

Iterated Prisoner's Dilemma with Choice and Refusal of Partners

E. Ann Stanley and Dan Ashlock

Department of Mathematics

Iowa State University

and

Leigh Tesfatsion

Department of Economics and Department of Mathematics

Iowa State University

ISU Economic Report No. 30

Revised February 1993

Iterated Prisoner's Dilemma with Choice and Refusal of Partners

running title: *choice and refusal of partners*

E. Ann Stanley and Dan Ashlock

Department of Mathematics

Iowa State University, Ames, IA 50011

and

Leigh Tesfatsion

Department of Economics and Department of Mathematics

Iowa State University, Ames, IA 50011

ISU Economic Report No. 30

Revised February 1993

Abstract

This paper studies the effects of partner selection on cooperation in an artificial ecology. Agents, represented by finite automata, interact with each other through an iterated prisoner's dilemma (IPD) game with the added feature that players choose and refuse potential game partners on the basis of continually updated expected payoffs. Analytical studies reveal that the subtle interplay between choice and refusal in N -player IPD games can result in various long-run player interaction patterns: e.g., mutual cooperation; mixed mutual cooperation and mutual defection; parasitism; and/or wallflower seclusion. Simulation studies indicate that choice and refusal can accelerate the emergence of cooperation in evolutionary IPD games. More generally, however, choice and refusal can result in the emergence and persistence of multiple payoff bands, reflecting the possible existence of ecological attractors characterized by play behavior that is not entirely cooperative. The existence of a spectrum of payoff bands in turn leads to the emergence of new ecological behaviors such as band spiking and band tunneling.

Key Words: Iterated prisoner's dilemma; choice and refusal of partners; evolution of cooperation; genetic algorithm; sequential game; artificial ecology; finite automata.

To appear in: Artificial Life III, Santa Fe Institute Studies in the Sciences of Complexity, ed. by Chris Langton, Addison-Wesley, Vol. 16, 1993.

1. INTRODUCTION

This paper investigates cooperative behavior in an artificial ecology in which egoistic agents interact with each other through a tournament of iterated games. The tournament played is a variant of the Axelrod (1984) iterated prisoner's dilemma (IPD) game, modified to permit players to choose and refuse potential game partners. In each iteration, players use expected payoffs updated on the basis of past encounters to make prisoner's dilemma game offers to a limited number of preferred players, and to refuse prisoner's dilemma game offers from unacceptable players.

The introduction of choice and refusal fundamentally modifies the way in which players interact in the IPD game, and the characteristics which result in high payoff scores. Choice allows players to increase their chances of encountering other cooperative players. Refusal gives players a way to protect themselves from defections without having to defect themselves. Ostracism of defectors occurs endogenously as an increasing number of players individually refuse their game offers. But choice and refusal also permit clever ripoff players to home in quickly on exploitable players and form parasitic relationships.

Following Miller (1989), player strategies for the IPD game with choice and refusal (or "IPD/CR game" for short) are represented by means of finite automata.¹ This representation has a number of advantages over the original code-based formulation of Axelrod (1984). It permits the same modelling of complex strategic behavior, but it is simpler and cleaner to program. Moreover, behavior modification in response to endogenously occurring events can readily be incorporated in the form of genetic algorithms.² Consequently, the basic single-tournament IPD/CR game is easily generalized to an "evolutionary" IPD/CR game, i.e., a multiple-tournament IPD/CR game in which the strategies of the players evolve between tournaments.

The choice/refusal mechanism is characterized by six parameters: initial expected payoff; the minimum tolerance (expected payoff) level below which game offers will be refused; the maximum number of game offers which can be made in each iteration; the rejection payoff received when a game offer is refused; the wallflower payoff received when game offers are neither made nor accepted; and a memory weight which determines the relative importance of distant to recent payoffs in the calculation of updated expected

¹A "finite automaton" is a system specified by a finite collection of internal states together with a state transition function, driven by input, which gives the next internal state the system will enter. See section 5, below, for a detailed discussion of the specific finite automaton representation used in this paper.

²A "genetic algorithm" uses a Darwinian selection principle to optimize a solution to a problem with respect to various selected problem features. Genetic algorithms are powerful tools for evolving high-performance strategies from simple representative strategy types; see Holland (1992).

payoffs. Analytical parameter sensitivity studies are undertaken for a variety of single-tournament 2-player IPD/CR games. To simplify and systematize the analysis, we restrict our attention to six illustrative player types that roughly span the range from uncooperative to cooperative behavior. For each pair of player types, we determine the precise conditions under which refusal first occurs, and the average payoff scores which are achieved, as a function of the choice/refusal parameters.

We then extend the analysis to single-tournament N -player IPD/CR games. The interplay between choice and refusal can be quite subtle for such games. We illustrate this with a detailed study of a single-tournament 5-player IPD/CR game in which the ultimate pattern of player interactions can include the formation of successful long-term parasitic relationships unless the minimum tolerance level is set suitably high.

We also report on simulation experiments carried out for the evolutionary 30-player IPD game studied by Miller (1989), modified to allow for choice and refusal of game partners. The experimental results indicate that, in comparison to Miller's findings, the emergence of cooperation is accelerated over much of the choice/refusal parameter space. However, high enough values for the minimum tolerance level and the wallflower payoff can result in a "wallflower trap" ecology consisting primarily of antisocial hermits.

Moreover, for non-extreme settings of the choice/refusal parameters, our simulation studies reveal an interesting clustering effect presaged by our analysis of single-tournament N -player IPD/CR games. In Miller's simulation experiments, most ecologies evolve to a set of players whose average payoff scores are near the mutual cooperation payoff. The average payoff scores of players who are not ultimately cooperative end up scattered between the mutual defection and mutual cooperation payoffs. In contrast, in our studies with choice and refusal we typically observe the ultimate formation of two or more distinct tight bands of average payoff scores reflecting the emergence of stable behavioral patterns that are not entirely cooperative. The existence of this payoff band spectrum in turn leads to the emergence of new ecological behaviors. For example, we see "band spiking" in which an ecology abruptly moves from one payoff band to another and then back again, and "band tunneling" in which an ecology that has long resided in one payoff band suddenly traverses to another payoff band and remains there.

The addition of choice and refusal to the Axelrod/Miller IPD game is motivated by an interest in human interactions, particularly in the sexual partner selection process which leads to the spread of AIDS and other sexually transmitted diseases [Hyman and Stanley (1988)]. Not only do rates of sex and new partner acquisition influence the spread of this epidemic, but the structure of the contacts determines who becomes infected. Current models of the spread of AIDS typically assume that behavior is predetermined and responds in exogenously determined ways to changing circumstances. This makes it difficult to

understand the impact of intervention strategies, and to predict the conditions under which particular types of sexual market places will arise.

The choice/refusal ecology studied in this paper is a first step towards a model of AIDS transmission in the context of an endogenously evolving social milieu. This is a long-term development project in which elements of realism will be introduced one step at a time. The current model focuses on a key feature essential for the final ecology: the possibility of choice and refusal in social interactions.

The relation of this paper to previous work on cooperation in IPD games is outlined in section 2. Section 3 reviews the essential features of the basic IPD game set out in Axelrod (1984), and extends this framework to include a choice/refusal mechanism. Section 4 briefly describes some simulation experiments with an IPD game studied in Miller (1989) which illustrate the potential sensitivity of evolutionary IPD game outcomes to the introduction of choice and refusal. An analytical study of single-tournament IPD/CR games is undertaken in sections 5 and 6. These results are used in section 7 to provide a more careful interpretation of various simulation results obtained for evolutionary IPD/CR games. Concluding comments are given in section 8.

2. RELATION TO PREVIOUS WORK

In a series of pathbreaking studies, Axelrod (1981,1984,1988) has explored the initial emergence and viability of cooperative behavior in the absence of either altruism or binding commitments, using the IPD game as a paradigm for social interactions. In each iteration, each player plays one prisoner's dilemma game with each other player in a fixed pool of N players. The only possible choice for each player in each two-player game is either to cooperate or to defect, and both players must choose simultaneously.

As discussed by Axelrod (1984, Chapter 3) and by May (1987), the cooperative Tit-for-Tat strategy is a collectively stable strategy for the IPD game if the number of iterations is either uncertain or infinite and the probability that any two players meeting in a current iteration will meet again in a future iteration is sufficiently high.³ In an IPD game with a known finite number of iterations, however, cooperation is hard to sustain. Mutual defection occurs in the final iteration of the game because no player foresees any future gains to cooperation, and this typically leads by backwards recursion to mutual defection

³The *Tit-for-Tat* strategy is defined as follows: cooperate initially, and thereafter do whatever the other player did in his previous move. A single mutant strategy introduced into a pool of identical native strategies is said to *invade* the native strategy if the newcomer receives a higher payoff from playing against a native strategy than a native strategy receives from playing against another native strategy. A native strategy is said to be *collectively stable* if no mutant strategy can invade it..

in every iteration.

A number of modifications have been proposed which permit the viability of cooperative strategies in IPD games even if the number of iterations is known to be finite. Kreps et al. (1982) establish that mutual cooperation can be sustained in every iteration up to some iteration close to the end of the game if one player assigns positive probability to the possibility that the strategy followed by the other player is Tit-for-Tat. Thompson and Faith (1981) and Hirshleifer (1987), among others, have shown that cooperation can be sustained if players can credibly commit themselves to use retaliatory strategies in response to defections by opposing players. Finally, Hirshleifer and Rasmusen (1989) use the possibility of group ostracism to sustain mutual cooperation in all but the last iteration. Empirical support for the cooperation-inducing effects of group ostracism can be found in a case study by Barner-Barry (1986).

Although such modifications do enhance the viability of cooperative strategies in IPD games, a major difficulty remains. As pointed out by Axelrod and Hamilton (1981), cooperative strategies cannot successfully invade a population of defectors playing the basic IPD game unless the initial frequency of interactions between cooperative strategies is sufficiently large. Consequently, it is difficult for cooperation to emerge spontaneously from noncooperation in the basic IPD framework.

One limitation of many iterated game studies of social interaction which hampers the emergence of cooperation is the implicit assumption that individual players have no control over which opponents they play; see, for example, the models reviewed by Maynard Smith (1982). Players either engage in a round-robin tournament—i.e., each player in each iteration plays one game with each other player in a pre-determined set of players—or games occur through random encounters. In actuality, however, social interactions among organisms are typically characterized by the choice and refusal of partners rather than by a random or deterministic matching mechanism. How do herds form for foraging and protection? How do animals choose mates? How do family and social structures protect cooperative players from noncooperative players? And how do humans choose their friends and sexual partners? The question thus arises whether the long-run viability of cooperation in the IPD game would be enhanced if players were more realistically allowed to choose and refuse their potential game partners.

Conjectures along these lines have been explored by a number of previous researchers. For example, in the context of a Darwinian fitness model, Eshel and Cavalli-Sforza (1982, p. 1333) show that full cooperativeness is the only evolutionarily stable strategy if encounter probabilities are sufficiently biased in favor of meeting an individual using the same strategy. Feldman and Thomas (1987) investigate conditions under which multiple IPD strategies can coexist in a stable equilibrium, assuming the probability that a player

stays in the IPD game depends either on his own current play or on the current play of his opponent. Dugatkin and Wilson (1991) examine the ability of a roving always-defect player to invade a Tit-for-Tat player population that is partitioned into "patches" of different externally-specified size and duration. Kitcher (1992) argues that altruistic play (giving weight to the payoffs of other players) can evolve more readily in contexts where game play is optional and the possibility exists for taking actions whose fitness effects are independent of the actions of others.

The present paper complements and extends this work by allowing players to choose and refuse potential game partners on the basis of continually updated expected payoffs. Player encounters are thus determined by anticipated rewards rather than by an encounter probability biased towards cooperative behavior per se. Moreover, ostracism (end of game play) for noncooperative players occurs endogenously as an increasing number of players individually refuse their game offers.

3. THE BASIC IPD GAME WITH CHOICE AND REFUSAL

In this section we first review the essential features of the basic single-tournament IPD game set out in Axelrod (1984). We then extend this framework to include a choice and refusal mechanism.

3.1 The Basic IPD Game

The prisoner's dilemma (PD) game is a game with two players. Each player has two possible moves, "cooperate" or "defect," and each player must move without knowing the move of the other player. If both players defect, each receives a payoff D . If both cooperate, each receives a payoff C which is strictly greater than D . Finally, if one defects and the other cooperates, the cooperating player receives the lowest possible payoff B and the defecting player receives the highest possible payoff S ; where $B < D < C < S$. For reasons clarified below, the payoffs are also restricted to satisfy $(S + B)/2 < C$.

The dilemma is that, if *both* players defect, both do worse than if both had cooperated; yet there is always an incentive for an individual player to defect. More precisely, the payoffs (C, C) achieved with mutual cooperation are higher than the payoffs (D, D) achieved with mutual defection. Nevertheless, defection is the best response to any move an opponent might make. The best response to defection is to defect, because this avoids the lowest possible payoff B ; and the best response to cooperation is to defect, because this achieves the highest possible payoff S .

The *iterated prisoner's dilemma (IPD) game* is a tournament consisting of the repeated

(iterated) play of a round-robin of PD games at discrete time intervals. In each iteration, each player plays one PD game with each other player in a fixed pool of N players. The only information a player has about another player is the history of payoffs achieved in previous game plays with that player. The restrictions on payoffs guarantee that the players cannot escape their dilemma by taking turns exploiting each other. For any two players, the average (per game) payoffs (C, C) achieved with mutual cooperation over the course of the IPD game are higher than either the average payoffs (D, D) achieved with mutual defection or the average payoffs $((S + B)/2, (S + B)/2)$ achieved with alternating plays of cooperation against defection and defection against cooperation.

3.2 Introduction of Choice and Refusal

The *IPD game with choice and refusal*, henceforth abbreviated by IPD/CR, is an IPD game with the added feature that players can choose and refuse game partners in each iteration. A fixed pool of N players engages in a tournament consisting of indefinitely many iterations. Each iteration, in turn, consists of five stages: (1) a *choice stage* in which each player makes PD game offers to a limited number of potential game partners with high expected payoffs; (2) a *refusal stage* in which each player refuses PD game offers with unacceptably low expected payoffs; (3) a *play stage* in which non-refused PD game offers are played out as PD games; (4) a *cleanup stage* in which payoffs are calculated and recorded for both active and inactive players; and (5) an *update stage* in which expected payoffs are updated for active players on the basis of newly received payoff information.

The logical progression of the IPD/CR game will now be more fully described. At the beginning of the first iteration 1, all players are assumed to assign the same *initial expected payoff* π^0 to each possible play of a game. Then, in each iteration $i = 1, 2, \dots$, the five component stages take the following form:

Choice Stage: Each player determines which other players are tolerable game partners. Given any player n , a player $m \neq n$ is *tolerable* for player n in iteration i if

$$\pi^{i-1}(m|n) \geq \tau, \quad (1)$$

where $\pi^{i-1}(m|n)$ denotes the the expected payoff to n from playing a PD game with player m in iteration i , and τ denotes an exogenously given *minimum tolerance level*. If the number of tolerable players for player n is no greater than an exogenously given *upper choice bound* K , where $1 \leq K \leq N - 1$, then player n makes a PD game offer to each tolerable player. If the number of tolerable players exceeds K , player n makes a PD game offer to the K tolerable players m for whom his expected payoff $\pi^i(m|n)$ is highest. Ties are settled by a random draw.

Refusal Stage: Each player then examines the PD game offers he has received. Any offer coming from an intolerable player is refused, and any offer coming from a tolerable player is accepted. Thus, a player gets to reject odious offers; but he cannot opt out of an offer received from someone he has judged to be tolerable in the Choice Stage.

Play Stage: All non-refused PD game offers are played out as PD games. Even if there are mutual offers between two players, only one PD game is played.

Clean-up Stage: A player receives a *rejection payoff* R for each PD game offer he made in iteration i which was refused, and a PD payoff (depending on the actual play of the game) for each PD game offer he made or received in iteration i which was not refused. An inactive player—i.e., a player who neither made nor accepted offers in iteration i —is assigned a *wallflower payoff* W . Note that a player makes no offers if and only if he judges all other players to be intolerable.

Update Stage: Consider any two players n and m . If n neither made nor accepted a PD game offer from m in the current iteration i , then n 's expected payoff $\pi^{i-1}(m|n)$ for the play of a PD game with m in iteration i is trivially updated to

$$\pi^i(m|n) = \pi^{i-1}(m|n) \quad (2)$$

for play in the next iteration $i+1$. On the other hand, suppose player n either made a PD game offer to m (who subsequently either accepted or rejected it) or accepted a PD game offer from m . In the former case, the payoff to n is either a PD payoff or the rejection payoff R ; in the latter case the payoff to n is a PD payoff. In either case, let this payoff be denoted by U . Then player n 's updated expected payoff for making a PD game offer to player m in the next iteration $i+1$ takes the form of a weighted average over player n 's payoff history with player m ,

$$\pi^i(m|n) = \omega \pi^{i-1}(m|n) + (1 - \omega)U, \quad (3)$$

where the *memory weight* ω controls the relative weighting of distant to recent payoffs.⁴ Note that an increase in ω implies an increase in the weight put on past payoffs relative to current payoffs, which in turn leads to more inertia in the partner selection process.

⁴This mechanism for updating expected payoffs is a special case of "criterion filtering"—i.e., the direct updating of expected return functions on the basis of past return observations. Criterion filtering is an operationally feasible alternative to the indirect updating of criterion functions via Bayes' rule which, given an appropriate specification for the filter weights, can yield a strongly consistent estimate for the true expected return function; see Tesfatsion (1979).

In summary, the choice and refusal mechanism for the IPD/CR game is characterized by the following six parameters:

| | |
|--------------------------|----------|
| Initial Expected Payoff: | π^0 |
| Minimum tolerance level: | τ |
| Upper choice bound: | K |
| Rejection payoff: | R |
| Wallflower payoff: | W |
| Memory weight: | ω |

At the end of the IPD/CR game, the overall success of each player is measured by his average payoff score, calculated as the total *sum* of his payoffs divided by the total *number* of his payoffs.⁵

It is assumed that players in IPD/CR games do not anticipate future conditions under which PD game play with a current game partner could end, implying that they play each PD game with another player as if the total number of PD games to be played with that player is indefinite. Since the choice/refusal mechanism itself is exogenously determined, this implies that players can then be distinguished from one another on the basis of the rules they use to play an indefinite sequence of PD games with an arbitrary opposing player, i.e., on the basis of their *PD rules*. As will be seen in section 5, below, a rich class of PD rules can be represented as finite automata.

Note that the IPD/CR game reduces to the basic Axelrod IPD game in two cases: either (i) $K = N - 1$ and $\tau = 0$; or (ii) $K = N - 1$, $\omega = 1$, and $\tau < \pi^0$. In either case, each player ends up playing one PD game with each other player in each iteration, i.e., each iteration reduces to a round-robin among the N players.

4. IPD WITH CHOICE AND REFUSAL: AN ILLUSTRATION

An intriguing artificial life experiment run with prisoner's dilemma is reported in an evolutionary IPD game study by Miller (1989). A population of thirty finite automata playing an IPD game were allowed to evolve by means of a genetic algorithm that used high payoff scores as its selection principle. To illustrate the potential sensitivity of evolutionary outcomes to the introduction of choice and refusal, this section briefly describes some simulation results obtained for Miller's evolutionary IPD game after the introduction of a

⁵In an IPD/CR game, the players have some degree of control over the number of PD games they play—equivalently, over the number of moves they make—and players not participating in PD games can still receive wallflower and rejection payoffs. Consequently, average payoff per payoff made is used as a measure of overall success rather than total payoff or average payoff per game.

choice/refusal mechanism. A more careful discussion of these results is given in section 7, following a preliminary analytical investigation of single-tournament IPD/CR games.

