Evolving Aspirations and Cooperation*

Rajeeva Karandikar

Indian Statistical Institute, New Delhi 110 016, India

Dilip Mookherjee

Boston University, Boston, Massachusetss 02215

Debraj Ray

Boston University, Boston, Massachusetts 02215; and
Instituto de Análisis Económico (CSIC), Bellaterra 08193, Barcelona, Spain

and

Fernando Vega-Redondo

Facultad de Económicas and Instituto Valenciano de Investigaciones Económicas,
Universidad de Alicante, Alicante 03071, Spain

A 2×2 game is played repeatedly by two satisficing players. The game considered includes the Prisoner's Dilemma, as well as games of coordination and common interest. Each player has an *aspiration* at each date, and takes an action. The action is switched at the subsequent period only if the achieved payoff falls below aspirations; the switching probability depends on the shortfall. Aspirations are periodically updated according to payoff experience, but are occasionally subject to trembles. For sufficiently slow updating of aspirations and small tremble probability, it is shown that both players must ultimately cooperate most of the time. *Journal of Economic Literature* Classification Numbers C72, D83. © 1998 Academic Press

1. INTRODUCTION

We consider a 2×2 game played repeatedly by two satisficing players. The class of games considered is general enough to include the Prisoner's Dilemma, as well as 2×2 symmetric games of coordination and common interest.

Each player has an *aspiration* at each date, and takes an action. The action is switched at the subsequent period only if the achieved payoff falls below the aspiration level, with a probability that depends on the shortfall. Aspirations are updated in each period, depending on the divergence of achieved payoffs from aspirations in the previous period.

The aspiration-based process exhibits two specific features: (i) *inertia*: every action is repeated with at least a certain probability bounded away from zero; and (ii) *experimentation*: with a small probability, aspiration levels experience small random trembles around the going aspiration, thus preventing them from being perpetually satisfied with any given action. We examine the long run outcomes that are induced by vanishingly small tremble probabilities. This paper therefore builds on [3], in which a model of consistent aspirations-based learning was introduced.[1]

We make precise and prove the following result (Theorem 1 below):

If the speed of updating aspiration levels is sufficiently slow, *then the outcome in the long run must involve both players cooperating most of the time.*

The model therefore describes an adaptive learning process where individuals not only cooperate, but play strictly dominated strategies of the stage game for most of the time. While players may (and occasionally do) profit by deviating from cooperative behavior, the dynamics of the process ultimately lead back to mutual cooperation. It should be stressed that the result does not rest on a conventional repeated game argument (i.e., based on players' rational responses to others' strategies of conditional cooperation), since players exhibit a form of nonstrategic behavior which is entirely myopic. Morever, cooperative behavior emerges in the long run, irrespective of initial conditions.

Section 2 introduces the model, and describes some of its preliminary properties. Section 3 presents the main results. Section 4 provides an informal discussion of the reasoning underlying the main results. Section 5 discusses related literature, Section 6 concludes, and Section 7 collects all proofs.

---

[1] That paper did not consider the updating of aspirations within a game. For discussion of this and related literature, see Section 5.

2. THE MODEL

Consider the following 2×2 game:

	C	D
C	(σ, σ)	$(0, \theta)$
D	$(\theta, 0)$	(δ, δ)

where $\sigma > \delta > 0$, and $0 \leqslant \theta \neq \delta$.[2] If $\theta > \sigma$, this is a Prisoner's Dilemma.[3] If $\theta < \sigma$, this is a game of common interest. The case $\theta = 0$ corresponds to a game of pure coordination, where (C, C) and (D, D) are both Nash equilibria.

Player 1's state at date t is given by an action $A_t \in \{C, D\}$, and an aspiration level α_t, which is a real number. The corresponding objects for player 2 are given by B_t and β_t. A *state s* is the pair made up of player 1's state and player 2's state. Thus at date t, $s_t = (A_t, \alpha_t; B_t, \beta_t)$. The state at date t determines the actions chosen and hence the payoffs π_t^1 and π_t^2 for the two players. We now describe the *updating* of each player's state.

Consider the rule followed by player 1; an analogous rule is followed by player 2. There are two features. First, *actions are updated* as follows: if $\pi_t^1 \geqslant \alpha_t$, then player 1 is *satisfied*, and $A_t = A_{t+1}$. Otherwise player 1 is *disappointed*, the action is switched ($A_{t+1} \neq A_t$) with probability $1 - p$, where p is an indicator of *inertia*. It is assumed that p is a nonincreasing function of the extent of disappointment ($\alpha_t - \pi_t^1$), satisfying:

1. $p = 1$ if $\alpha_t - \pi_t^1 \leqslant 0$,

2. $p \in (\tilde{p}, 1)$ otherwise, for some $\tilde{p} \in (0, 1)$, and

3. p is continuous and the rate at which it falls is bounded, i.e., for all $x > 0$, $1 - p(x) \leqslant Mx$ for some $M < \infty$.[4]

Figure 1 describes p. In words, for any given positive degree of disappointment, the player will switch his action with positive probability.

