and perhaps a class mate. In the Machiguenga experiment, individuals were told that they played with someone else in their community (of 300 people). How would it affect the results if we controlled for players assumptions about the sphere of exchange? Do cultural differences affect how people react to different spheres? Many villages (in Amazonian, Africa and New Guinea, for example) are subdivided into clans or moieties. What if we changed the sphere from the village (say about 500 people) to the clan (150 people) or moiety (250 people)? Does the sphere affect the tendency to make equal divisions or the desire to punish 'unfair' offers? How does ethnicity affect the equation: What if we told Mongolian pastoralists they were playing against neighboring but ethnically distinct Khazaks? Would they play differently against other Mongols, then they would against anonymous Khazaks?

Situational characteristics. Moreover, the situational characteristics of exchangers may influence ultimatum game behavior. Individuals may use contextual clues about the status or character of the other anonymous player in accessing their behavior choices, and the rules used in assessing others may be culturally transmitted. In Western industrial society, for example, Hoffman et al. (1991) shows that when participants believe that some individuals have 'earned the right' to be proposers, offers tend to be lower, and responders are more likely to accept low offers, than when roles are perceived as randomly assigned. The perceived status of the proposer may suggest to westerners that this individual deserves more than would otherwise be acceptable. Both the contextual situation (e.g. 'having earned the right') which evokes the modification in what is normatively acceptable, and the strength of that modification may be culturally transmitted. Among some cultures, it may be that apparently skilled or 'deserving' individuals are not permitted larger shares in exchange situations. In many foraging societies for example, the hunter that made the kill (the more 'deserving' by our norms) is not permitted to partake of the meat, or he may be only entitled to eat only a small or equal portion of the meat. Meanwhile, unsuccessful hunters are often apportioned generous shares. We can test this kind of effect by incorporating the same kind of information (about the status of

characteristics of other players, but keeping them anonymous) as Hoffman et al., but do it in culturally different situations. The details of such an experiment would be worked out by the research group, thereby ensuring the feasibility, comparability and cultural sensitivity of the test.

Economic and ecological correlates. In addition to exploring the cultural variability of context-specific factors, this research will explore how different ecological factors direct the evolution of norms and behavior that influence ultimatum performance. For example, economic anthropologists (Plattner 1989, Johnson 1989) have suggested that people's economic behavior is strongly affected by their degree of economic independence or self-sufficiency. Nomadic peoples like the Machiguenga who rely almost entirely on family-based production and kin-based exchange, and thus are much less tightly imbedded in anonymous exchange networks, will be less likely to expect 'fair' (according to our cultural norms) behavior in the ultimatum game. Sedentary, more economic interdependent, peoples establish longer term relationships with their neighbors, and therefore are more likely to expect fair behavior in the ultimatum game. Note, we are not proposing that participating in different economic systems evokes different individual-level economic calculations or notions of fairness. We are saying that certain economic systems provide conditions in which different norms about fairness and expectations of fairness will evolve in time through the differential transmission of behaviors/ideas (see Boyd and Richerson 1985).

To explore these possibilities, we propose to choose field sites so that people share some ecological and economic features but not others. For example, to test the our hypothesis that a group's degree of economic independence affects norms about sharing, we would gather ultimatum game data from Peruvian peasant agriculturalists who share an identical environment (both economically and ecologically) with the Machiguenga, Tibetan pastoralists who share little with the Machiguenga except a nomadic life style, family-based mode of production and kin-based exchange networks, Mapuche households who are highly economically independent but sedentary, and Cuban socialists. If our understanding of the relationship between culturally transmitted things, like expectations of fairness, and economic systems is correct, then the Machiguenga's peasant neighbors should behave like American undergraduates and Cuban socialists, while Tibetan pastoralists and Mapuche farmers should behave like Machiguenga. Each population was selected to inform a specific aspect of the overall project, as well as the practical difficulties of population access, contacts and familiarity; this research will not be done through universities, with students or in urban centers. Each experiment will use similar stakes, the identical protocol as

was used with both UCLA graduate students and Machiguenga Indians.

Working together, the research group will design an experimental protocol for the ultimatum game based on protocol used by Henrich among the Machiguenga. This group of experienced field workers will be a able to anticipate potential difficulties in administering the game in their particular communities, and the group will be able to discuss and agree on strategies for adapting the game to particular ethnological circumstances.