We first implemented our own version of Miller's experiment, without noise. As in Miller, fifty successive tournaments were conducted. Each tournament was separated from the next by a genetic step in which only the twenty most successful players among the thirty automata constituting the current player set were allowed to reproduce, resulting in a modified player set of thirty automata for the next tournament. The fifty-tournament run thus resulted in an "ecology" consisting of an evolved population of thirty automata.

Each tournament in turn consisted of one hundred and fifty iterations of round robin PD games among thirty automata. The four possible per-game payoffs were $B = 0$ (for cooperating against a defecting player), $D = 1$ (for mutual defection), $C = 3$ (for mutual cooperation), and $S = 5$ (for defecting against a cooperating player). The entire run of fifty tournaments separated by genetic steps was repeated forty times to obtain forty distinct ecologies.

Figure 1 shows, for each of the fifty tournaments, the average payoff score obtained by the thirty automata over all forty ecologies. These results generally conform to the results obtained by Miller. Note the initial "dip" in the average payoff score. This dip reflects the exploitation of some players by other opportunistic players, until the implacable forces of evolution eliminate the chumps. The upward progress seen thereafter is the result of cooperative but relatively unexploitable players—of which Tit-for-Tat is a sterling example—beating out the opportunists now that the opportunists' victims have died out.

—Figure 1 About Here—

We then incorporated our choice and refusal mechanism into Miller's experimental set up. In Figure 2 we see the analogous diagram to Figure 1 for simulations of the resulting modified Miller IPD game in which, in each iteration: (i) players have initially rosy payoff expectations prior to play ($\pi^0 = 3$); (ii) each player makes a PD game offer to at most one other preferred player ($K = 1$); (iii) received offers with expected payoffs greater than or equal to $\tau = 1.6$ are played out, and received offers with expected payoffs less than this value are refused; (iv) a player whose offer is refused receives a rejection payoff $R = 1$; (v) a player neither making nor accepting any offers receives a wallflower payoff $W = 1.6$; and (vi) players have reasonable but not excellent memories of their past payoff outcomes ($\omega = 0.7$).

—Figure 2 About Here—

It is interesting to note in Figure 2 that the initial dip in average payoff score seen

in Miller's experimental data has vanished. The Miller dip occurred because various nice players were fatally exploited by predatory players, which lowered the average payoff score. In contrast, in an IPD/CR game it sometimes happens that players, exploitable in the basic IPD game, can protect themselves via refusal of predatory players rather than dying, and hence remain in the ecology indefinitely. One type of highly exploitable player is a player whose PD rule closely approximates "always cooperate." Such players can attain high payoff scores when playing one another, and hence can dominate the genetic step reproduction process when protected from predatory players by a refusal mechanism. As a result, unlike Miller, an initial dip in average payoff score need not occur in an evolutionary IPD/CR game.

While replicating Miller's work, we became curious about the behavior of individual ecology payoffs. To this end, we plotted the average tournament payoffs for each of the forty ecologies in our replication of Miller's experiment and obtained Figure 3. As one can see, nine of the forty Miller ecologies did not end in the cluster of essentially cooperative ecologies, hence the standard deviation of average tournament payoffs across ecologies is rather large.

—Figure 3 About Here—

Figure 4 gives the average tournament payoffs for each of the forty ecologies in the choice/refusal experiment depicted in Figure 2. These payoffs display an interesting behavior when compared to the noise-like outlier payoffs obtained for Miller's experiment in Figure 3. Notice that *two* distinct payoff bands have now emerged, one close to 3 and the other just below 2.8. The ecologies corresponding to the payoff band near 3 ultimately consist of a single large group within which essentially random partner selection takes place. In contrast, the ecologies corresponding to the payoff band just below 2.8 ultimately consist of several small player groups whose members only choose to play one another.

—Figure 4 About Here—

We then raised both the minimum tolerance level and the wallflower payoff from 1.6 to 2.5. As depicted in Figure 5, this change results in the gain of a third tight payoff band. The ecologies associated with this band ultimately consist of many players who neither make nor accept game offers, choosing instead to survive on the wallflower payoff.

—Figure 5 About Here—

As these examples indicate, the addition of choice and refusal to Miller's original IPD

experiment induces the formation of an interesting new ecological feature—the formation of multiple distinct payoff bands. Somewhat at odds with our initial expectations, choice and refusal do not merely speed up the emergence of cooperation; they also allow, in some parameter regimes, the appearance of stable player interaction patterns that are not entirely cooperative.

5. SINGLE-TOURNAMENT 2-PLAYER IPD/CR GAMES

The simulation results presented in the previous section for an evolutionary IPD/CR game are suggestive but preliminary. This section undertakes a more careful analysis of the role of choice and refusal in the simpler context of a single-tournament two-player IPD/CR game. For such games, the upper choice bound $K \in \{1, \dots, N - 1\}$ is forced to equal $1 = N - 1$. The following section 6 takes up the more general case of a single-tournament N -player IPD game with $K = 1$. The results of sections 5 and 6 are then used in section 7 to provide a more careful interpretation of the simulation results obtained for evolutionary IPD/CR games.

Since even a single-tournament two-player IPD/CR game can be very complicated to analyze, we first describe six example player types. We then analyze single-tournament two-player IPD games using various combinations of these player types.

5.1 Example Player Types

As discussed in section 3.2, players in IPD/CR games can be identified with the PD rules they use in playing an indefinite sequence of PD games with an arbitrary opposing player. We will study the pairwise interactions of six types of players (PD rules): (1) Always Defect (AllD); (2) Ripoff-Artist (Rip); (3) Gentle Ripoff (GRip); (4) Tit-for-Tat (TFT); (5) Tit-for-Two-Tats (TFTT); and (6) Always Cooperate (AllC).

These six players roughly span the range from uncooperative to cooperative behavior. With the exception of GRip, all of these players have previously been used in studies of IPD games. GRip was invented in order to have a relatively subtle opportunistic player who initially appears cooperative, but who repeatedly sneaks in defections after an opposing player has built up a rosy payoff expectation.

Finite automaton representations for these six players are depicted in Figure 6, ordered by their complexity. All six players make an opening move, either cooperate (c) or defect (d), and then enter state 1; this opening move is indicated next to the arrow entering state 1. Thus, GRip, TFT, TFTT, and AllC initially cooperate, while AllD and Rip initially defect.

Once one of the six players, say n , has arrived at a current state, his next move is conditioned on the previous move of the opposing player. This move sequence then determines a transition to a new state. A transition to a new state is indicated by an arrow, and the move sequence(s) which result in this transition are indicated beside the arrow in a move-slashmark-move format. The previous move of the opposing player appears to the left of the slashmark and the next move of player n appears to the right of the slashmark; i.e., moves are time-sequenced from left to right.

In particular, TFT begins by cooperating and then mimics whatever move his opponent made in the previous PD game. Consequently, once in the initial state 1, two move sequences are possible for TFT, either $c|c$ or $d|d$, and each move sequence results in a transition back to state 1; see Figure 6(a). The finite automaton representations for AllC and AllD are similarly straightforward; see Figures 6(b) and 6(c).

TFTT begins by cooperating, subsequently defects only if his opponent defects twice in a row, and quickly reverts back to cooperation as soon as his opponent cooperates. TFTT's behavior is depicted in Figure 6(d) by a finite automaton with two states. Initially, TFTT cooperates and enters state 1; and TFTT continues to cooperate and return to state 1 as long as his opponent cooperates. The first time his opponent defects, if ever, TFTT cooperates but enters a new state 2. If his opponent then defects again, TFTT defects and returns to state 2; but if his opponent cooperates, TFTT cooperates and returns to state 1. Thus, TFTT is less provokable than TFT, requiring two successive defections before retaliating with a defection; but TFTT is equally quick to forgive as soon as the other player is nice.

Rip is more complex than TFT or TFTT, requiring three states for its finite automaton representation in Figure 6(e). Rip evolved in an evolutionary IPD game which included TFT, TFTT, AllC, and AllD in the initial player pool. Consequently, it is not surprising that Rip takes advantage of TFT, TFTT, and AllC, while protecting itself from AllD. More precisely, Rip first tests the mettle of an opponent with a defection and enters state 1. If the opponent ever defects, Rip cooperates and enters the "TFT" state 3, resulting in TFT play for all further moves against this player. Thus, after first defecting and then cooperating, Rip attains mutual cooperation with TFT, with other Rips, and with GRip, and mutual defection with AllD. However, against opponents such as AllC and TFTT which never retaliate after a single defection, Rip alternately cooperates and defects, moving from state 1 to state 2 and back again. Rip thus takes advantage of the low provocability of AllC and TFTT—full advantage in the case of TFTT.

GRip is more complex yet, requiring four states for its finite automaton representation

in Figure 6(f). GRip initially cooperates, but eventually GRip tries out a defection against “nice” players. Against TFFT and AllC, GRip cooperates for his first four moves; but then GRip enters a 3-move cycle (d, c, c) while TFFT and AllC continue to cooperate. Thus, like Rip, GRip rips off TFFT and AllC, although not as frequently. Against another GRip, each GRip immediately enters into a 5-move cycle (c, c, c, c, d) . Against TFT and Rip, GRip ultimately enters a 5-move cycle (d, c, c, c, c) against which TFT and Rip play (c, d, c, c, c) .

5.2 Analytical Results for Two-Player IPD/CR games

An analytical study was undertaken of single-tournament two-player IPD/CR games between all possible pairs of players described in section 5.1. Table 1 presents the key results from these studies, and Figures 7 and 8 highlight some of these results.

Reading the columns of Table 1 from left to right, we have: (1) the player pairs; (2) the players’ (updated) expected payoffs at the end of iteration i ; (3) sufficient conditions for PD game play to come to a halt in finitely many iterations; (4) the maximum (possibly infinite) number i^* of iterations during which PD games will be played (i.e., the last iteration before either a refusal of a PD game offer occurs, or no PD game offers are made); (5) the player(s) who would refuse any PD game offer made by the opposing player in iterations subsequent to i^* , if any; and (6) the average payoff score of each player in each iteration through i^* . The quantities in columns two through four and in column six are determined as functions of the PD game payoffs and the parameters describing the choice/refusal mechanism.

The results in Table 1 are presented in normalized form. More precisely, the PD game payoffs $\{B, D, C, S\}$ and the initial expected payoff π^0 are normalized by subtracting B and dividing by $D - B$. The resulting normalized values are indicated below as hatted quantities:

$$\hat{B} = 0; \hat{D} = 1; \hat{C} = \frac{C - B}{D - B}; \hat{S} = \frac{S - B}{D - B}; \hat{\pi}^0 = \frac{\pi^0 - B}{D - B}. \quad (4)$$

Thus, after normalization, the lowest possible PD game payoff B becomes 0, the mutual defection payoff D becomes 1, and the non-oscillation condition $(S + B)/2 < C$ becomes $\hat{S}/2 < \hat{C}$. This normalization decreases the number of parameters by two, and demonstrates that the signs of the payoffs are not important for the IPD/CR game. However, since it is not possible to normalize \hat{C} and \hat{S} any further, changing the relative distance between the payoffs can have an effect. For expositional simplicity, the hats are dropped from the normalized payoffs and initial expected payoff in Table 1 and throughout the remainder of this paper.

In obtaining the results presented in Table 1, we assumed that the initial expected payoff π^0 was not unrealistically high; but that it was high enough to guarantee that at

least one game would always be played. Specifically, we assumed that $\tau < \pi^0 \leq S$. We also assumed that the minimum tolerance level τ was not unrealistically high or low, i.e., $0 \leq \tau \leq C$, and that the memory weight ω satisfied $0 < \omega < 1$. Finally, we dropped the conditional player indexing on expected payoffs, since it is generally obvious who is playing whom when there are only two players; subscripts are used for clarification when necessary.

Before we discuss the specific results in Table 1, we demonstrate how they were obtained by describing a couple of the calculations.

AllD versus AllD:

Suppose two AllD players are playing a two-person IPD/CR game. Since $\tau < \pi^0$ by assumption, both players make PD game offers to each other in the first iteration. In the subsequent play of the PD game, both players defect and receive a payoff of $D = 1$. Their expected payoffs for the next iteration then take the form $\pi^1 = \omega\pi^0 + (1 - \omega)$.

Suppose the players play a PD game in i successive iterations, $i \geq 1$. They will then each have an expected payoff of

$$\begin{aligned}\pi^i &= \omega\pi^{i-1} + (1 - \omega) \\ &= \omega^i\pi^0 + (1 - \omega^i).\end{aligned}\tag{5}$$

PD game play stops at the end of iteration i if and only if the expected payoff π^i drops below τ , i.e., if and only if

$$\omega^i\pi^0 + (1 - \omega^i) < \tau.\tag{6}$$

If $\tau \leq 1$, this inequality is not satisfied for any i , implying that $i^* = \infty$. Suppose $1 < \tau$. Then $1 < \pi^0$ must hold as well. Solving the inequality for i then yields

$$\ln\left(\frac{\tau - 1}{\pi^0 - 1}\right) \frac{1}{\ln(\omega)} < i.\tag{7}$$

Letting i^* denote the smallest integer value of i for which this inequality is true, i^* is the last iteration during which a PD game is played between AllD and AllD. For example, if $\tau = 2$, $\pi^0 = 3$, and $\omega < 0.5$, then $i^* = 1$, implying that exactly one PD game is played.

If $i^* < \infty$, the expected payoff which AllD associates with another play of a PD game with AllD remains frozen at $\pi^{i^*} < \tau$ for all $i \geq i^*$, and no more PD game offers are made. Each player then simply receives the wallflower payoff in all subsequent iterations.

Note that an optimistically high initial expected payoff π^0 will result in many PD game plays (large i^*) unless the memory weight ω is small. That is, the specification of a high initial expected payoff tends to encourage repeated play.

Rip versus Rip:

Both Rips defect against each other on the first iteration, which results in an updated expected payoff $\pi^1 = \omega\pi^0 + (1-\omega)$ for each player. If $\pi^1 < \tau$, they will stop playing without rejection. On the other hand, if $\pi^1 \geq \tau$, they will continue playing forever, because they will cooperate with each other on all PD games except the first. Thus, play between two Rip players will stop after a finite number of iterations if and only if the first PD game results in an updated expected payoff $\pi^1 < \tau$, in which case $i^* = 1$.

Rip versus AllC or TFTT:

AllC and TFTT always cooperate with Rip, while Rip alternately defects and cooperates. Without loss of generality, consider TFTT. TFTT receives C in even iterations and 0 in odd iterations, while Rip receives C in even iterations and S in odd iterations. Given any payoffs E and F , the expected payoffs when receiving alternating payoffs of E in even iterations and F in odd iterations take the form

$$\pi^{2i} = \omega\pi^{2i-1} + (1-\omega)E; \quad \pi^{2i+1} = \omega\pi^{2i} + (1-\omega)F. \quad (8)$$

Solving these coupled difference equations gives

$$\pi^{2i} = \omega^{2i}\pi^0 + \frac{1-\omega^{2i}}{1+\omega}(E+\omega F) \quad (9)$$

and

$$\pi^{2i+1} = \omega^{2i+1}\pi^0 + \omega\frac{1-\omega^{2i}}{1+\omega}(E+\omega F) + (1-\omega)F. \quad (10)$$

If $(C\omega)/(1+\omega) \geq \tau$, then TFTT and Rip never refuse each other because their expected payoffs remain greater than τ in each iteration i . Consequently, $i^* = \infty$. If $(C\omega)/(1+\omega) < \tau$ and $\pi^0 \leq C/(1+\omega)$, the expected payoff of TFTT falls below τ after the first PD game play with Rip: $\pi^1 = \omega\pi^0 \leq (C\omega)/(1+\omega) < \tau$. Thus $i^* = 1$.

Finally, suppose $(C\omega)/(1+\omega) < \tau$ and $C/(1+\omega) < \pi^0$. In this case the expected payoff of TFTT falls below τ in the first odd iteration $2i+1$ that satisfies

$$\ln\left(\frac{\tau(1+\omega)-C\omega}{\pi^0(1+\omega)-C}\right)\frac{1}{\ln(\omega)} < 2i+1. \quad (11)$$

Using the difference equation (9), together with the assumption that $\tau \leq C$, it can be shown that the expected payoff of TFTT in all previous even iterations is still greater than τ . Moreover, for all previous iterations through this odd iteration, the expected payoff of Rip exceeds τ , implying that Rip has never refused TFTT. Consequently, this odd iteration gives i^* , the iteration of last PD game play between Rip and TFTT.

Once TFFT begins to refuse Rip, Rip receives the rejection payoff R . The expected payoff for Rip is updated according to the equation

$$\pi_R^i = \omega^{i-i^*} \pi_R^{i^*} + (1 - \omega^{i-i^*})R. \quad (12)$$

If $R \geq \tau$, π_R^i never falls below τ , and Rip will receive the rejection payoff in each subsequent iteration. If $R < \tau$, Rip will stop choosing to play TFFT when π_R^i declines below τ . Thereafter, both players will receive the wallflower payoff.

Play Stoppage Conditions:

Examining the third column in Table 1, we see that the minimum tolerance and memory parameters τ and ω play a crucial role in determining whether or not play eventually stops for our six player types. Figures 7(a) through 7(c) show the regions of the $\omega - \tau$ plane in which play stoppage occurs for AllD, Rip and GRip, respectively, in IPD/CR game play with other player types. These regions are determined for the particular parameter specifications $\pi^0 = C = 3$ and $S = 5$.

—Figure 7 About Here—

Note that some results are determined by τ alone, regardless of all other parameter and payoff values:

- $\tau > 1$: Play stops for all two-player IPD/CR games involving AllD;
- $\tau \leq 1$: Play never stops for AllD v. AllD, Rip v. Rip, or GRip v. GRip;
- $\tau > 0$: AllC eventually refuses AllD.

When memory fades quickly, i.e., when ω is small, defection can trigger refusal by fairly tolerant players despite a long history of mutual cooperation or a high initial expected payoff. For example, as seen in Figure 7(c), Rip eventually refuses GRip for small enough ω (if GRip has not already refused Rip) even for minimum tolerance levels τ less than the mutual defection payoff $D = 1$. Note that the line demarking when Rip refuses GRip is independent of π^0 .

If defections occur early and then stop, play behavior depends critically on simple relationships among the minimum tolerance level τ , the memory weight ω , and the initial expected payoff π^0 . A high initial expected payoff π^0 tends to decrease a player's sensitivity to early defections, whereas a low memory weight ω tends to increase his sensitivity to early defections by downgrading the importance of the initial expected payoff in all subsequent updated expected payoffs. Thus, for example, a defection on the first iteration will result

in a refusal in the next iteration by any opposing player who: (i) cooperated in the first iteration, if (ω, τ) lies above the line $\tau = \omega\pi^0$ in the $\omega - \tau$ plane; or (ii) defected in the first iteration, if (ω, τ) lies above the line $\tau = \omega[\pi^0 - 1] + 1$ in the $\omega - \tau$ plane. Consequently, the lower the initial expected payoff π^0 , the larger is the (admissible) region in the $\omega - \tau$ plane over which Rip, TFT, and GRip will play only one PD game with any player who defects in the first iteration.