---

[2] Our main result does not hold in the case where $\theta = \delta$, for reasons explained further in Section 4.

[3] Note however that no condition has been imposed on the relative values of δ and θ, so the model allows for the case in which alternation between (C, D) and (D, C) yields a higher payoff than (C, C).

[4] This last set of conditions is inessential for our main result, though employed in the proof. As will be discussed in Section 6, modified arguments apply in the case where the inertial probability is discontinuous at the point of zero disappointment, i.e., is bounded away from 1 as well as 0 for any positive disappointment.
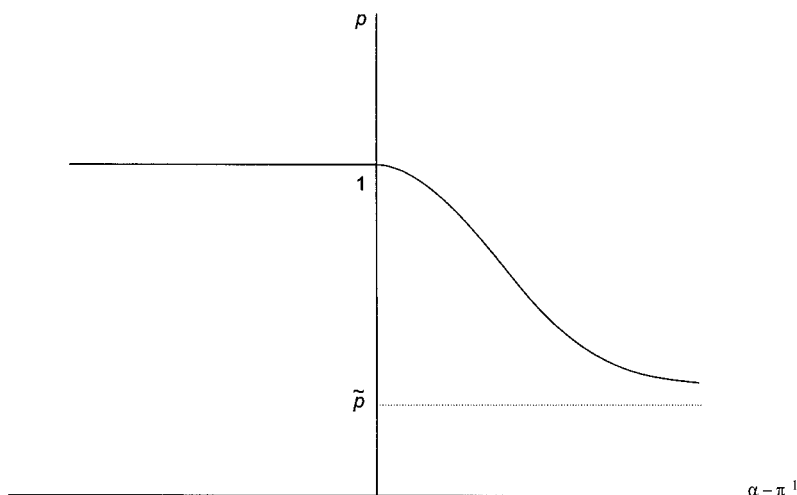
FIG. 1. The function p.

However, owing to inertia the probability of *not* switching is bounded away from zero.

Second, with respect to the *updating of aspirations*, it is convenient to first consider the case without any "trembles".

2.1. *The Model without Trembles*

Aspirations are updated as an average of the aspiration level and the achieved payoff at the previous play. For player 1, this yields

$$\alpha_{t+1} = \lambda\alpha_t + (1-\lambda)\,\pi_t^1, \tag{1}$$

where $\lambda \in (0, 1)$ may be thought of as a *persistence* parameter, assumed equal for both players for simplicity. A parallel equation applies to player 2.

These updating rules (i.e., without trembles) define a Markov process over the state space, which we may identify with the set $E \equiv \{C, D\}^2 \times \mathbb{R}^2$. This process will be denoted P, and will be referred to as the *untrembled process*.

Given any action pair (A, B), let the corresponding *pure strategy state* (pss) refer to the state where this action pair is played with aspiration levels exactly equal to the achieved payoffs: $\alpha = \pi^1(A, B)$, $\beta = \pi^2(A, B)$. It is clear that every pss is an absorbing state of the untrembled process: if players are satisfied with the payoffs they receive in an ongoing action pair, they have no reason to alter their actions or aspirations. Indeed, it is for this very reason that it is necessary to explore the robustness of any absorbing state by admitting the possibility of perturbations.

Before proceeding to the case of trembles, however, it is useful to settle a preliminary question: does the untrembled process always converge? This is addressed in

PROPOSITION 1.   *From any given initial state, the untrembled process P converges almost surely to some pure strategy state.*

This result is discussed in Section 4.

2.2. *The Model with Trembles*

While the untrembled process always converges to some pss, one suspects that some of these may not be robust to small perturbations in a player's state. To model such phenomena, we introduce trembles in the formation of aspirations.[5]

With probability $1 - \eta$, aspirations are formed according to the deterministic rule (1), while with the remaining probability η, the updated deterministic aspiration α is perturbed according to some density $g(., \alpha)$. Assume, again for simplicity, that η is the same for both players.

Informally, we would like small perturbations on either side of α to be possible, but at the same time, uninteresting technical complexities would be introduced by allowing aspirations to wander too far from the payoffs of the game, and we want to avoid these. Assume, then, that there exists some compact interval Λ which contains all feasible payoffs such that for each $\alpha \in \Lambda$, the support of $g(. \,|\alpha)$ is contained in Λ, and moreover, that $g(\alpha' \mid \alpha) > 0$ for all α' in some nondegenerate interval around α (relative to Λ). Furthermore, suppose that g is continuous as a function of α. Finally, assume that all initial aspiration vectors lie in the compact region Λ^2, and that all perturbations are independent over time and across players.[6]

Denote the resulting stochastic process by P^η. A standard theorem (see, e.g., [14, Theorem 16.2.5]) guarantees that the process has a well defined long run outcome:

PROPOSITION 2.   *For each $\eta > 0$, the process P^η converges (strongly) to a unique limit distribution μ^η, irrespective of the initial state.*

---

[5] Starting with any pss, a small upward push to a player's aspiration level will cause that player to be disappointed, and hence induced to experiment with other actions. An alternative modeling approach would directly allow experimentation with different actions. We suspect that the results would be the same in such an approach.