We have designed this research around ongoing projects and experienced researchers for a number of reasons. First, using existing projects greatly facilitates access to populations. Research among real populations, inhabiting inhospitable regions in politically delicate climates (where most people live) can be extremely difficult, time consuming and dangerous without an experienced researcher who has friends and contacts in the area, speaks the language, knows the local customs and political pitfalls. Second, ongoing research supplies key ethnographic data that will be essential to contextualizing and interpreting our results. Third, gaining a sufficient trust from a group takes months, and sometimes years. But, by cooperating with anthropologists who have accumulated a group's trust through years of interaction, we will be able to use the Foundation's money and our time most efficiently. So far we have found a number of interesting and well-qualified researchers who can give us access to nomadic pastoralists in Tibet, sedentary pastoralists in Mongolia and Tanzania, Cuban collectivists, and horitculturalists living ing Peru and on the Indonesian island of Sulawesi.

6 Topic IV: The Implications for Social Policy

Egalitarian sentiment is often based on a commitment to 'reciprocal fairness.' Social policies designed to tap this sentiment motives may succeed where others, equally egalitarian in intent, are politically unpopular.

Both historical and contemporary experimental evidence support this position. In his magisterial Injustice: the Social Bases of Obedience and Revolt, Barrington Moore (1978) sought to discern if there might be common motivational bases—"general conceptions of unfair and unjust behavior" (21)—for the moral outrage fueling struggles for justice that have recurred throughout human history. "There are grounds," he concludes from his wide-ranging investigation,

> for suspecting that the welter of moral codes may conceal a certain unity of original form...a general ground plan, a conception of what social relationships ought to be. It is a conception that

> by no means excludes hierarchy and authority, where exceptional qualities and defects can be the source of enormous admiration and awe. At the same time, it is one where services and favors, trust and affection, in the course of mutual exchanges, are ideally expected to find some rough balancing out. (4-5,509)

Moore termed the general ground plan he uncovered

> ...the concept of reciprocity—or better, mutual obligation, a term that does not imply equality of burdens or obligations... (506)

One standard explanation for the fact that individuals vote for redistributive expenditure and voluntarily contribute to private charities is that the welfare of others is an argument in the utility function (Hochman and Nitzan 1985). However it is well known that predictions based on this concept of altruism do not describe charitable behavior well (Sugden 1982, Sugden 1985) and this form of altruism has perverse efficiency effects (Buchanan 1975, Bernheim and Stark 1988, Lindbeck 1988, Stark 1990, Bruce and Waldman 1990, Hori and Kanaya 1989, Kranich 1988). Another standard explanation is insurance: people vote for redistributive expenditures that they might not current receive, but could receive under some likely future states of nature. However much of the support for redistribute expenditure for the poor is difficult to justify in this manner.

Can a theory of reciprocity contribute to an explanation of charitable giving and support for redistributive expenditure? We intend to commission two or three papers to address this issue. Potential researchers in this area are Robert Sugden and Martin Gilens, who have written extensively on the problem, and Christopher Jencks, who uses the concept in his analysis of the 'revolt against the welfare state,' but does not employ formal models of reciprocal fairness.[12]

We expect that reciprocal fairness may be able to explain both the rise of the welfare state after World War II, and the more recent tide of opposition to redistributive policies in recent years. Specifically, in light of the experimental regularities outlined above, we suspect the following to be true as well: redistributive policies that reward people independent of whether and how much they contribute to society are considered unfair and are not supported, even if the intended recipients are otherwise worthy of support,

[12]Samuel Bowles and Herbert Gintis have argued for a reciprocity theory of giving (Bowles and Gintis 1998c), but have not developed formal models of the phenomenon.

and even if the incidence of defection in the target population is not particularly high. This would explain the opposition to many welfare measures for the poor, particularly since such measures have been to some extent opportunistically exploited, and are thought to have facilitated various social pathologies. At the same time it explains the continuing support for social security and medicare in the United States, since the public perception is that the recipients are 'deserving.'[13]

Our artificial life simulations suggest that the moral bases of social policy may follow a dynamic cycle. In modeling the repeated Prisoner's Dilemma, we have found that in periods in which levels of defection are high, reciprocal strategies (e.g., tit-for-tat) spontaneously emerge as individually successful strategy that leads to a very high level of cooperation. In this highly cooperative state, however, the attractive features of tit-for-tat disappear, and there is a movement towards unconditional cooperation that in turn invites high levels of defection and non-cooperative behavior. This state of affairs is then conducive to the return of reciprocating strategies, thus completing the dynamic cycle.