The fourth column of Table 1 highlights another interesting characteristic of two-player IPD/CR games: the maximum (possibly infinite) number of iterations during which PD game play will take place, denoted by i^* . Figure 8 depicts regions of the $(1 - \omega) - \tau$ plane where the iteration number i^* takes on a constant value, conditional on $\pi^0 = C = 3$ and $S = 5$, for various player pairs. In each case, the region where i^* takes on its smallest value is the largest region, and the regions get ever smaller as i^* increases. In general, the regions where i^* takes on large but finite values are very small indeed; that is, for most player pairs, intolerance (play stoppage) either occurs after only a few iterations, or it does not occur at all.

—Figure 8 About Here—

Recall that, given $K = N - 1$, the IPD/CR game reduces to the IPD game if either $\tau = 0$ or $\omega = 1$ and $\tau < \pi^0$. The analytical results of this section indicate that, near these boundaries in parameter space, the IPD/CR game with $K = N - 1$ will behave much like the IPD game: either players will play each other forever without refusal, or they will play many PD games before refusal. However, as τ increases and/or ω decreases, refusal becomes more likely, and we move away from the IPD game.

Generic Results:

Certain parameter regimes can be directly associated with particular types of play behavior. In section 7, below, these parameter regimes will be used to help predict and explain the ultimate outcomes in our evolutionary (multi-tournament) IPD/CR game simulations. The most important parameter regime is when any defection against a cooperation results in future refusal by the cooperating player. The next most important regime is when any defection results in immediate play stoppage.

If the first defection in a two-player tournament occurs on iteration i , with one player cooperating and the other defecting, the cooperating player will refuse further play if

$$\tau > \omega^i \pi^0 + \omega(1 - \omega^{i-1})C. \quad (13)$$

Examining this inequality, we see that if $\tau > \max\{\omega\pi^0, \omega C\}$, then any defection at any time

against a cooperation results in immediate refusal to play. For example, in the evolutionary simulations illustrated in section 4 and detailed in section 7, below, we set $\pi^0 = C = 3$ and $\omega = 0.7$; and for these fixed parameter values, defections against cooperation are not tolerated for any minimum tolerance level τ greater than 2.1.

If instead the first defection is mutual, both players will stop choosing each other if

$$\tau > \omega^i \pi^0 + \omega(1 - \omega^{i-1})C + 1 - \omega. \quad (14)$$

Comparing (14) with (13), we see that *any* initial defection results in refusal to play if $\tau > (1 - \omega) + \max\{\omega\pi^0, \omega C\}$. In particular, for $\pi^0 = C = 3$ and $\omega = 0.7$, any initial defection results in immediate play stoppage if $\tau > 2.4$.

6. SINGLE-TOURNAMENT N-PLAYER IPD/CR GAMES

The analysis in section 5 describes the characteristics of various representative two-player IPD/CR games with $K = 1$. This two-player analysis can be extended in a straightforward way to analyze any N -player IPD/CR game in which $K = N - 1$. Given this specification of K , each player chooses to play a PD game with every tolerable player in each iteration, implying that the N -player IPD/CR game decomposes into a collection of $N(N - 1)/2$ two-player IPD/CR games of the type analyzed in section 5. Consequently, setting $K = N - 1$ trivializes the role of choice in the determination of PD game partners; it is refusal which is critical.

On the other hand, when the number N of players is greater than 2 and K is less than $N - 1$, players will not necessarily choose to play all tolerable players during each iteration. Choice then becomes more important, and one would expect to see choice and refusal working together in a more subtle fashion. In fact, as we shall see, in many interactions choice is much more important than refusal.

The outcomes of N -player IPD/CR games can be characterized in a number of different ways, including the average payoff scores both for individual players and for the entire ecology, the fraction of game plays of each type, and the formation of long-run player interaction patterns. For example, who ends up playing whom when a steady state is reached, if ever? To what extent is the ultimate steady state ecology characterized by a hierarchy of social cliques, ranging from high-status player groups which everyone else would like to join, all the way down to ostracized individual players whom all other players avoid?

In order to see how play in a multiple-player tournament proceeds, and how long-run player networks form, consider an illustrative 5-player IPD/CR game with three TFTT players and two Rip players played for infinitely many iterations. As in section 5.2, let

payoffs be normalized so that $B = 0$ and $D = 1$. For simplicity, suppose that each player can choose at most one potential game partner in each iteration ($K = 1$), that the initial expected payoff, common to all players, is $\pi_0 = C$, that the wallflower payoff W coincides with the minimum tolerance level τ , and that the rejection payoff R does not exceed 1. Finally, assume that when a player needs to choose among partners with the same expected payoff, the choice is made by a random draw. Since the initial expected payoff π^0 is the same for all potential partners, this implies in particular that each player chooses each of the other four players as a potential first partner with probability $1/4$.

Recall from section 5 that two TFTTs always cooperate, and that a TFTT receives a payoff sequence $(0, C, 0, C, \dots)$ in an IPD game with a Rip. Moreover, a Rip receives a payoff sequence (S, C, S, C, \dots) in an IPD game with a TFTT, and a payoff sequence $(1, C, C, \dots)$ in an IPD game with another Rip. Observe, also, that $\pi^0 = C$ implies $C\omega/(1+\omega) < \omega\pi^0 + (1-\omega)$. Consequently, there are three possible parameter regimes for this 5-player IPD/CR game. If $\tau \leq C\omega/(1+\omega)$, all players remain tolerable to all other players. If $C\omega/(1+\omega) < \tau \leq \omega\pi^0 + (1-\omega)$, Rips are eventually refused by TFTTs, but remain tolerable to each other forever. And if $\tau > \omega\pi^0 + (1-\omega)$, Rips are refused by TFTTs after only one PD game, and they also find each other mutually intolerable after only one PD game. What happens in each of these three parameter regimes will now be described in more detail.

Case I: $\tau \leq C\omega/(1+\omega)$

In this case no player ever refuses to play another player in repeated PD game plays. Play eventually stabilizes in one of two possible player interaction patterns. In both cases, each TFTT chooses randomly in each iteration between the other two TFTTs, and each Rip chooses to play one TFTT repeatedly. But in one case the two Rips both repeatedly play the *same* TFTT player, while in the other case they each repeatedly play *different* TFTT players. Which network forms is determined by the particular realizations of the random draws used to break expected payoff ties in the first few iterations.

We will now look in greater detail at the sequences of plays. First note that the probability that the two Rips ever play each other is $7/16$, the probability that they play each other in the first iteration. This occurs because both Rips defect if they play each other in the first iteration, implying that their expected payoffs from playing each other decrease below C . Since each Rip always has an expected payoff of at least C from any TFTT, each Rip then chooses a TFTT in all future iterations. On the other hand, if the two Rips do not play each other in the first iteration, each Rip plays a PD game with a TFTT and receives a payoff of S . The expected payoff for playing additional PD games

with this TFFT is then larger than C , and the Rip prefers this TFFT to all other players, including the other Rip. In any subsequent PD game with this TFFT, it receives either C or S , and its expected payoff from this TFFT remains above C . From then on it always prefers the TFFT that it played in the first iteration to the other Rip; hence, while it may eventually learn to prefer a different TFFT to the first one it played, it always chooses a TFFT. Since neither Rip chooses the other Rip after the first iteration, they never play each other unless they do so in the first iteration.

The expected payoff that a TFFT has from another TFFT stays at C . The expected payoff that a TFFT has from a Rip is also C until the TFFT first plays that Rip, after which it is less than C no matter how many times the TFFT plays that Rip again. A TFFT therefore chooses randomly among all players until he encounters a Rip (either because he makes a PD game offer to a Rip or because he receives a PD game offer from a Rip). If enough games are played, a TFFT eventually plays both Rips at least once. Once a TFFT plays a Rip, it never chooses that Rip again; but the Rip may choose the TFFT on future plays.

Rips may oscillate among different TFFT game partners for several iterations, but eventually each Rip settles into playing only one particular TFFT. Which TFFT each Rip ends up playing is a matter of chance.

For example, suppose that Rip #1 by chance plays TFFT #1 in the first iteration and gets S . He then chooses TFFT #1 again in iteration 2, and gets C . If TFFT #2 by chance chooses Rip #1 in iteration 2, Rip #1 gets S from TFFT #2, and so chooses TFFT #2 in iteration 3. After getting C from TFFT #2 in iteration 3, Rip #1 is indifferent between TFFT #1 and TFFT #2, but prefers either of these TFFTs to the remaining TFFT #3. If TFFT #3 by chance chooses Rip #1 in iteration 3, Rip #1 then prefers TFFT #3 in iteration 4 and subsequently chooses randomly among all three TFFTs in iteration 5. Since Rip #1 receives S from the TFFT it chooses in iteration 5, and the other TFFTs never choose Rip #1 again, Rip #1 chooses this iteration 5 TFFT in all future iterations. This particular sequence of events is depicted in Figure 9.

—Figure 9 About Here—

Note that each Rip in the long-run player interaction pattern has a higher average payoff than any TFFT. The average payoff of a Rip eventually approaches $(S + C)/2 > C$, as a result of the successful long-run parasitic relationship established with a TFFT. On the other hand, the average payoff for any TFFT who is playing only other TFFTs in the final network approaches C from below as the number of iterations increases; the average payoff for TFFTs playing one Rip are lower than this; and TFFTs playing two Rips do

the worst.

In summary, in case I the low minimum tolerance level allows each opportunistic Rip player to find and exploit a nice TFTT player while avoiding the other Rip. Moreover, not all TFTT players achieve the same average payoff: at least one TFTT player will manage by chance to avoid Rip parasites, and at least one will not.

Case II: $C\omega/(1+\omega) < \tau \leq \omega\pi^0 + (1-\omega)$

In this case a TFTT eventually refuses a Rip in repeated PD game plays, but Rips remain tolerable to each other. The players thus eventually break into two groups, one consisting of the three TFTTs and the other consisting of the two Rips. Average payoffs approach the mutual cooperation payoff C for all players.

How long the final network takes to form depends on the magnitude of τ . If τ is close to the lower bound $C\omega/(1+\omega)$, then the pattern from case I holds until the payoffs that TFTTs expect from Rips all drop below τ . Once this occurs, Rips begin receiving rejection payoffs R from TFTTs. Eventually the Rips stop choosing the TFTTs and settle on each other. On the other hand, if τ is large enough (see section 5), TFTTs refuse Rips after only one PD game and the long-run player interaction pattern gels after only a few iterations. The reader is referred to Stanley et al. (1992) for further details.

Note that the TFTTs primarily use refusal to protect themselves from Rips and use choice to rotate among each other. Rips are able to take advantage of TFTTs only for a limited number of iterations. Ultimately, if the tournament lasts long enough, each Rip is refused by all TFTTs, and the Rips end up choosing to play only with each other.

Case III: $\omega\pi^0 + (1-\omega) < \tau$

In this case, recalling the analysis of two-player IPD/CR games in section 5.2: (i) a Rip never refuses a TFTT in repeated PD game plays; (ii) a TFTT refuses a Rip after only one PD game play (since $\pi^1 = \omega\pi^0 < \tau$); and (iii) a Rip refuses another Rip after only one PD game play. The final player interaction pattern is the group of TFTTs and two ostracized Rip players. Average payoffs approach the mutual cooperation payoff C for the TFTTs and the wallflower payoff W for the Rips. Choice plays almost no role, and refusal works so strongly that Rips never discover that they could receive C from repeated plays with each other in all but the first PD game play.

The role of π^0 .

Suppose the initial expected payoff π^0 in the 5-player IPD/CR game described above

is less than C . Then, once a TFTT choses to play another TFTT, he never chooses to play a Rip in any future iteration. Lowering the initial expected payoff thus cools the system in the sense that exploration is reduced. Rips behave the same as before.

If the initial expected payoff π^0 is greater than C , the long-run player interaction patterns are unchanged. However, each TFTT chooses to play each other player at least once by iteration 5, and plays more PD games with Rip. Moreover, the region in the $\omega - \tau$ plane where Rips refuse each other is smaller than before. Thus, as one would anticipate, increasing the initial expected payoff results in increased exploration.

7. EVOLUTIONARY IPD/CR GAMES

This section describes the software used to run the evolutionary IPD/CR game simulations briefly reported on in section 4, and provides a more careful discussion of the simulation findings.

7.1 Software and Hardware

The software used to replicate the experimental results of Miller (1989), and to modify Miller's framework to include a choice and refusal mechanism, was developed in Turbo Pascal 5.5. We used the random number generator included with the Pascal compiler. The actual runs were done on a number of machines, all 80386-based PC-compatibles with 80387 coprocessors.

Our algorithm is initialized with thirty finite automata of the sort described and illustrated in section 5.1. Each automaton possesses sixteen states and a uniformly distributed random assignment of state transitions, indicated by arrows, together with an initial move and allowable move sequences appearing as arrow labels. Each automaton has access to the values for the parameters $(\pi^0, \tau, K, R, W, \omega)$ characterizing the choice/refusal mechanism, and has a memory comprising: (i) its play history with each other automaton; and (ii) its current expected payoff for a PD game play with each other automaton.

The thirty automata take part in an evolutionary IPD/CR game consisting of fifty successive single-tournament IPD/CR games separated by genetic steps. Each single-tournament IPD/CR game (or "tournament" for short) consists of one hundred and fifty iterations, where each iteration constitutes one pass through the five-stage choice and refusal mechanism described in section 3.2. In the genetic step at the end of each tournament, the current population of thirty automata is replaced via reproduction by a new population of thirty automata. Each run of the evolutionary IPD/CR game thus results in a distinct "ecology" consisting of an evolved population of thirty automata.

The genetic step separating successive tournaments proceeds as follows. First, the automata are sorted by the average (per payoff) payoff score they achieved during the previous tournament. Second, as in Miller (1989), the ten automata scoring the worst are discarded and the twenty automata scoring the best are retained. Finally, the ten emptied slots are filled by “sexual reproduction” of the best twenty automata, with “mutation” of the resulting offspring. The resulting thirty finite automata then constitute the genetically altered population of players for the next tournament.

The “sexual reproduction” is accomplished as follows: First, take the initial move (one bit), 32 arrows (each arrow represented by four bits to describe the “next state”), and 32 arrow labels (one bit each) of each automata and turn them into a bit string 161 bits long.⁶ Second, select two automata to be parents, where each automaton’s probability of selection is directly proportional to the average payoff score it received in the preceding tournament. Third, generate a random variable q distributed uniformly over the discrete range $1, 2, \dots, 161$. Fourth, exchange the bits in positions q through 161 of the parental bit strings to obtain the bit strings for two offspring. Finally, repeat the second and third stages for five pairs of parents, thus obtaining five pairs of offspring.

“Mutation” simply consists of flipping the value of each of the bits of each offspring’s bit string with probability five in one thousand. Once the final bit string for any offspring is obtained, reverse the original conversion process from automaton to bit string to obtain the finite automaton representation for the offspring.

7.2 Simulation Results

Based on the analysis in section 5.2, we decided to focus on the role of the minimum tolerance level τ . In all reported experiments, below, we use the following fixed PD game payoff specifications,

$$B = 0; D = 1; C = 3; S = 5, \quad (21)$$

and we set

$$\pi^0 = 3; K = 1; R = 1; \omega = 0.7. \quad (22)$$

The wallflower payoff W is always set equal to τ . Forty runs are carried out for each tested τ value to allow for stochastic variability.

⁶A bit can take on only two values, 0 or 1. A string of four bits can therefore represent $2^4 = 16$ different states. The arrow labels require only one bit, which encodes the current move to be made in response to a previous move of an opponent, because the information about the previous move of an opponent is coded by position in the string. The finite automaton and bit string representations used in this paper differ from the representations used in Miller (1989).

Figures 10 through 13 display some of the simulation results obtained for a fifty-tournament evolutionary IPD/CR game as the minimum tolerance level τ was set at 0.5, 1.5, 1.6, and 2.4, respectively. More precisely, each figure displays the following information for the indicated τ value: (a) individual ecology payoff trajectories obtained for forty different runs of the fifty-tournament game; (b) the ecology payoffs obtained for each of the fifty tournaments, averaged across the forty runs; and (c) the fraction of various types of play behavior (mutual cooperation CC , successful defection CD , mutual defection DD , and rejection RJ) exhibited in each of the fifty tournaments, averaged across the forty runs. Figure 14 highlights characteristics of the ecologies associated with the lowest payoff band appearing in Figure 13(a).

As will be clarified in the following discussion, Figures 10 through 14 illustrate many of the interesting results of our simulation study. A more detailed discussion of these results can be found in Stanley et al. (1992). A complete data set, together with the PC-compatible software needed to view it, is available; please contact the second author for information.

—Figure 10 About Here—

Consider, first, the payoff band patterns exhibited by the payoff trajectories for individual ecologies in Figures 10(a) through 13(a). Given $\tau = 0.5$ [Figure 10(a)], we see many of these trajectories clustered into three payoff bands at heights of roughly 3.0, 2.8, and 2.3, as well as a fairly large number of trajectories in no apparent band. This is typical of the runs with low τ values. Band formation seems to be primarily a choice-moderated phenomenon. Payoffs with no apparent band association become more common with decreases in τ , hence with decreases in rejections; but payoff bands still form even at $\tau = 0$ where all rejections cease. Notice [Figure 10(c)] that the fraction of rejections RJ is effectively zero for $\tau = 0.5$.

As τ increases to 1.5 and to 1.6 [Figures 11(a) and 12(a)], two of the three payoff bands persist; but the lowest payoff band at level 2.3 is only faintly apparent at best. Fewer ecology payoffs lie outside of a payoff band, and those that do lie outside appear to be either on their way into or out of a band. As τ further increases to 2.4 [Figure 13(a)], the payoff band at 2.3 disappears altogether, the payoff band at 2.8 becomes extremely sparse, and a new payoff band arises in the vicinity of 2.1. The ecologies with payoffs lying in the latter “wallflower isolation” band consist of players that engage in mutual defection until they are rejected by all other players, and that thereafter collect the wallflower payoff $W = 2.4$. [Recall that the wallflower payoff W is always set equal to the minimum tolerance level τ .] A moment’s reflection suggests that such play behavior is perfectly reasonable,

given the high value for the wallflower payoff. Characteristics of these wallflower isolation ecologies are separately displayed in Figure 14.

Preliminary analysis of the genetic diversity of ecologies in the terminal (fiftieth) tournament has been undertaken in an attempt to obtain a better understanding of this payoff band phenomenon. As noted above, payoff bands occur at levels 2.3 and 2.8 for relatively low τ values, and the terminal ecologies associated with each of these payoff bands appear to split into two genetically distinct groups. One group has a relatively low genetic diversity; that is, each ecology in this group is dominated by players with identical or closely related finite automaton structures, leading to highly synchronized play behavior. The stable average payoff behavior exhibited by this group is intuitively understandable; for a synchronized pattern of play can result in fixed fractions of distinct play behaviors, and hence in a roughly constant average payoff. The other group of ecologies has a relatively high genetic diversity, and the stable average payoff behavior exhibited by this group is harder to understand. The terminal ecologies associated with the mutually cooperative payoff band 3.0 and the wallflower isolation band are not as strongly bimodal as the ecologies associated with the middle bands 2.3 and 2.8.

The payoff trajectories for individual ecologies exhibit other interesting features as well. For $\tau < 2.1$, a number of ecologies exhibit an abrupt payoff collapse, a feature also present in the Miller replication results. For example, one of the payoff trajectories depicted in Figure 10(a) abruptly collapses in generation 15 to a level well below that of all other trajectories, after which it slowly recovers. Examining in detail the ecology associated with this particular payoff trajectory, we saw that a spike occurred in the fraction of successful defections on the downward leg of the collapse, mutual defection became the dominant mode of play at the bottom of the collapse, and successful defection again became frequent during the subsequent payoff recovery. Note that the payoff of this ecology is roughly stable over generations 24 to 30 at the level 2.5, the average payoff level obtained by IPD players engaging in alternating *cd* and *dc* play; but the ecology payoff level subsequently climbs to the mutually cooperative payoff level 3.0.