The cross-cultural research we will undertake also has important policy implications. Distinguishing the effects of culture from pan-human cognitive characteristics has important consequences for economic theory. The rational actor model carries with it the implicit assumption that people everywhere reason the same way, and the results of economic experiments are sensibly interpreted as revealing facts about these reasoning processes. However, if human economic behavior is affected by the cultural milieu, then theories of human economic reasoning that neglect of the cultural environment are incomplete. Thus, such experiments suggest that economic theory should be extended to consider the reciprocal effects of economic institutions and culturally transmitted beliefs.

Application-oriented scientists seeking to positively influence economic policy will have to model both innate human economic calculations and population-level process of cultural evolution. In general it will not be possible to predict patterns of behavior without understanding the cultural transmission of the norms, values, and rules that influence economic behavior. For example, development economists need to understand that millions of people in underdeveloped and developing nations may not respond as they anticipate because these peoples possess culturally evolved norms and

[13]Evidence for this view is presented in Gilens (1996). Similarly, Citrin and Green (1990) uses voting data and other statistical data to argue that only to a very small degree does self-interest explain to the political preferences of Americans in many policy areas, including cash grants based on means-tested criteria.

rules that contrast greatly with the norms and rules possessed by western industrial peoples; the environment in which most of these theories were generated. Further, culture is an evolutionary process, and consequently the norms and behavioral rules that govern ultimatum game performance in western society may rapidly change with the rise of novel social intuitions. Economic theories founded on culturally-transmitted behavioral rules, but lacking any mechanisms for the temporal dynamics of those rules, will lose their explanatory power as our culture evolves and our society changes.

References

Abbink, Klaus, Gary E. Bolton, Abdolkarim Sadrieh, and Fang-Fang Tang, "Adaptive Learning versus Punishment in Ultimatum Bargaining," 1996. Discussion Paper No. B0-381, University of Bonn.

Alcock, John, *Animal Behavior: An Evolutionary Approach* (Sinauer Associates, 1993).

Andreoni, James, "Why Free Ride? Strategies and Learning in Public Good Experiments," *Journal of Public Economics* 37 (1988):291–304.

Arrow, Kenneth J. and Frank Hahn, *General Competitive Analysis* (San Fransisco: Holden-Day, 1971).

— and Gerard Debreu, "Existence of an Equilibrium for a Competitive Economy," *Econometrica (*1954):265–290.

Axelrod, Robert, *The Evolution of Cooperation* (New York: Basic Books, 1984).

— and Douglas Dion, "The Further Evolution of Cooperation," *Science* 242 (1988):1385–1390.

— and William D. Hamilton, "The Evolution of Cooperation," *Science* 211 (1981):1390–1396.

Bendor, Jonathan and Dilip Mookherjee, "Institutional Structure and the Logic of Ongoing Collective Action," *American Political Science Review* 81,1 (March 1987):129–154.

Bergstrom, Theodore C., "On the Evolution of Altruistic Ethical Rules for Siblings," *American Economic Review* 85,1 (March 1995):58–81.

— and Oded Stark, "How Altruism can Prevail in an Evolutionary Environment," *American Economic Review* 83,2 (May 1993):149–155.

Bernheim, B. Douglas and Oded Stark, "Altruism Within the Family Reconsidered: Do Nice Guys Finish Last?," *American Economic Review* 78,5 (December 1988):1034–45.

Bolton, Gary E. and Rami Zwick, "Anonymity versus Punishment in Ultimatum Games," *Games and Economic Behavior* 10 (1995):95–121.

— , Elena Katok, and Rami Zwick, "Dictator Game Giving: Rules of fairness versus random acts of kindness," *International Journal of Game Theory* (forthcoming).

Boorman, Scott A. and Paul Levitt, *The Genetics of Altruism* (New York: Academic Press, 1980).

Bowles, Samuel, "The Production Process in a Competitive Economy: Walrasian, Neo- Hobbesian, and Marxian Models," *American Economic Review* 75,1 (March 1985):16–36.

— and Herbert Gintis, "The Revenge of Homo Economicus: Contested Exchange and the Revival of Political Economy," *Journal of Economic Perspectives* (Winter 1993).