It is also possible for ecologies to collapse to, and remain at, low payoff levels for an extended period of time. A dramatic example appears in Figure 11(a), where an ecology suffers a payoff collapse at about generation 22 and remains below the payoff level 2.0 for the remainder of the simulation.

"Band spiking" is another interesting feature exhibited by the payoff trajectories for individual ecologies for $\tau < 2.1$. An ecology sometimes exhibits an abrupt jump from one payoff band to another, followed by an abrupt jump back to its original payoff band. Preliminary evidence suggests that this spiking might be caused in part by the following type of population dynamics. Suppose the average payoff achieved by a type A player

is higher than that of a type *B* player as long as type *A* players are rare, but that the average payoff of type *A* players drops below that of type *B* players as type *A* players become more prevalent in the population due to reproduction or fortuitous crossover. The ecology payoff might then systematically oscillate between the type *A* and type *B* payoff levels. When an ecology of this type was saved and used to “seed” a set of forty new runs, the resulting payoff trajectories exhibited a great deal of spiking.

In addition, the payoff trajectories for individual ecologies exhibit “band tunneling,” particularly for high τ values. That is, ecologies that have been long-term residents of one payoff band suddenly exhibit an abrupt change in the gradient of their payoff score, followed by a rapid and smooth traverse to another payoff band. For example, in Figure 13(a) a number of ecologies tunnel directly from the wallflower isolation band to the mutually cooperative payoff band at 3.0; and one ecology tunnels from the wallflower isolation band to the payoff band at 2.8, remains at this payoff band for fifteen generations, and then tunnels to the mutually cooperative payoff band. These tunneling events appear to be correlated with the fraction of rejections. That is, as the fraction of rejections increases for an individual ecology, the payoff level of the individual ecology increases; and the fraction of rejections recedes as the ecology merges into a new payoff band. In addition, partial tunneling events also take place. For example, in Figure 14(a) an ecology rises out of the wallflower isolation band between generations 15 and 20, but then rapidly recedes back into it.

The properties displayed by *average* ecology payoffs are also interesting. Comparing Figure 10(b) with the replicated Miller results reported in Figure 1, the average ecology payoffs obtained for the low minimum tolerance level $\tau = 0.5$ display less of an initial dip than in the Miller replication, and do not attain as high an ultimate payoff level. These differences can be traced to the emergence of new payoff bands. As seen in Figure 10(a), these new bands are at low enough levels, and contain enough ecology payoffs, to decrease markedly the average payoff across ecologies. While a substantial fraction of the individual ecology payoff trajectories appear to converge—indeed, as many as in the Miller replication—many of these trajectories do *not* converge to the mutually cooperative payoff level. For low τ values, the evolution of cooperation is in fact retarded as compared to the Miller replication.

When τ is increased to 1.5 and to 1.6, the initial dip in average ecology payoffs completely disappears and the evolution of cooperation is substantially accelerated [Figures 11(b) and 12(b)]. Moreover, the fraction of *CC* play behavior markedly increases [Figures 11(c) and 12(c)]. The minimum tolerance level τ is now high enough to give the refusal mechanism some teeth. As τ continues to increase, however, the wallflower isolation payoff band eventually appears. The initial average ecology payoffs begin to flatten out again

and, by $\tau = 2.4$, an initial dip has reappeared [Figure 13(b)]. This dip is smaller than for the Miller replication and has an entirely different cause. Rather than representing the destruction of chump players by rapacious players, the dip results from the early formation of wallflower isolation ecologies which drag down the initial average ecology payoffs.

Finally, the fractions of distinct play behavior exhibited under different τ values are also of interest [Figures 10(c)-13(c)]. These fractions appear to undergo a phase change around $\tau = 2.1$. For $\tau < 2.1$, the simulation runs settle to a state where, on average, three types of PD game play behavior persist. The most common behavior is mutual cooperation CC , followed by mutual defection DD , and then by successful defection CD . For $\tau > 2.1$, however, successful defection effectively vanishes after the first few generations of players. This finding is predictable; as noted in section 5.2, for $\tau > 2.1$, any successful defection results in the refusal of all future PD game offers. Mutual defection also becomes less common, appearing primarily in those ecologies constituting the wallflower isolation payoff band [Figure 14(c)].

8. CONCLUDING REMARKS

The incorporation of a choice/refusal mechanism into the standard IPD framework has resulted in the emergence of several new ecological features. First, even without evolution, choice and refusal permit a broader range of interactions between mutually cooperative players, between potential hosts and would-be parasites, and between like-minded opportunists: for example, internally generated ostracism, persistent parasitism, and assortative partner selection. Second, with evolution, choice and refusal can accelerate the emergence of cooperation. More generally, however, choice and refusal can lead to the emergence of multiple tight payoff bands, reflecting the possible existence of multiple ecological attractors. The existence of a spectrum of payoff bands in turn permits the emergence of interesting new ecological behaviors, such as band spiking and band tunneling.

We hope to obtain a better quantitative understanding of these new ecological features as we undertake more elaborate studies of our choice and refusal mechanism. For example, it now seems clear that the genetic diversity of an ecology should be tracked throughout each simulation run, and that more detailed records should be maintained concerning which types of ecologies exhibit band spiking, band tunneling, and payoff collapse. At present, we have examples of ecologies that belong to the various payoff bands; but only the mutual cooperation and wallflower isolation payoff bands are well understood. The emergence of the middle payoff bands remains somewhat mysterious. We are currently conducting extensive sensitivity studies in order to ascertain the conditions under which these middle payoff bands form. Preliminary studies suggest that the number of automaton states used

by players in their interactions with other players is an important factor affecting the formation and stability of these bands.

In addition, a number of revisions to our model and methods are currently under consideration. For example, the minimum tolerance level τ and the initial expected payoff π^0 —both currently set as fixed exogenous constants—should instead be allowed to evolve over time along with the structure of the PD rules. As seen in both the analytical and simulation studies reported here, higher levels of τ result in increased resistance to parasitic relationships; and differing resistance to parasites has been conjectured as one reason why populations divide into smaller reproductive groups [Rennie (1992)]. Moreover, π^0 turns out to be a critical parameter that affects both the willingness of players to tolerate defections and the willingness of players to seek out new potential partners.

Also, using a fixed convex combination of past expected payoff and current payoff in order to obtain an updated expected payoff is too simplistic, even in the present context in which players know nothing a priori about other player types. For example, surely a player could learn, through observation, that the current move of an opponent is functionally dependent on the previous move he made against that opponent. The memory weights used to obtain updated expected payoffs could reflect this dependence. For example, if a player's previous move against a current opponent was c , the updated expected payoff for this opponent could be a weighted average only over all past payoffs obtained under the same circumstance, i.e., obtained when the previous move against this opponent was c . Moreover, once a player learns that his current moves affect the later moves of other players, deliberate attempts to modify these later moves become possible.

Finally, the analysis of individual finite automata is difficult and time consuming. We are currently attempting to develop automatic methods for detecting unusual or interesting play behavior. We are also attempting to adapt some of the techniques of systematics from evolutionary biology in order to develop better descriptive methods for studying our artificial ecologies.

ACKNOWLEDGMENTS

Partially supported by ISU University Research Grant No. 430-17-03-91-0001. A preliminary version of this paper was presented at the ALife III Conference sponsored by the Santa Fe Institute, Santa Fe, New Mexico, June 13-19, 1992. The authors are grateful to conference participants and to three anonymous referees for helpful comments.

REFERENCES

1. Axelrod, R. "The Emergence of Cooperation Among Egoists." *American Political Science Review* 75 (1981): 306-318.
2. Axelrod, R. *The Evolution of Cooperation*. New York: Basic Books, 1984.
3. Axelrod, R., and W. D. Hamilton. "The Evolution of Cooperation." *Science* 211 (1981): 1390-1396.
4. Axelrod, R., and Douglas Dion. "The Further Evolution of Cooperation." *Science* 242 (December 1988): 1385-1390.
5. Barner-Berry, C. "Rob: Children's Tacit Use of Peer Ostracism to Control Aggressive Behavior." *Ethology and Sociobiology* 7 (1986): 281-293.
6. Dugatkin, L. A., "Rover: A Strategy for Exploiting Cooperators in a Patchy Environment," *The American Naturalist* 138 (1991), 687-701.
7. Eshel, I. and L. L. Cavalli-Sforza. "Assortment of Encounters and Evolution of Cooperativeness." *Proceedings of the National Academy of Sciences, U.S.A.* 79 (February 1982): 1331-1335.
8. Feldman, M., and E. A. C. Thomas. "Behavior-Dependent Contexts for Repeated Plays of the Prisoner's Dilemma II: Dynamical Aspects of the Evolution of Cooperation." *Journal of Theoretical Biology* 128 (1987): 297-315.
9. Hirshleifer, J. "On the Emotions as Guarantors of Threats and Promises." In *The Latest on the Best: Essays on Evolution and Optimality*, edited by J. Dupre, Cambridge: MIT Press, 1987.
10. Hirshleifer, D., and E. Rasmusen. "Cooperation in a Repeated Prisoners' Dilemma with Ostracism." *Journal of Economic Behavior and Organization* 12 (1989): 87-106.
11. Holland, J. *Adaptation in Natural and Artificial Systems*. Cambridge: The MIT Press, 1992.
12. Hyman, J. M., and E. A. Stanley. "Using Mathematical Models to Understand the AIDS Epidemic." *Mathematical Biosciences* 90 (1988): 415-473.
13. Kitcher, Philip, "Evolution of Altruism in Repeated Optional Games," Working Paper, Department of Philosophy, University of California at San Diego, July 1992.
14. Kreps, D., Milgrom, P., Roberts, J., and R. Wilson. "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma." *Journal of Economic Theory* 27 (1982): 245-252.
15. May, R. M. "More Evolution of Cooperation." *Nature* 327 (May 1987): 15-17.
16. Maynard Smith, J. *Evolution and the Theory of Games*. United Kingdom: Cambridge University Press, 1982.

17. Miller, J. H. "The Coevolution of Automata in the Repeated Prisoner's Dilemma," Revised Working Paper No. 8903, Santa Fe Institute, July 1989.
18. Rennie, J. "Trends in Parasitology: Living Together." *Scientific American* 266 (January 1992): 122-133.
19. Stanley, A., Ashlock, D., and L. Tesfatsion, "Iterated Prisoner's Dilemma with Choice and Refusal," Economic Report No. 30, Iowa State University, July 1992.
20. Tesfatsion, L. "Direct Updating of Intertemporal Criterion Functions for a Class of Adaptive Control Problems." *IEEE Transactions on Systems, Man, and Cybernetics*. SMC-9 (1979): 143-151.
21. Thompson, E., and R. Faith. "A Pure Theory of Strategic Behavior and Social Institutions." *American Economic Review* 71 (1981): 366-380.