— and — , "Optimal Parochialism: The Dynamics of Trust and Exclusion in Communities," June 1997. University of Massachusetts Working Paper.

— and — , "How Communities Govern: The Structural Basis of Prosocial Norms," in Louis Putterman and Avner Ben-Ner (eds.) *Economics, Values and Organizations* (New York: Cambridge University Press, 1998).

— and — , "The Moral Economy of Community: Structured Populations and the Evolution of Prosocial Norms," *Evolution & Human Behavior* (1998).

— and — , *Recasting Egalitarianism: New Rules for Markets, States, and Communities* (London: Verso, 1998). Erik Olin Wright (ed.).

Boyd, Robert and Peter J. Richerson, "Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizeable Groups," *Ethology and Sociobiology* 113 (1992):171–195.

Bruce, Neil and Michael Waldman, "The Rotten-Kid Theorem Meets the Samaritan's Dilemma," *Quarterly Journal of Economics* 105,1 (February 1990):155–65.

Buchanan, James, "The Samaritan's Dilemma," in Edmund S. Phelps (ed.) *Altruism, Morality, and Economic Theory* (Russell Sage, 1975).

Camerer, Colin and Richard Thaler, "Anomalies: More Ultimatum and Dictator Games," *Journal of Economic Perspectives* 9,2 (1995):209–219.

Cameron, Lisa, "Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia," 1995. Discussion Paper #345, Department of Economics, Princeton University.

Caporael, L. R., "Homo Sapiens, Homo Faber, Homo Socians: Technology

and the Social Animal," in W. Callebaut and R. Pinxten (eds.) *Evolutionary Epistemology: A Multiparadigm Program* (Reidel, 1987).

Citrin, Jack and Donald Philip Green, "The Self-Interest Motive in American Politics," *Research in Micropolitics* 3 (1990):1–28.

Cosmides, Leda and John Tooby, "Cognitive Adaptations for Social Exchange," in Leda Cosmedes Jerome H. Barkow and John Tooby (eds.) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (New York: Oxford University Press, 1992) pp. 163–228.

Cremer, Jacques, "Cooperation in Ongoing Organizations," *Quarterly Journal of Economics* 101 (1986):33–49.

Davis, Douglas D. and Charles A. Holt, *Experimental Economics* (Princeton: Princeton University Press, 1993).

Dawes, Robyn M., J. C. Van de Kragt, and John M. Orbell, "Not Me or Thee, but We: The Importance of Group Identity in Eliciting Cooperation in Dilemma Situations: Experimental Manipulations," *Acta Psychologica* 68 (1988):83–97.

Dawkins, Richard, *The Selfish Gene, 2nd Edition* (Oxford: Oxford University Press, 1989).

Debreu, Gerard, *Theory of Value* (New York: Wiley, 1959).

Eckel, Catherine and Philip Grossman, "The Relative Price of Fairness: Gender Differences in a Punishment Game," *Journal of Economic Behavior and Organization* 30,2 (1996):143–158.

— and —, "Chivalry and Solidarity in Ultimatum games," February 1997. Virginia Polytechnic Institute, Working Paper #E92-23.

— and Phillip Grossman, "Are Women Less Selfish Than Men?: Evidence from Dictator Games," June 1996. Department of Economics Working Paper, Virginia Polytechnic Institute.

Fehr, Ernst and Jean-Robert Tyran, "Institutions and Reciprocal Fairness," *Nordic Journal of Political Economy* (1996).

— and Simon Gächter, "Cooperation and Punishment," 1996. Working Paper, Institute for Empirical Economic Research, University of Zürich.

—, —, and Georg Kirchsteiger, "Reciprocity as a Contract Enforcement Device," *Econometrica* (forthcoming 1997).

Forsythe, Robert, Joel Horowitz, N. E. Savin, and Martin Sefton, "Replicability, Fairness and Pay in Experiments with Simple Bargaining Games," *Games and Economic Behavior* 6,3 (May 1994):347–369.

Fudenberg, Drew and Eric Maskin, "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica* 54,3

(May 1986):533–554.

— and Jean Tirole, *Game Theory* (Cambridge, MA: MIT Press, 1991).

Gale, John, Kenneth Binmore, and Larry Samuelson, "Learning to be Imperfect: the Ultimatum Game," *Games and Economic Behavior* 8,1 (January 1995):56–90.

Gardner, Roy, Elinor Ostrom, and James Walker, "The Nature of Common Pool Resource Problems," *Rationality and Society* 2 (1990):335–358.

Gilens, Martin, "Economic Self-Interest and Opposition to Welfare," 1996. Annual Meetings of the APSA.

Gintis, Herbert, "Financial Markets and the Political Structure of the Enterprise," *Journal of Economic Behavior and Organization* 1 (1989):311–322.

Grafen, Alan, "The Hawk-Dove Game Played between Relatives," *Animal Behavior* 27,3 (1979):905–907.

Grossman, Sanford and Oliver Hart, "An Analysis of the Principal-Agent Problem," *Econometrica* 51 (1983):7–45.

Grossman, Sanford J. and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94,4 (August 1986):691–719.

Guttman, Joel M., "Matching Behavior and Collective Action," *Journal of Economic Behavior and Organization* 7 (1986):171–198.

— , "A Non-Cournot Model of Voluntary Collective Action," *Economica* 54,213 (February 1987):1–19.

Hamilton, W. D., "The Evolution of Altruistic Behavior," *American Naturalist* 96 (1963):354–356.

— , "The Genetical Evolution of Social Behavior," *Journal of Theoretical Biology* 37 (1964):1–16,17–52.

Hansson, Ingemar and Charles Stuart, "Malthusian Selection of Preferences," *American Economic Review* 80,3 (June 1990).

Hart, Oliver and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy* 98,6 (December 1990):1119–1158.

Henrich, Joe, "Market Incorporation, Agricultural Change and Sustainability among the Machiguenga Indians of the Peruvian Amazon," *Human Ecology (*forthcoming).

Hirshleifer, Jack and Eric Rasmusen, "Cooperation in a Repeated Prisoners' Dilemma with Ostracism," *Journal of Economic Behavior and Organization* 12 (1989):87–106.

Hochman, Harold M. and Shmuel Nitzan, "Concepts of Extended Preference," *Journal of Economic Behavior and Organization* 6,2 (June

1985):161–76.

Hoffman, Elizabeth, Kevin McCabe, and Vernon L. Smith, "On Expectations and the Monetary Stakes in Ultimatum Games," *International Journal of Game Theory* (1994).

— , — , and — , "Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology," April, 1996. Unpublished.

Hölmstrom, Bengt, "Moral Hazard and Observability," *Bell Journal of Economics* 10,1 (Spring 1979):74–91.

— , "Moral Hazard in Teams," *Bell Journal of Economics* 7 (1982):324–340.

— and Paul Milgrom, "The Firm as an Incentive System," *American Economic Review* 84,4 (September 1994):972–991.

Hori, Hajime and Sadao Kanaya, "Utility Functionals with Nonpaternalistic Intergenerational Altruism," *Journal of Economic Theory* 49,2 (December 1989).

Isaac, R. Mark and James M. Walker, "Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism," *Economic Inquiry* 26 (1988):585–608.

— and — , "Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism," *Quarterly Journal of Economics* 103 (1988):179–200.

Johnson, Allen, "Horticulturalists: Economic Behavior in Tribes," in Stuart Plattner (ed.) *Economic Anthropology* (Stanford: Stanford University Press, 1989) pp. 49–77.

Kachelmaier, S. J. and M. Shehata, "Culture and Competition: A Laboratory Market Comparison between China and the West," *Journal of Economic Behavior and Organization* 19 (1992):145–168.

— and — , "Examining Risk Preferences Under High Monetary Incentives: Evidence from the People's Republic of China," *American Economic Review* (1992):1120–1141.

Kranich, Jaurence J., "Altruism and Efficiency: A Welfare Analysis of the Walrasian Mechanism with Transfers," *Journal of Public Economics* 36,3 (August 1988):369–386.

Kreps, David M., *A Course in Microeconomic Theory* (Princeton: Princeton University Press, 1990).

Ledyard, J. O., "Public Goods: A Survey of Experimental Research," in J. H. Kagel and A. E. Roth (eds.) *The Handbook of Experimental Economics* (Princeton: Princeton University Press, 1995) pp. 111–194.

Levine, David K., "Modeling Altruism and Spitefulness in Experiments," 1996. Department of Economics, UCLA.

Lindbeck, Assar, "Altruism and Time Consistency: The Economics of Fait Accompli," *Journal of Political Economy* 96,6 (December 1988):1165–82.