FIGURE 1
Average Ecology Payoffs for the Replicated Miller Game

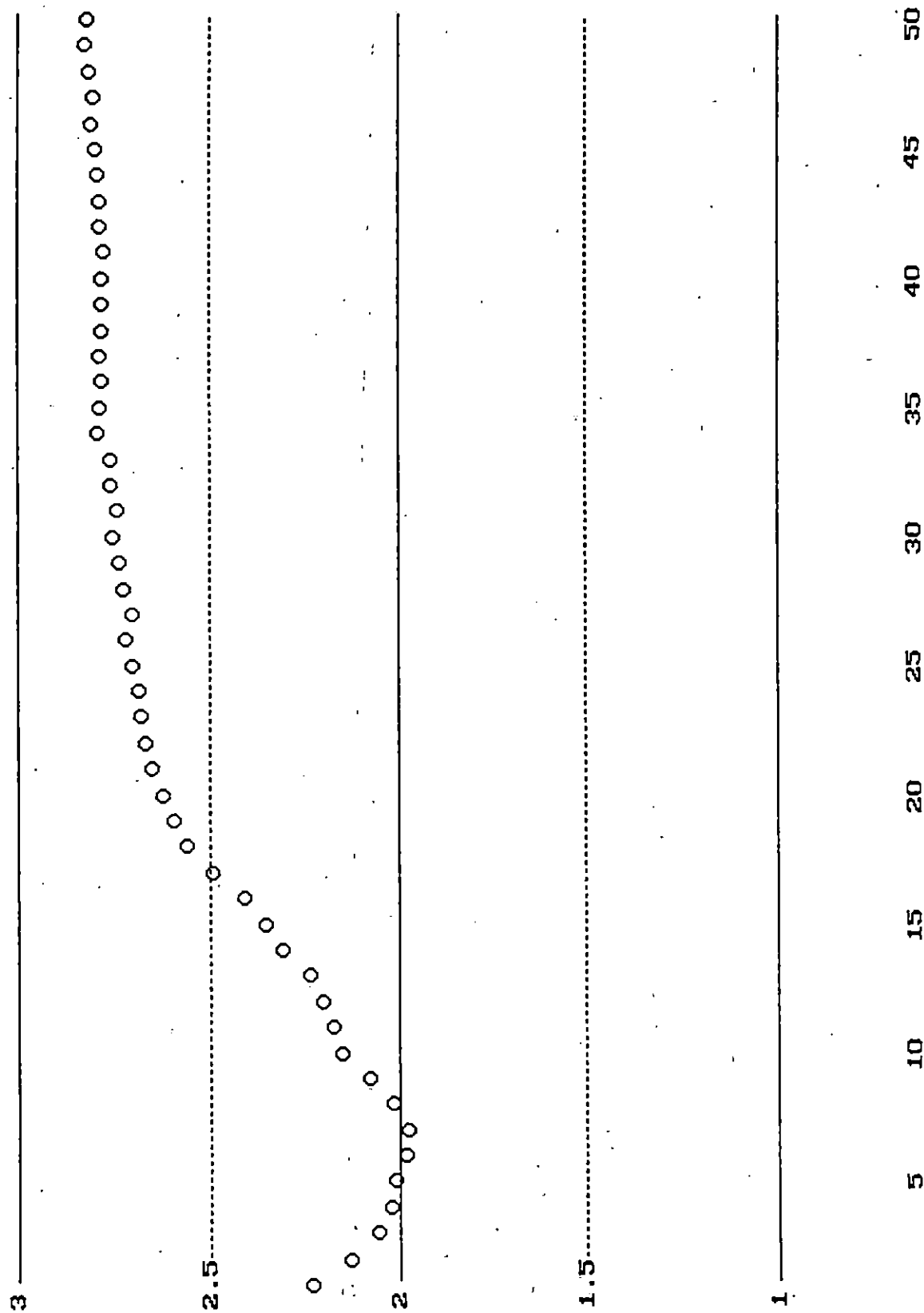


FIGURE 2
Average Ecology Payoffs for the IPD/CR Game with $\tau=1.6$

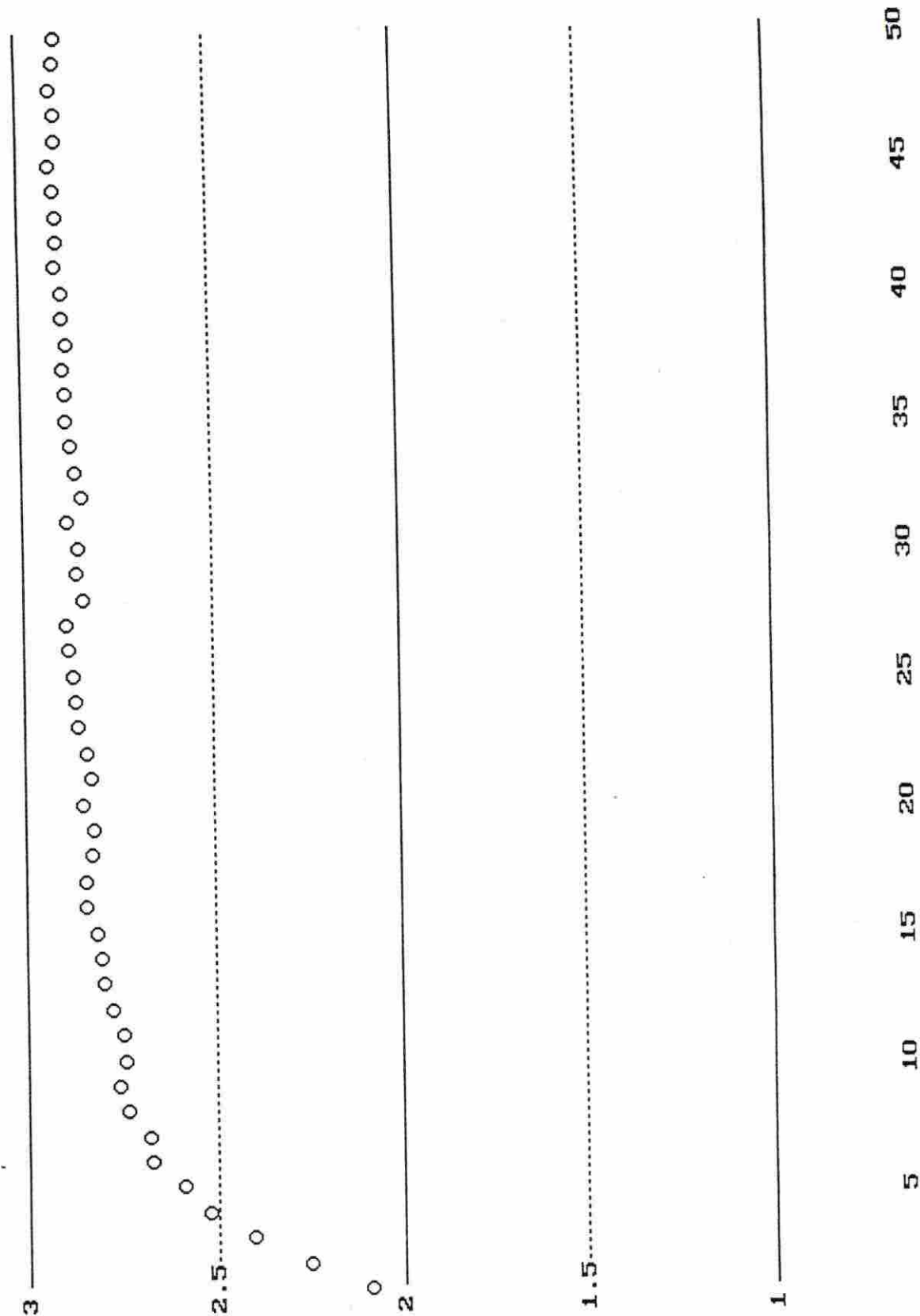


FIGURE 3
Individual Ecology Payoffs for the Replicated Miller Game

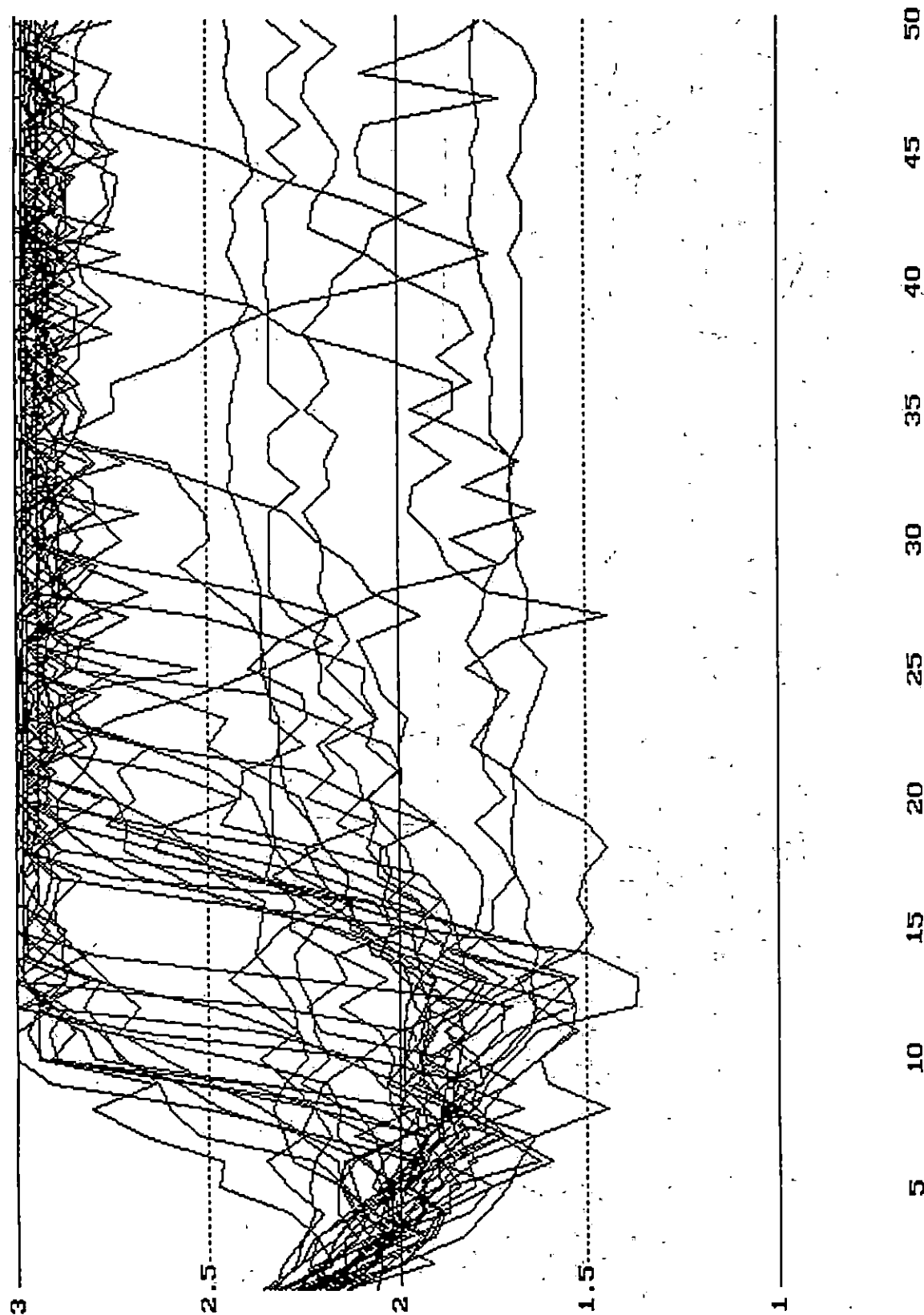


FIGURE 4
Individual Ecology Payoffs for the IPD/CR Game with $\tau=1.6$

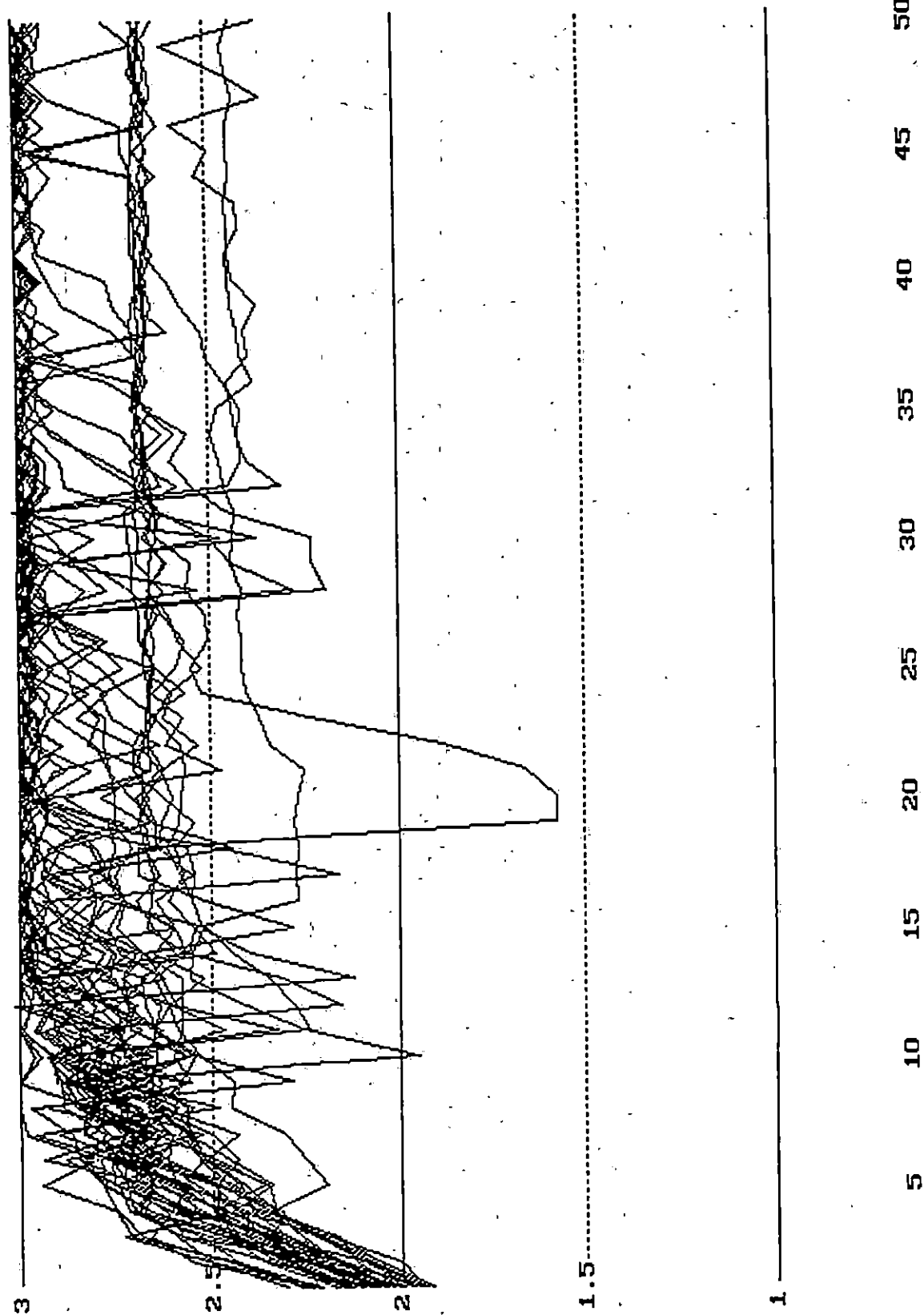
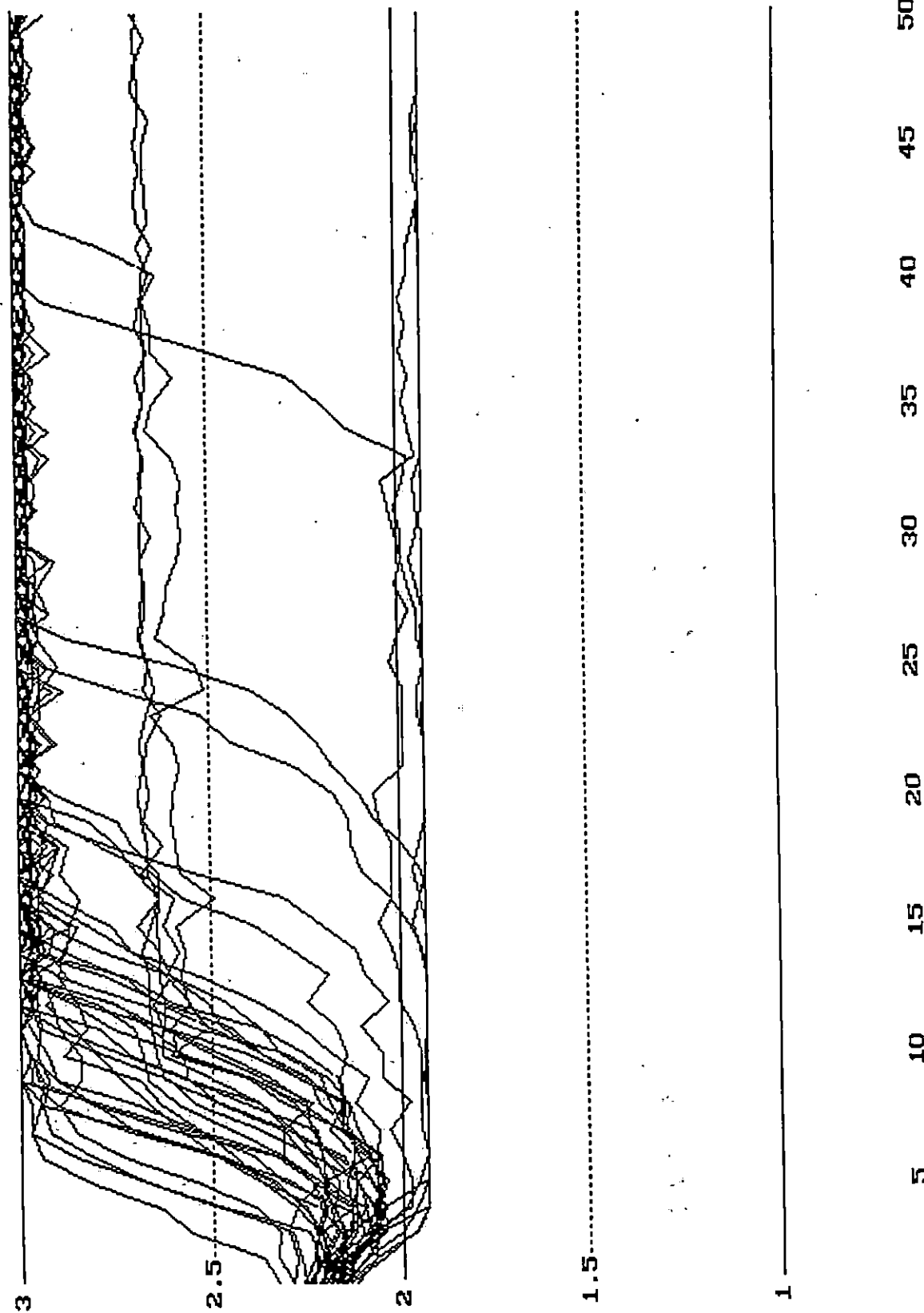


FIGURE 5
Individual Ecology Payoffs for the IPD/CR Game with $\tau=2.5$



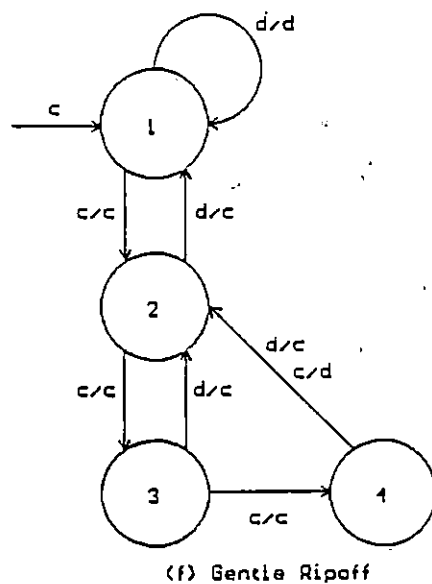
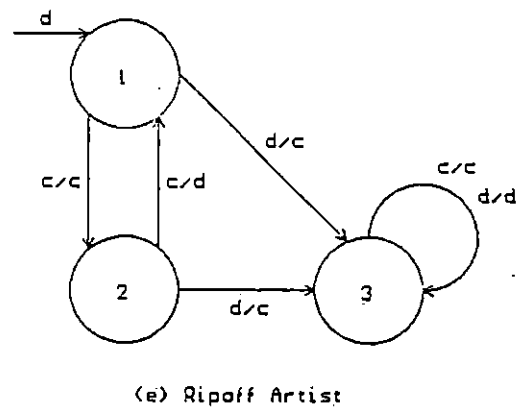
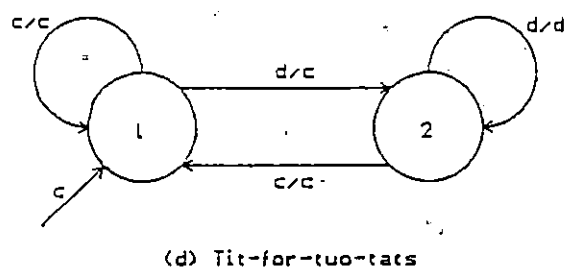
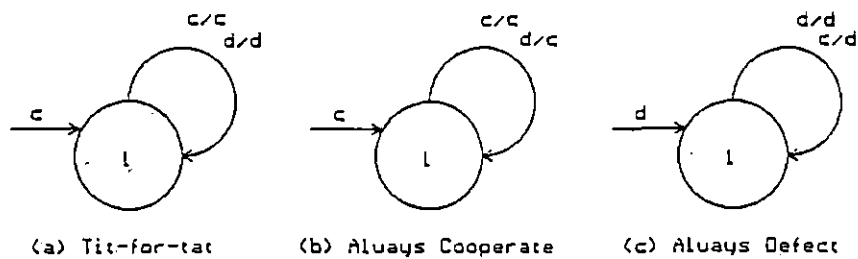


FIG. 6: Example Players (a) Tit-for-Tat; (b) Always Cooperate; (c) Always Defect; (d) Tit-for-Two-Tats; (e) Ripoff-Artist; (f) Gentle Ripoff

Table 1: Pairwise interactions prior to the first refusal

| players in 2x2 games | π^i for $i \leq i^*$ | Conditions for PD games to stop | i^* | Refuser | Average payoff for $i \leq i^*$ |
|-----------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| AllD v. AllD | $\omega^i \pi^0 + (1 - \omega^i)$ | $1 < \tau$ | $\left\lceil \ln\left(\frac{\tau-1}{\pi^0-1}\right) \frac{1}{\ln(\omega)} \right\rceil$ | both players | 1 |
| Ripoff v. Ripoff | for $i=1$: $\omega \pi^0 + (1-\omega)$; for $i>1$: $\omega^i \pi^0 + \omega^{i-1} (1-\omega) + C (1 - \omega^{i-1})$ | $\omega \pi^0 + (1-\omega) < \tau$ | if $\omega \pi^0 + (1-\omega) < \tau$: $i^*=1$ otherwise: $i^* = \infty$ | both players | for $i=1$: 1; for $i>1$: $\frac{1}{i} + \frac{i-1}{i} C$. |
| Any pair of AllC, TFT, or TFTT | $\omega^i \pi^0 + C (1 - \omega^i)$ | never | $i^* = \infty$ | neither | C |
| AllD v. Ripoff | Let: X = S for AllD and 0 for Ripoff for $i=1$: $\omega \pi^0 + (1-\omega)$ for $i=2$: $\omega^2 \pi^0 + \omega (1-\omega) + (1-\omega)X$ for $i \geq 3$: $\omega^i \pi^0 + \omega^{i-2} (1-\omega)X$ $+ (1-\omega)^{i-2} + \omega^{i-1} (1-\omega)$ | $\omega^2 \pi^0 + \omega (1-\omega) < \tau$ or $1 < \tau$ | if $\omega \pi^0 + (1-\omega) < \tau$: $i^*=1$; else, if $\omega^2 \pi^0 + \omega (1-\omega) < \tau$: $i^*=2$; else, if $1 < \tau$: $i^* = \left\lceil \ln\left(\frac{\omega^2(\tau-1)}{\omega^2 \pi^0 - (1-\omega+\omega^2)}\right) \frac{1}{\ln(\omega)} \right\rceil$ otherwise: $i^* = \infty$ | if $i^*=1$: both players if $i^*=2$ or $i^*=3$: Rip only if $i^*=3$: Rip and maybe AllD[1] | for $i=1$: 1 for $i=2$: $(1+X)/2$ for $i \geq 3$: $((i-1)+X)/i$ |
| AllD v. TFT or GRip | Let: X = S for AllD and 0 for TFT and GRip for $i=1$: $\omega \pi^0 + (1-\omega)X$ for $i \geq 2$: $\omega^i \pi^0 + \omega^{i-1} (1-\omega)X + (1-\omega)^{i-1}$ | $\omega \pi^0 < \tau$ or $1 < \tau$ | if $\omega \pi^0 < \tau$: $i^*=1$ else, if $1 < \tau$: $i^* = \left\lceil \ln\left(\frac{1}{\ln(\omega)} \ln\left(\frac{\omega(\tau-1)}{\omega \pi^0 - 1}\right)\right) \right\rceil$ otherwise: $i^* = \infty$ | TFT or GRip only[2] | for $i=1$: X for $i \geq 2$: $(X+(i-1))/i$ |
| AllD v. AllC | Let: X = S for AllD and 0 for AllC $\omega^i \pi^0 + X (1 - \omega^i)$ | $0 < \tau$ | if $0 < \tau$: $i^* = \left\lceil \ln\left(\frac{\tau}{\pi^0}\right) \frac{1}{\ln(\omega)} \right\rceil$ | AllC only | X |
| AllD v. TFTT | Let: X = S for AllD and 0 for TFTT for $i=1$: $\omega \pi^0 + (1-\omega)X$ for $i=2$: $\omega^2 \pi^0 + (1-\omega^2)X$ for $i \geq 3$: $\omega^i \pi^0 + \omega^{i-2} (1-\omega^2)X + (1-\omega)^{i-2}$ | $\omega^2 \pi^0 < \tau$ or $1 < \tau$ | if $\omega \pi^0 < \tau$: $i^*=1$; else, if $\omega^2 \pi^0 < \tau$: $i^*=2$; else, if $1 < \tau$: $i^* = \left\lceil \ln\left(\frac{1}{\ln(\omega)} \ln\left(\frac{\omega^2(\tau-1)}{\omega^2 \pi^0 - 1}\right)\right) \right\rceil$ otherwise: $i^* = \infty$ | TFTT only[2] | $i=1, 2$: X $i \geq 2$: $\frac{2X+(i-2)}{i}$ |

| | | | | | |
|---------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Ripoff v. AIIc or TFTT | <p>Let: X = S for Ripoff and 0 for AIIc and TFTT</p> <p>for i = 2k: $\omega^2 \pi^0 + \frac{1 - \omega^{2k}}{1 + \omega} (C + \omega X)$;</p> <p>for i = 2k+1: $\omega^{2k+1} \pi^0 + \omega \frac{1 - \omega^{2k}}{1 + \omega} (C + \omega X) + (1 - \omega) X$.</p> | <p>if $\pi^0(1 + \omega) > C$ (π^{2k+1}_{TFTT} declines): $\frac{C\omega}{1 + \omega} < \tau$ otherwise, if $\pi^0(1 + \omega) \leq C$ (π^{2k+1}_{TFTT} increases): $\omega \pi^0 < \tau$</p> | <p>if $\pi^0(1 + \omega) > C$ and $\frac{C\omega}{1 + \omega} < \tau$: $i^* =$ The smallest odd integer greater than $\ln(\frac{\tau(1 + \omega) - C\omega}{\pi^0(1 + \omega) - C}) \frac{1}{\ln(\omega)}$ else if $\pi^0(1 + \omega) \leq C$ and $\omega \pi^0 < \tau$: $i^* = 1$ otherwise: $i^* = \infty$</p> | <p>AIIc or TFTT only</p> | <p>$i = 2k$: $\frac{C + X}{2}$ $i = 2k + 1$: $\frac{(k + 1)X + kC}{2k + 1}$</p> |
| Ripoff v. TFT | <p>Let: X = S for Ripoff and 0 for TFT Y = 0 for Ripoff and S for TFT</p> <p>for i = 1: $\omega \pi^0 + (1 - \omega) X$ for i = 2: $\omega^2 \pi^0 + \omega(1 - \omega) X + (1 - \omega) Y$ for i ≥ 3: $\omega^i \pi^0 + \omega^{i-1} (1 - \omega) X + \omega^{i-2} (1 - \omega) Y + C(1 - \omega^{i-2})$</p> | <p>$\omega \pi^0 < \tau$</p> | <p>if $\omega \pi^0 < \tau$: $i^* = 1$; otherwise: $i^* = \infty$</p> | <p>TFT only</p> | <p>$i = 1$: X $i \geq 2$: $\frac{S + (i - 2)C}{i}$</p> |
| GRip v. TFTT or AIIc | <p>Let: X = S for GRip, and 0 for TFTT and AIIc</p> <p>for i ≤ 4: $\omega^i \pi^0 + (1 - \omega^i) C$ for i = 4 + 3j, j = 0, 1, 2, ... $\omega^3 j \pi^4 + \frac{1 - \omega^{3j}}{1 + \omega + \omega^2} (\omega^2 X + (1 + \omega) C)$ for i = 5 + 3j, j = 0, 1, 2, ... $\omega^{3j+1} \pi^4 + \frac{\omega(1 - \omega^{3j})}{1 + \omega + \omega^2} (\omega^2 X + (1 + \omega) C) + (1 - \omega) X$ for i = 6 + 3j, j = 0, 1, 2, ... $\omega^{3j+2} \pi^4 + \frac{\omega^2(1 - \omega^{3j})}{1 + \omega + \omega^2} (\omega^2 X + (1 + \omega) C) + (1 + \omega) C + (1 - \omega)(\omega X + C)$</p> | <p>$\frac{\omega(1 + \omega) C}{1 + \omega + \omega^2} < \tau$</p> | <p>if $\frac{\omega(1 + \omega) C}{1 + \omega + \omega^2} < \tau$ $i^* = 2 + 3j$, where j is the smallest positive integer such that $3j > \frac{\ln(\frac{\tau - \omega(1 + \omega)(C - \tau)}{\omega^2(1 + \omega + \omega^2)(\pi^0 - C) + C})}{\ln \omega}$ otherwise $i^* = \infty$</p> | <p>TFTT or AIIc</p> | <p>$i \leq 4$: C for j = 0, 1, 2, ... $i = 4 + 3j$: $\frac{(5 + 2j)C + jS}{4 + 3j}$ $n = 5 + 3j$: $\frac{(5 + 2j)C + (j + 1)S}{5 + 3j}$ $i = 6 + 3j$: $\frac{(6 + 2j)C + (j + 1)S}{6 + 3j}$</p> |

| | | | | | |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Rip | <p>Let:</p> <p>$X=S, Y=0$ for GRip, and $X=0, Y=S$ for Rip $Z=Y+\omega X+(1+\omega+\omega^2)\omega^2 C$ $q=1+\omega+\omega^2+\omega^3+\omega^4$</p> <p>for $i=1$: $\omega\pi^0+(1-\omega)Y$ for $i=2$: $\omega^2\pi^0+(1-\omega)(\omega Y+X)$ for $i=2+5j, j=0,1,2,\dots$: $\omega^5j\pi^2+\frac{1-\omega^5}{q}Z$ for $i=2+m+5j, m=1,2,3$ and $j=0,1,2,\dots$ $\omega\pi^{2+5j}+(1-\omega^m)C$ for $i=6+5j, j=0,1,2,\dots$ $\omega^4\pi^{2+5j}+(1-\omega^3)\omega C+(1-\omega)X$ Special case: $\pi_{Rip}^6=\omega^6\pi^0+(1-\omega)[(\omega+\omega^2+\omega^3)C+\omega^5S]$</p> | $\omega\pi^0 < \tau$ or $\pi_{Rip}^0 \geq C$ and $\pi_{Rip}^6 < \tau$ or [3] (Rip's expected payoffs are decreasing each cycle of 5 games) $\pi_{Rip}^2 > Z_{Rip}/q$ and (eventually become less than τ) $((\omega+\omega^2+\omega^3)C$ $+\omega^4S)/q < \tau$ | <p>if $\omega\pi^0 < \tau$: $i^*=1$ else if $\pi_{Rip}^6 < \tau$: $i^*=6$; else if</p> <p>$\frac{\omega^4S+(1+\omega+\omega^2)\omega C}{1+\omega+\omega^2+\omega^3+\omega^4} < \tau$, and $[(1-\omega)S+(\omega^2+\omega^3+\omega^4)C+\omega^6S]/q$ $< \pi_{Rip}^0$, $i^*=1+5j$, j the smallest positive integer such that $\ln(\frac{\tau-\omega C+\omega^4(C-Z_{Rip}/q)}{\pi_{Rip}^2-Z_{Rip}/q})$ $5j > 1+\frac{\ln(\omega)}{\ln(\omega)}$</p> | <p>if $i^*=1$ GRip only otherwise Rip only</p> | <p>$i=1$: Y $i=2$: $\frac{S}{2}$ $i=2+5j+m, m=0,1,2$ and $j=0,1,2,\dots$ $\frac{(j+1)S+(3j+m)C}{2+m+5j}$ $i=5j, j=1,2,\dots$ $\frac{jS+3jC}{5j}$ $i=5j+1, j=1,2,\dots$ $\frac{jS+3jC+X}{5j+1}$</p> |
| GRip | <p>Letting</p> <p>$X=S, Y=0$ for GRip, and $X=0, Y=S$ for TFT $Z=Y+\omega X+(1+\omega+\omega^2)\omega^2 C$ $q=1+\omega+\omega^2+\omega^3+\omega^4$</p> <p>for $i=1$: $\omega\pi^0+(1-\omega)C$ for $i=5j+1, j=0,1,2,\dots$: $\omega^5j\pi^1+\frac{1-\omega^5}{q}Z$ for $i=5j+1+m, m=1,2,3$ and $j=0,1,2,\dots$ $\omega\pi^{5j+1}+(1-\omega^m)C$ for $i=5j+5, j=0,1,2,\dots$: $\omega^4\pi^{5j+1}+(1-\omega^3)\omega C+(1-\omega)X$</p> | $\omega^5\pi^0+C(1-\omega^4)\omega < \tau$ or $(\pi_{TFT}^5$ is decreasing with j) $S-C < \omega\pi^0+\omega^2(1+\omega+\omega^2+\omega^3)(\pi^0-C)$ and (eventually becomes less than τ) $\frac{(\omega+\omega^2+\omega^3)C+\omega^4S}{q} < \tau$ | <p>if $\omega^5\pi^0+C(1-\omega^4)\omega < \tau$: $i^*=5$ else if $S-C < \omega\pi^0+\omega^2(1+\omega+\omega^2+\omega^3)(\pi^0-C)$ and $\frac{\omega^4S+(1+\omega+\omega^2)\omega C}{1+\omega+\omega^2+\omega^3+\omega^4} < \tau$, $i^*=5j+5$, j the smallest positive integer such that $5j+4 > \ln((q-\omega^4S+(1+\omega+\omega^2)\omega C)/(\omega\pi^0+\omega^2(1+\omega+\omega^2+\omega^3)(\pi^0-C)-S+C))/\ln(\omega)$</p> | <p>TFT only</p> | <p>$i=1,2,3,4$: C $i=5j, j=1,2,\dots$ $\frac{(j-1)S+(3j+1)C+X}{5j}$ $i=5j+m, m=1,2,3,4$ and $j=1,2,\dots$ $\frac{jS+(3j+m)C}{1+5j}$</p> |
| GRip | <p>for $i=1,2,3,4$: $\omega^1\pi^0+(1-\omega)C$ for $i=4j+1, j=1,2,\dots$: $\omega^4j\pi^1+\frac{1-\omega^4}{1-\omega^4}((1-\omega)+\omega(1-\omega^3)C)$ for $i=4j+1+m, m=1,2,3$ and $j=0,1,2,\dots$: $\omega\pi^{4j+1}+(1-\omega^m)C$</p> | $\tau > A$ where $A = \frac{(1-\omega)+\omega(1-\omega^3)C}{1-\omega^4}$ | <p>if the condition holds: $i^*=4j+1$ where j is the smallest positive integer greater than $\frac{\ln(\frac{\tau-A}{\pi^1-A})}{4 \ln \omega}$, otherwise: $i^*=\infty$</p> | <p>Both players</p> | <p>$i=1,2,3,4$: C $i=4j+m, m=1,2,3,4$, and $j=1,2,\dots$: $\frac{j+(3j+m)C}{4j+1}$</p> |

¹ Both players (ALID and Ripoff) may decide not to play each other each other on iteration i^*+1 if both expected payoffs cross τ on iteration i^* , but not before. If a subscript D denotes the ALID expected payoff, then $\pi_D^i < \tau$. In order for both expected payoffs to be less than τ before iteration i^* and greater than τ after it, necessary, but not sufficient, conditions are: if $i^*=3, 0 > \omega(S-\omega\pi^0)-2\omega+1$. If $i^* \geq 4$, it is not possible since $\pi^0 \leq S$.

2 Since we assume $\pi^0 \leq S$, it is not possible for both players (AllD and TFT or AllD and TFTT) to refuse each other (mutual decision not to play) on iteration i^*+1 if $\pi_D^{i^*} < \tau$ (see [1]).

3 The condition that Ripoff's payoffs are decreasing is the condition $[(1-\omega)S + (\omega^2 + \omega^3 + \omega^4)C + \omega^6 S]/q < \pi^0$. Since Ripoff does not refuse after iteration $2+5j$, Z_{Rip} does not have to be less than τ for refusal to occur. However, in order for GRip to eventually refuse Rip, its payoffs must not only be decreasing, but they must drop below τ after a $2+5j$ iteration. Note also that since we assume $\pi^0 \leq S$, it is not possible for GRip's expected payoffs after the iterations where it receives less than those that Rip has after the previous iteration.

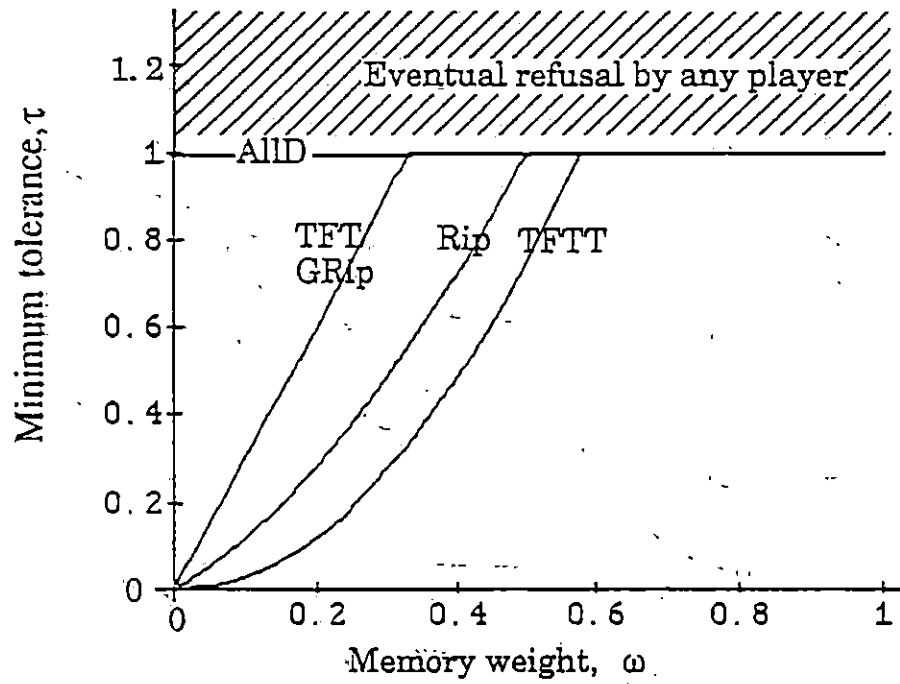


FIG. 7(a): A player eventually finds AllD intolerable in an IPD/CR game play if τ and ω lie in the region above the curve indicated for this player's type. [AllD is always eventually refused by AllC]. The parameter values are $B = 0$, $D = 1$, $C = 3$, $S = 5$, and $\pi^0 = 3$.

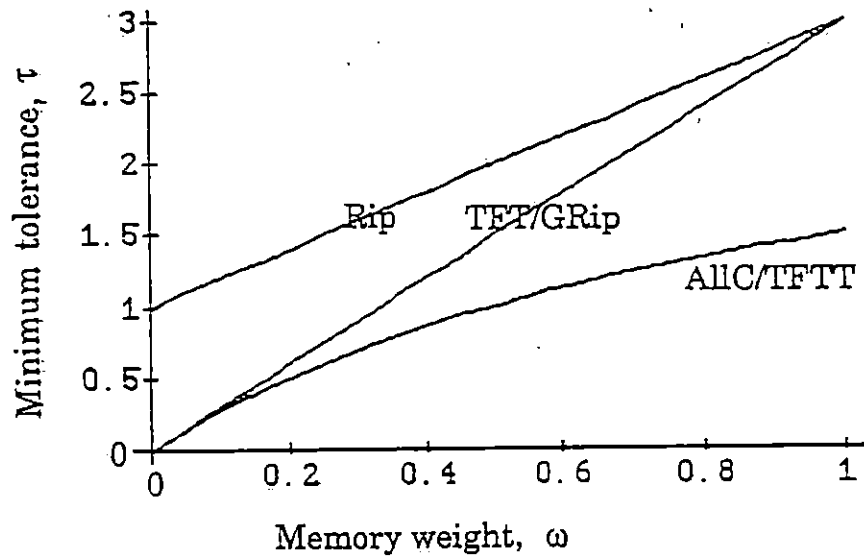


FIG. 7(b): A player eventually finds Rip intolerable in an IPD/CR game play if τ and ω lie in the region above the curve indicated for this player's type. Play stops between two Rips by mutual assent; but, in the other indicated cases, Rip is eventually refused. The parameter values are $B = 0$, $D = 1$, $C = 3$, $S = 5$, and $\pi^0 = 3$.

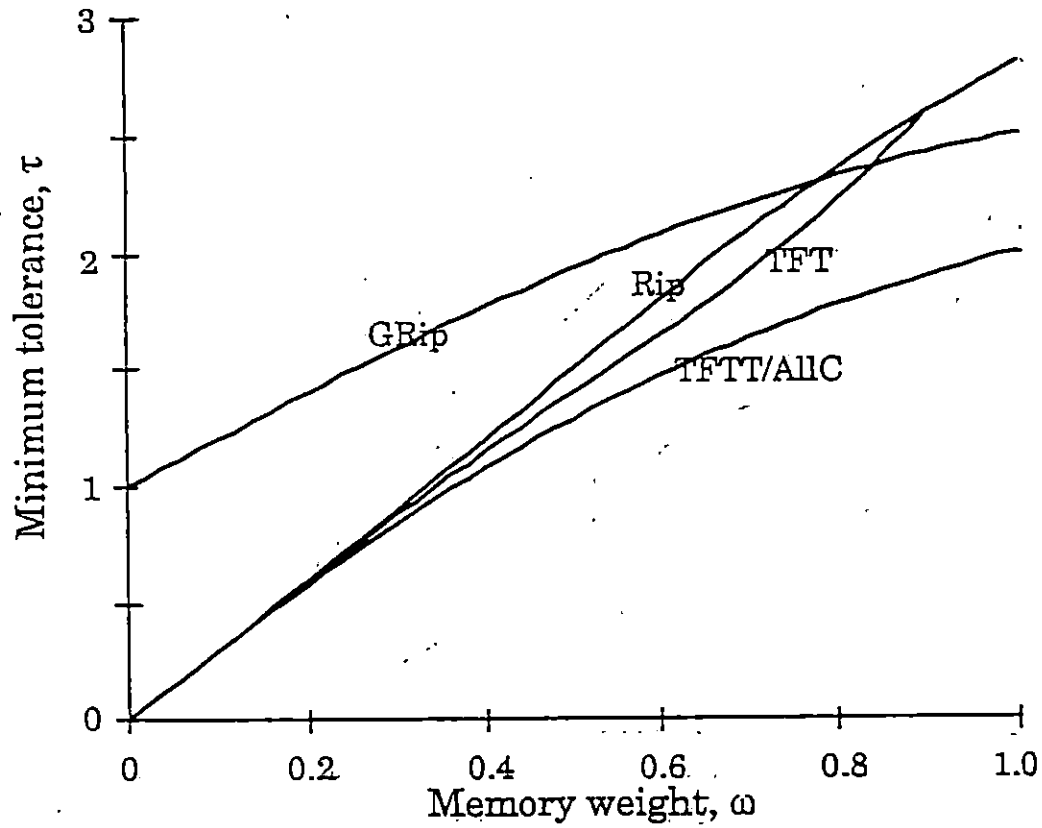


FIG. 7(c): A player eventually finds GRip intolerable in an IPD/CR game play if τ and ω lie in the region above the curve indicated for this player's type. The parameter values are $B = 0$, $D = 1$, $C = 3$, $S = 5$, and $\pi^0 = 3$.

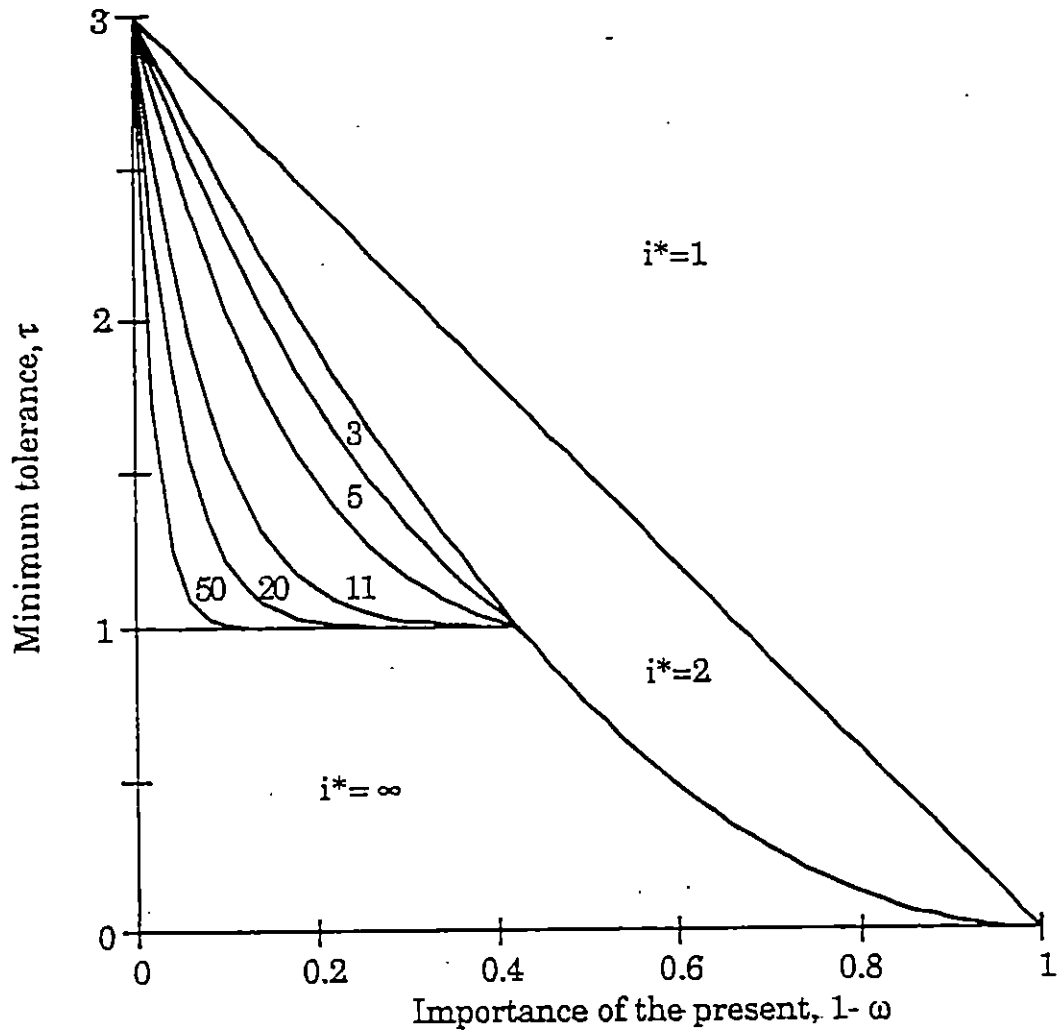


FIG. 8(a): The total number of PD games played in a 2-player IPD/CR game between AllD and TFTT before play stoppage, as a function of τ and ω . The number i^* gives the last iteration during which a PD game takes place in a region. The other numbers indicate the maximum number of games played in a region. Each region includes its upper boundary line but not its lower boundary line. Except for the top boundary line, all boundary lines cross each other exactly once, at $\omega = 1/\sqrt{3}$ and $\tau = 1$. The d/c move sequence in the first two iterations provokes refusal by TFTT in a large region of this plane.