Mailath, George J., Larry Samuelson, and Avner Shaked, "Evolution and Endogenous Interactions," 1995. Department of Economics, University of Pennsylvania.

Marwell, G. and R. Ames, "Experiments on the Provision of Public Goods," *American Journal of Sociology* 84 (May 1979):1335–1360.

Mas-Colell, Andreu, *The Theory of General Economic Equilibrium: A Differentiable Approach* (Cambridge: Cambridge University Press, 1985).

Maynard Smith, John, "Group Selection," *Quarterly Review of Biology* 51 (1976):277–283.

Moore, Jr. Barrington, *Injustice: The Social Bases of Obedience and Revolt* (White Plains: M. E. Sharpe, 1978).

Nowak, Martin and Karl Sigmund, "A Strategy of Win-Stay Lose-Shift that Outperforms Tit-for-Tat in the Prisoner's Dilemma Game," *Nature* 364 (July 1993):56–58.

Ostrom, Elinor, James Walker, and Roy Gardner, "Covenants with and without a Sword: Self-Governance is Possible," *American Political Science Review* 86,2 (June 1992):404–417.

Plattner, Stuart, "Markets and Marketplaces," in Stuart Plattner (ed.) *Economic Anthropology* (Stanford: Stanford University Press, 1989) pp. 171–208.

Robson, Arthur J., "A Biological Basis for Expected and Non-Expected Utility," March 1995. Department of Economics, University of Western Ontario.

—, Carl T. Bergstrom, and Jonathan K. Prichard, "Risky Business: Sex and Selfing in an Unpredictible World," 1996. Department of Economics, University of Western Ontario.

Rogers, Alan, "Evolution of Time Preference by Natural Selection," *American Economic Review* 84,3 (June 1994):460–481.

Roth, Alvin, "Bargaining Experiments," in John Kagel and Alvin Roth (eds.) *The Handbook of Experimental Economics* (Princeton: Princeton University Press, 1995).

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir, "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study," *American Economic Review*

81,5 (December 1991):1068–1095.

Schneider, Friedrich and Werner W. Pommerehne, "Free Riding and Collective Action: An Experiment in Public Microeconomics," *Quarterly Journal of Economics* 96 (November 1981):689–704.

Sethi, Rajiv and E. Somanathan, "The Evolution of Social Norms in Common Property Resource Use," *American Economic Review* 86,4 (September 1996):766–788.

Shapiro, Carl and Joseph Stiglitz, "Unemployment as a Worker Discipline Device," *American Economic Review* 74,3 (June 1984):433–444.

Simon, Herbert A., "Altruism and Economics," *American Economic Review* 83,2 (May 1993):156–61.

Soltis, Joseph, Robert Boyd, and Peter Richerson, "Can Group-functional Behaviors Evolve by Cultural Group Selection: An Empirical Test," *Current Anthropology* 36,3 (June 1995):473–483.

Stark, Oded, "On the Economics of Nonmarket Transfers: the Role of Altruism," *Economic Letters* 33,1 (May 1990):11–16.

Sugden, Robert, "On the Economics of Philanthropy," *Economic Journal* 92,366 (June 1982):341–50.

— , "Consistent Conjectures and Voluntary Contributions: Why Conventional Theory Does not Work," *Journal of Public Economics* 27,1 (June 1985):117–24.

Thaler, Richard, "Anomalies: The Ultimatum Game," *Journal of Economic Perspectives* 2,4 (1988):195–206.

Thompson, Earl and Roger Faith, "A Pure Theory of Strategic Behavior and Social Institutions," *American Economic Review* 71 (1981):366–380.

Trivers, R. L., "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology* 46 (1971):35–57.

Weissing, Franz and Elinor Ostrom, "Irrigation Institutions and the Games Irrigators Play: Rule Enforcement without Guards," in Reinhard Selten (ed.) *Game Equilibrium Models II: Methods Morals and Markets* (Berlin: Springer-Verlag, 1991) pp. 188–262.

Williams, G. C., *Adaptation and Natural Selection: A Critique of some Current Evolutionary Thought* (Princeton, NJ: Princeton University Press, 1966).

Wilson, David Sloan, *The Natural Selection of Populations and Communities* (Menlo Park, CA: Benjamin Cummings, 1980).

— and Elliott Sober, "Reintroducing Group Selection to the Human Behavioral Sciences," *Behavior and Brain Sciences* 17 (1994):585–654.