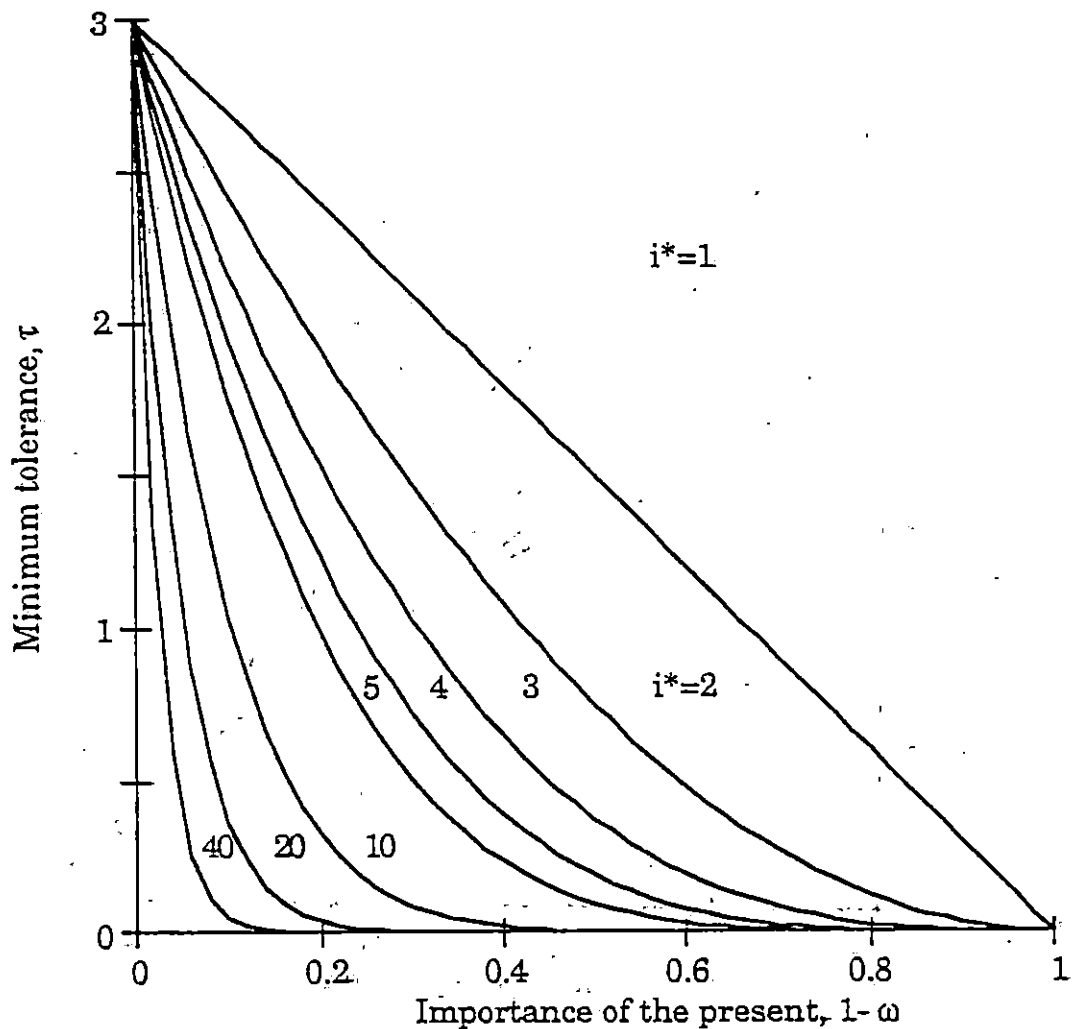


FIG. 8(b): The total number of PD games played in a 2-player IPD/CR game between AllD and AllC before play stoppage, as a function of τ and ω . The number i^* gives the last iteration during which a PD game takes place in a region. The other numbers indicate the maximum number of games played in a region. Each region includes its upper boundary line but not its lower boundary line.

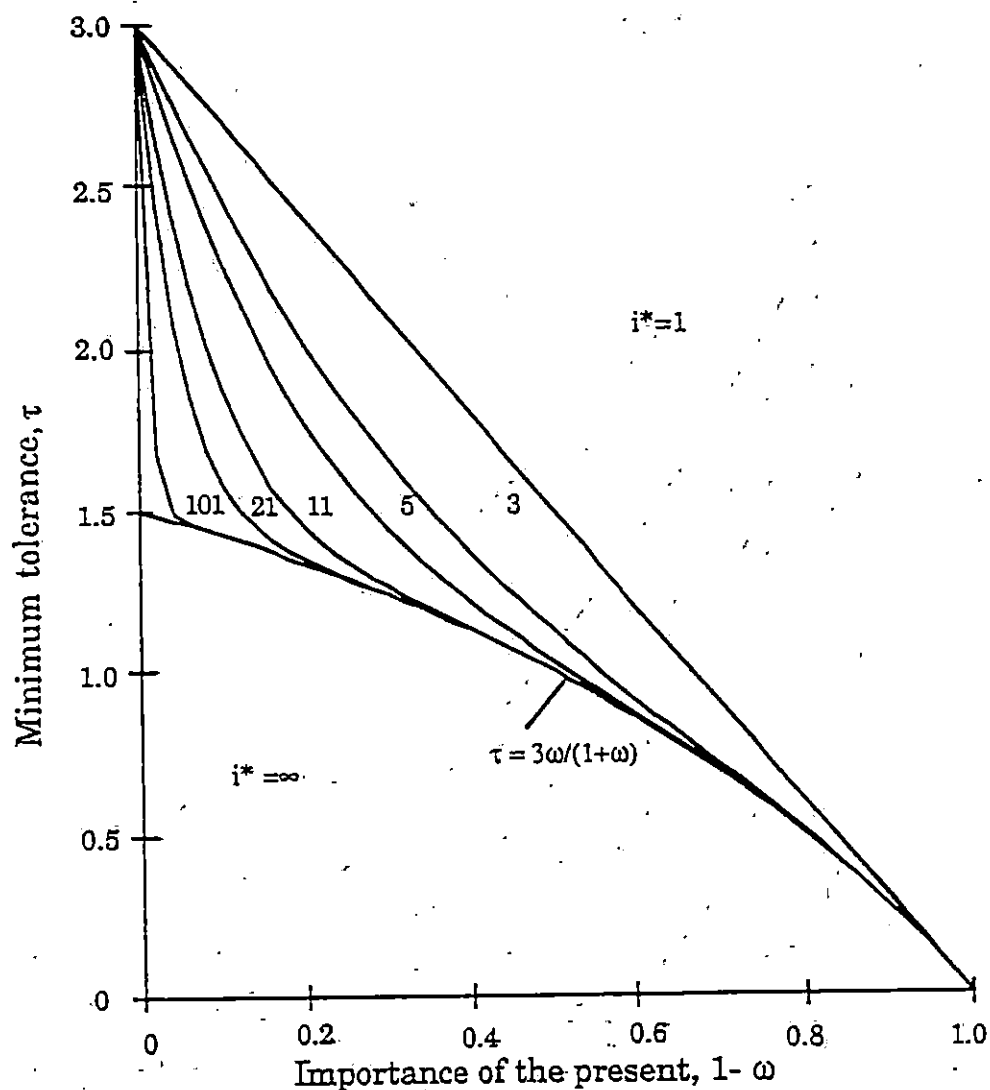


FIG. 8(c): The total number of PD games played in a 2-player IPD/CR game between Rip and either ALLC or TFFT before play stoppage, as a function of τ and ω . The number i^* gives the last iteration during which a PD game takes place in a region. The other numbers indicate the maximum number of games played in a region. Each region includes its upper boundary line but not its lower boundary line. The boundary lines cross the line $\tau = 3\omega/(1+\omega)$ only at the axis.

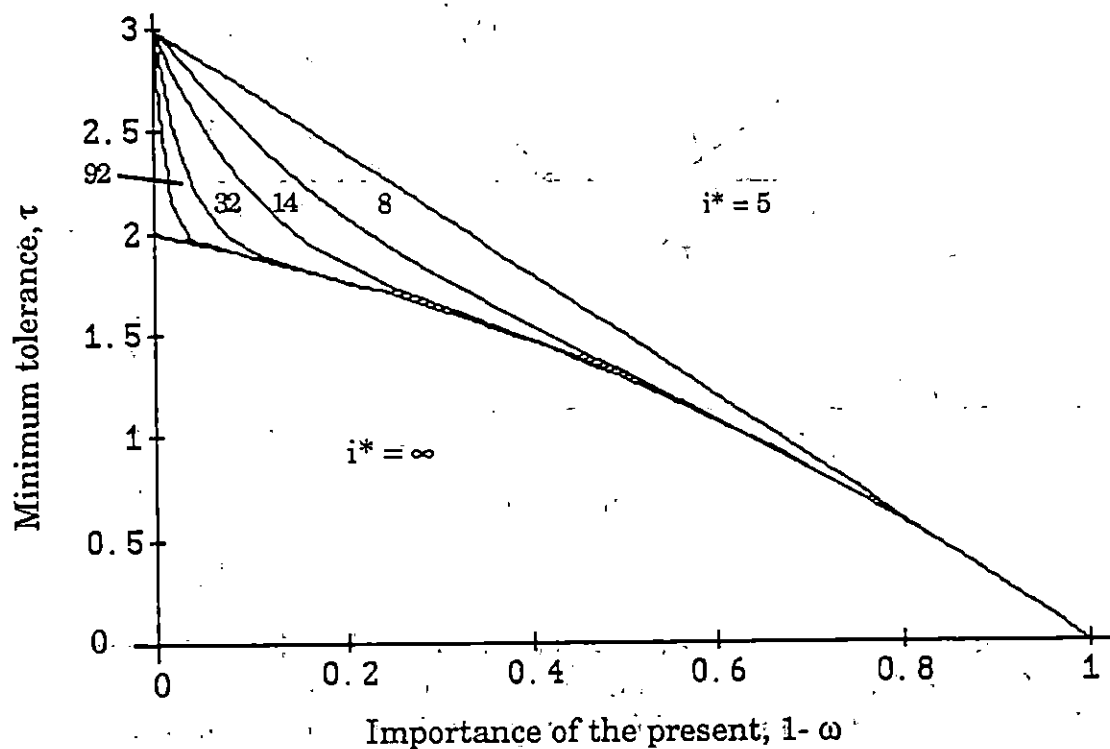


FIG. 8(d): The total number of PD games played in a 2-player IPD/CR game between GRip and either AllC or TFTT before play stoppage, as a function of τ and ω . The number i^* gives the last iteration during which a PD game takes place in a region. The other numbers indicate the maximum number of games played in a region. Each region includes its upper boundary line but not its lower boundary line. The boundary lines are tangent to the line dividing eventual play stoppage ($i^* < \infty$) from no stoppage ($i^* = \infty$), and they meet at each axis.

Rip versus 3 TFTT players

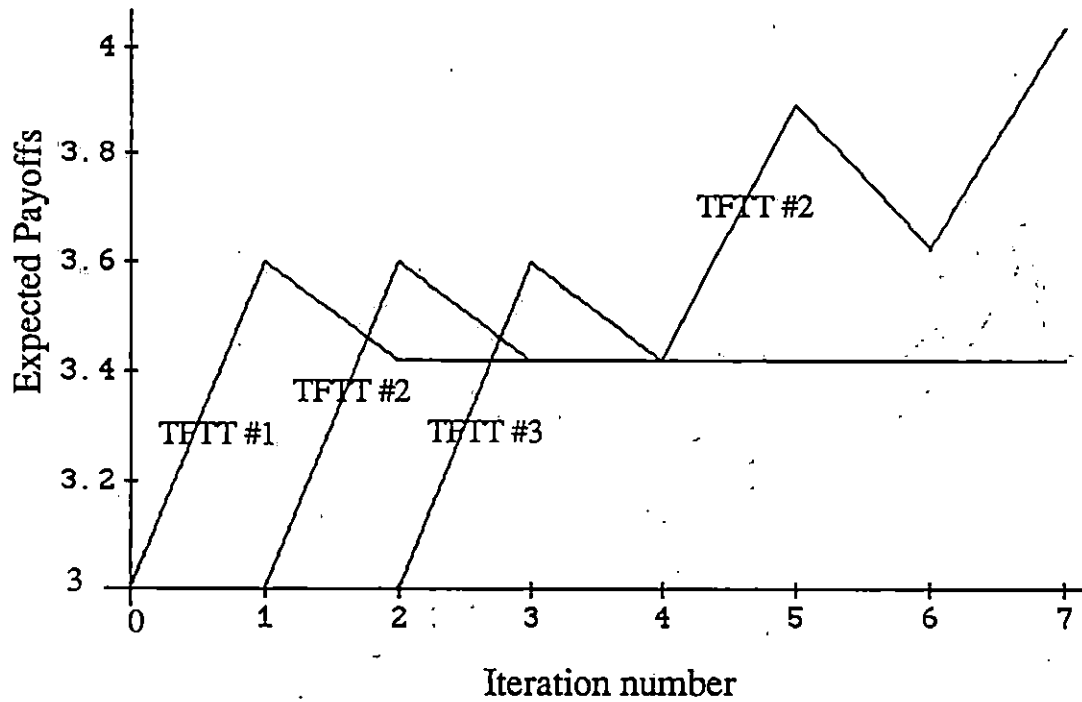


FIG. 9: A schematic of the expected payoffs that a Rip anticipates from three TFTT (or AllC) players when $K =$ and $\pi^0 = C$. In this sample scenario, Rip plays TFTT #1, TFTT #2, and TFTT #3 for the first time in iterations 1, 2, and 3, respectively. Before iteration 5, Rip has the same expected payoff for all three TFTTs and chooses TFTT # 2 at random. It is assumed that $i^* \geq 5$, i.e., that τ and ω are such that at least five PD games are played between Rip and a TFTT (or AllC) player before the TFTT (or AllC) player refuses Rip.

FIG. 10(a): Individual Ecology Payoffs for the IPD/CR Game with $\tau=0.5$

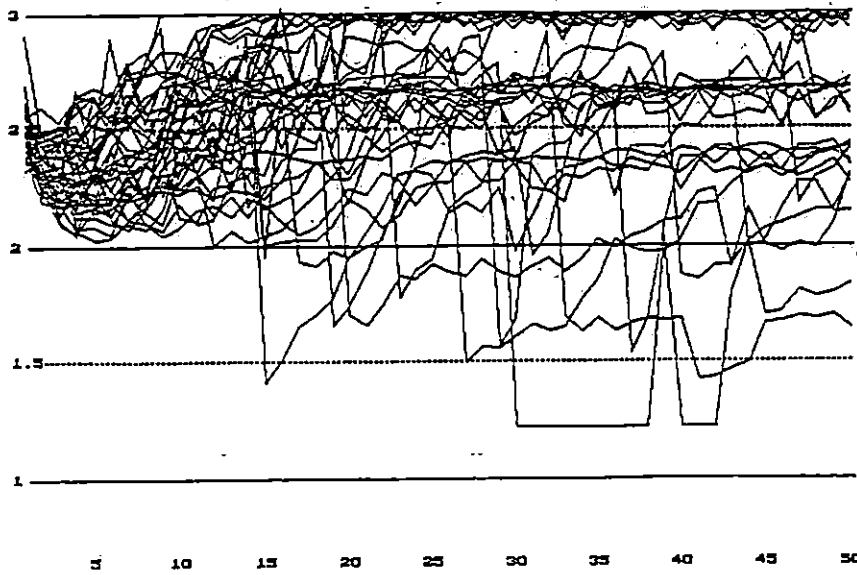


FIG. 10(b): Average Ecology Payoffs for the IPD/CR Game with $\tau=0.5$

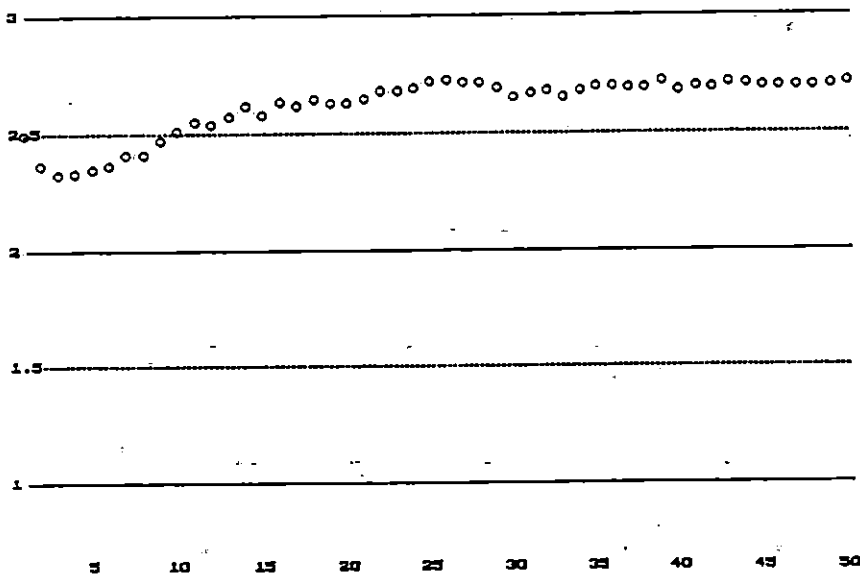


FIG. 10(c): Play Behavior Fractions for the IPD/CR Game with $\tau=0.5$

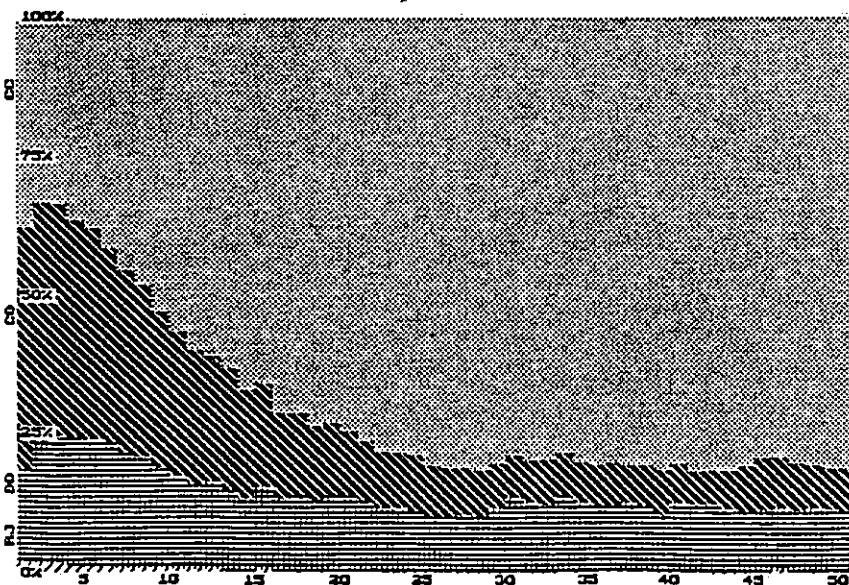


FIG. 11(a): Individual Ecology Payoffs for the IPD/CR Game with $\tau=1.5$

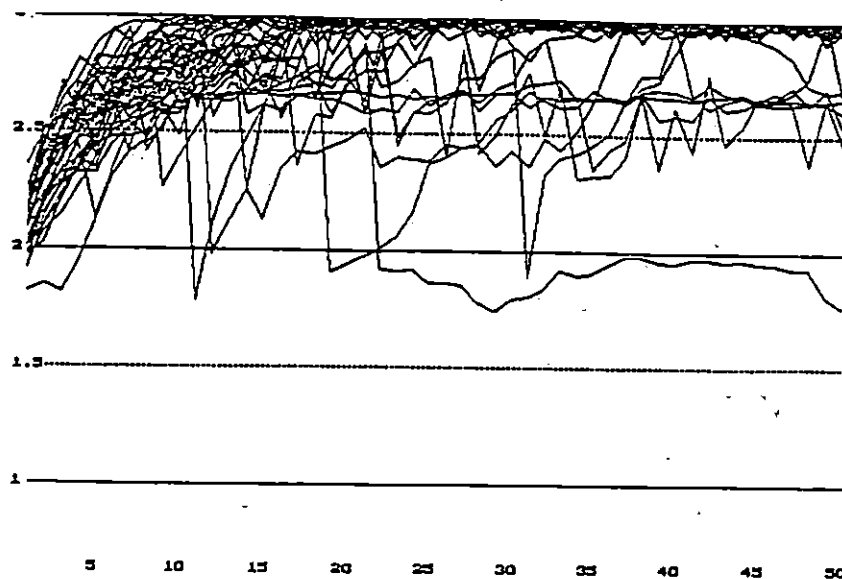


FIG. 11(b): Average Ecology Payoffs for the IPD/CR Game with $\tau=1.5$

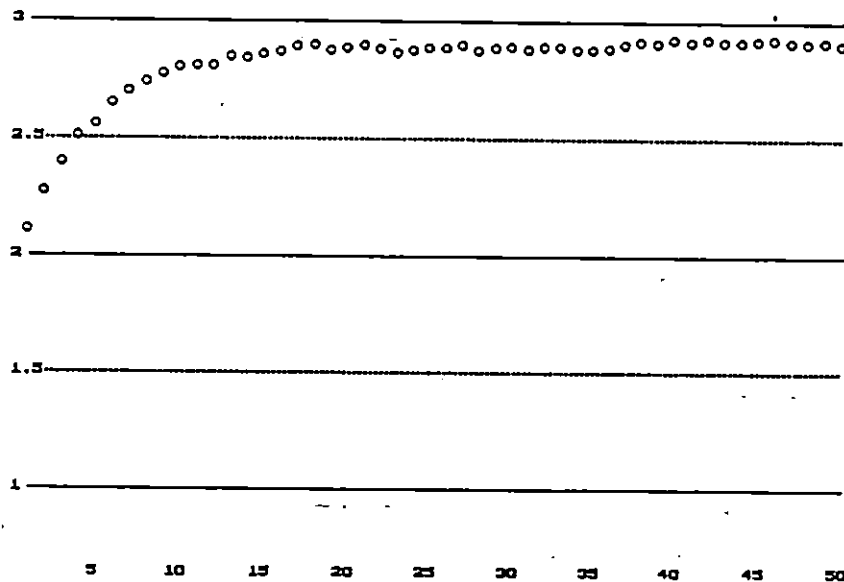


FIG. 11(c): Play Behavior Fractions for the IPD/CR Game with $\tau=1.5$

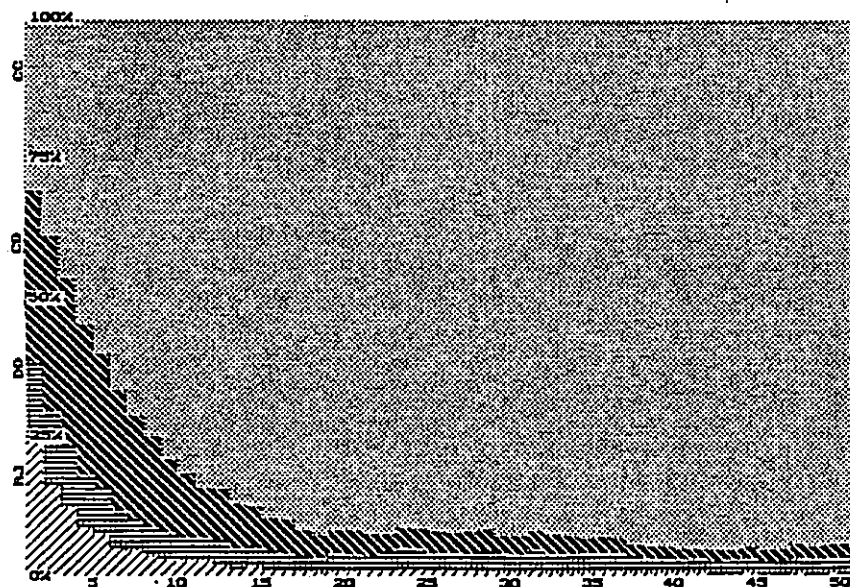


FIG. 12(a): Individual Ecology Payoffs for the IPD/CR Game with $\tau=1.6$

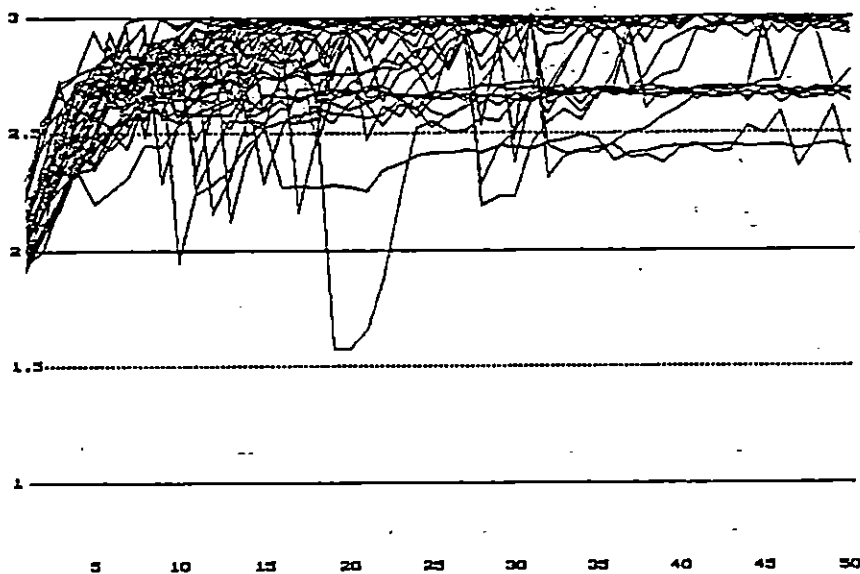


FIG. 12(b): Average Ecology Payoffs for the IPD/CR Game with $\tau=1.6$

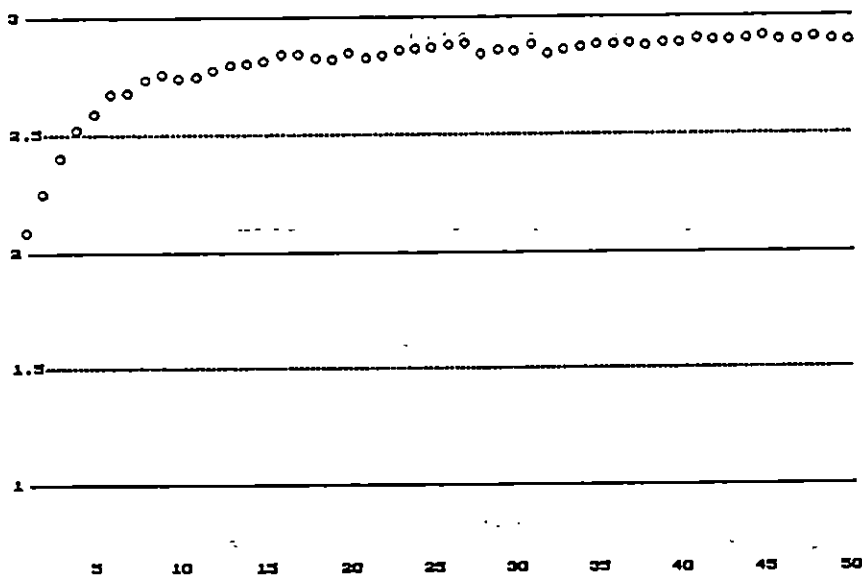


FIG. 12(c): Play Behavior Fractions for the IPD/CR Game with $\tau=1.6$

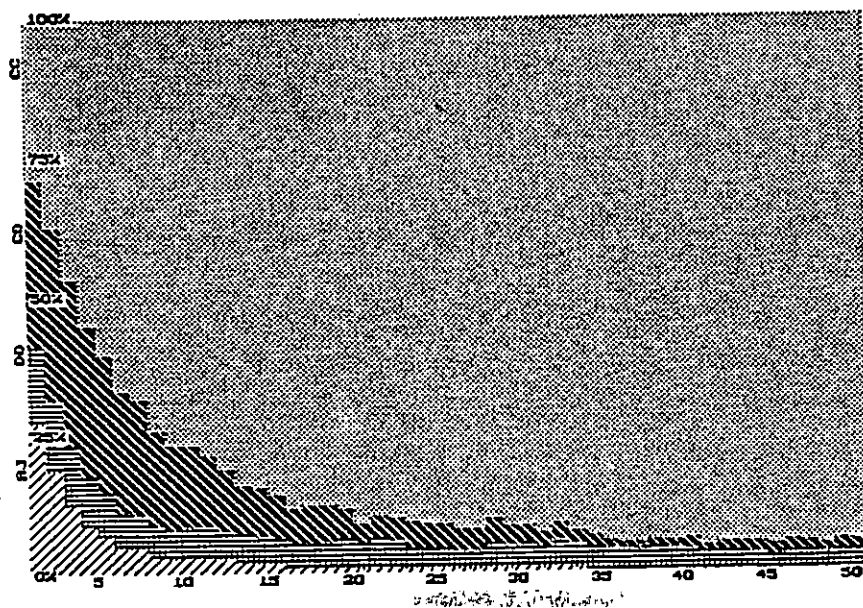


FIG. 13(a): Individual Ecology Payoffs for the IPD/CR Game with $\tau=2.4$

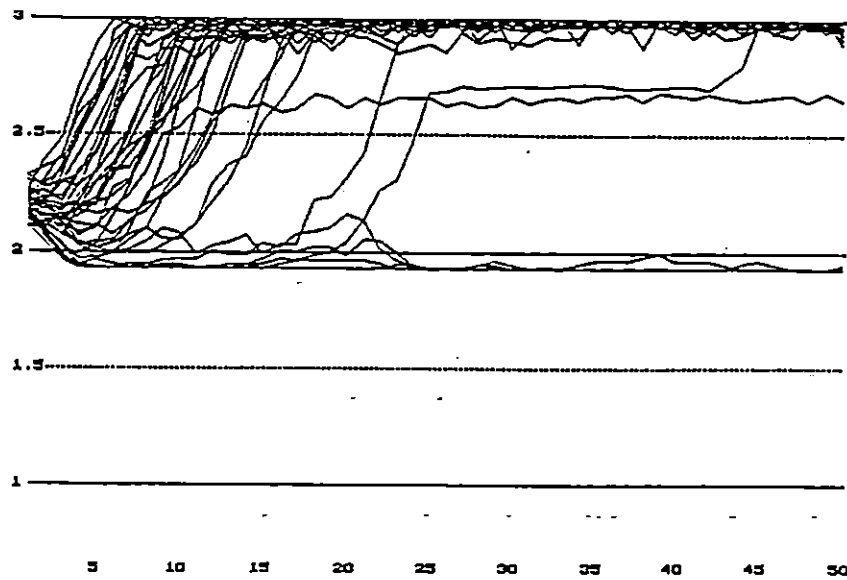


FIG. 13(b): Average Ecology Payoffs for the IPD/CR Game with $\tau=2.4$

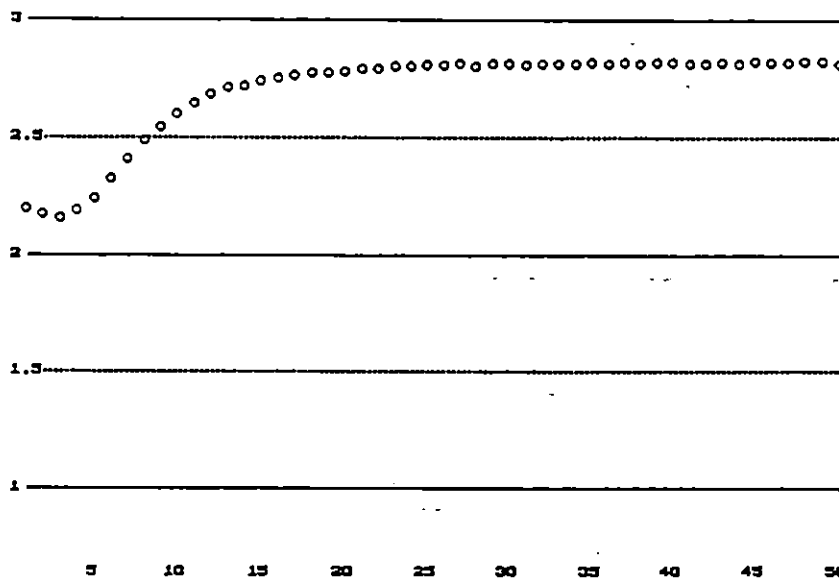


FIG. 13(c): Play Behavior Fractions for the IPD/CR Game with $\tau=2.4$

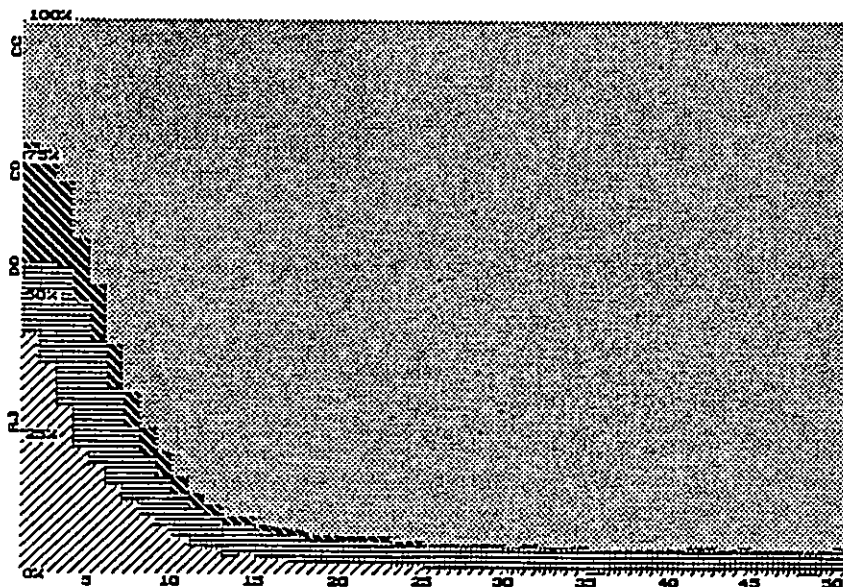


FIG. 14(a): Individual Ecology Payoffs for the Wallflower Band, $\tau=2.4$

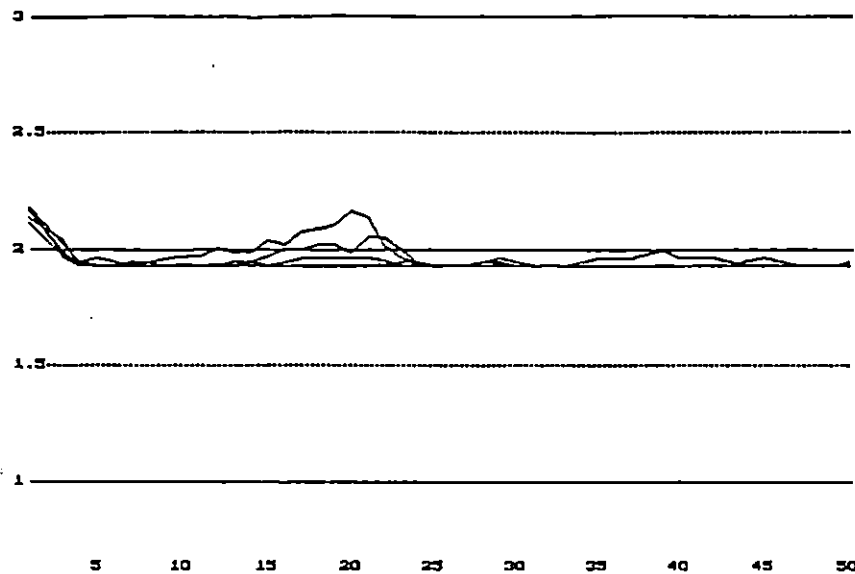


FIG. 14(b): Average Ecology Payoffs for the Wallflower Band, $\tau=2.4$

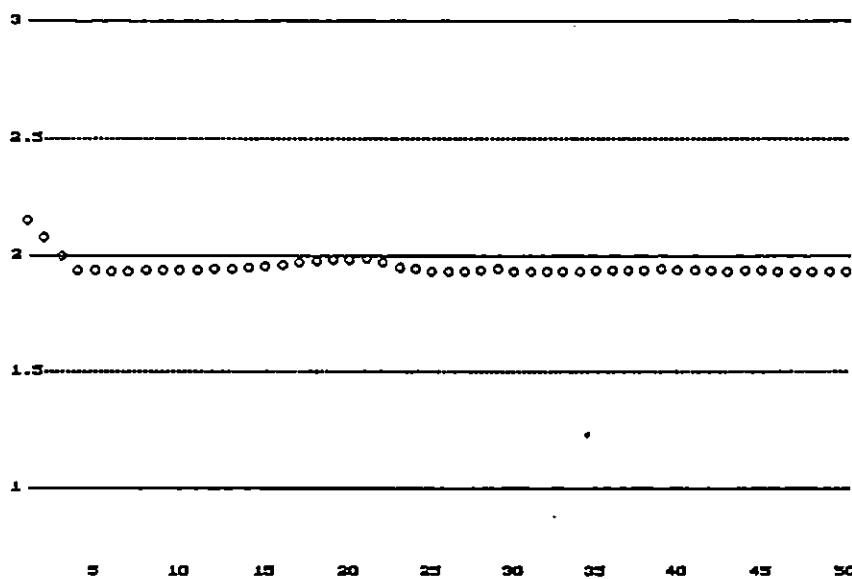


FIG. 14(c): Play Behavior Fractions for the Wallflower Band, $\tau=2.4$