[6] As usual, a similar definition holds for player 2. It is immaterial to the argument whether the function g is the same for both players.

3. MAIN RESULTS

By Proposition 3, the introduction of trembles serves to single out a unique (though probabilistic) long-run outcome. Obviously, one is interested in the nature of the long run distribution μ^η when the tremble probability η is close to zero, since this is likely to yield a selection from the multiple long-run limits of the untrembled process.

Some preliminary steps are needed before we can state such a result precisely. To begin with, one needs to ensure that the sequence of long run distributions μ^η settles down as η goes to zero:

PROPOSITION 3. *The sequence of distributions μ^η converges weakly to a distribution μ^* on E as $\eta \downarrow 0$.*

Proposition 3 is a corollary of a general observation on the long run behavior of Markov processes subjected to small stochastic perturbations, which also provides a precise characterization of the limiting distribution. Because this result may be of wider interest than the specific application studied here, we provide a self-contained statement and proof of this general theorem (Theorem 2) in Section 7. Theorem 2 also yields, as a corollary, a description of the limit distribution μ^*, which we now state.

Use Q_i to denote the one step transition probability of the stochastic process, conditional on the situation where *only* player i's aspiration is subjected to a tremble. Let Q denote $\frac{1}{2}(Q_1 + Q_2)$. We may interpret this as the transition rule when exactly one player trembles, with both players being equally likely to tremble.

Let R denote P^∞, the infinite step transition rule in the untrembled process. This is well-defined by Proposition 1. Finally, let QR denote the composition of Q and R. In words, the process QR refers to the effect of subjecting exactly one player (chosen randomly) to a tremble in her aspirations, followed by the untrembled process thereafter for ever.

PROPOSITION 4. *The limiting distribution μ^* is the unique invariant distribution of the process QR.*

By Proposition 1, the untrembled process converges to a pss. It follows that an invariant distribution of QR must be concentrated on the pss's. Proposition 4 says that QR has a unique invariant distribution, which is precisely the limit of the invariant distributions corresponding to vanishing tremble probabilities.

In words, the selected long run outcome can be obtained as the unique long run outcome of an "artificial" Markov process defined only over the four pss's, with the transition probability between pss's obtained as follows: Starting with any pss, subject one player chosen randomly to a single tremble

in aspiration to obtain a new state, from which the untrembled process is left to operate thereafter to arrive eventually at some pss.

Why do we need a precise characterization of the limit of the long run distributions, unlike the work of previous authors? The reason is that the long run distribution will *not* generally be concentrated on a single pss. This is in sharp contrast with random matching contexts considered by [12] and [24], where the corresponding process singles out a unique limit state.

In the informal discussion, which we postpone to Section 4, we attempt to provide a clearer intuitive explanation of this observation.

We are now in a position to state our main result. Let $\mathscr{C}$ denote the singleton set consisting of the (C, C) pss.

THEOREM 1. *The weight $\mu^*(\mathscr{C})$ placed by the limiting distribution μ^* on the mutual cooperation pure strategy state is close to* 1, *for persistence parameter λ sufficiently close to* 1. *Formally,*

$$\lim_{\lambda \to 1} \mu^*(\mathscr{C}) = 1.$$

4. INFORMAL DISCUSSION

The assertions thus far may be summarized as follows. First, the untrembled joint process of aspirations and actions always converges to a pure strategy state. Second, the trembled process is ergodic and converges to a unique invariant distribution. Third, these invariant distributions (viewed as functions of the tremble) themselves settle down to a "limit" invariant distribution as the tremble probability approaches zero. Finally —and this is the main result —the limit invariant distribution places almost all weight on the cooperative outcome, provided that the persistence parameter is sufficiently close to unity.

We discuss these informally in turn.

(i) Begin with the *convergence of the untrembled process*. The formal proof of Proposition 1, in the appendix, takes the easiest route towards establishing convergence, by exploiting a degree of inertia in the model that is built in by assumption. Specifically, given any state, there can be an infinite run on the current action pair, which would cause aspirations to converge to the corresponding payoffs.

To illustrate this, consider action pair (C, D) in the Prisoner's Dilemma where $\theta > \delta$, and suppose that initial aspirations of both players lie anywhere in the interior of the feasible payoff region. What is the probability of an infinite run on (C, D) thereafter in the untrembled process, which

would result in convergence to the (C, D) pss? Along such a path, player 2 would have no cause to switch away from D, so what is needed is for player 1 to stick with C perpetually despite being disappointed at every stage. At stage t, player 1's aspiration and hence disappointment level would be $\lambda^t \alpha$, if α denotes her initial aspiration. Hence the probability of converging to (C, D) is $\prod_{t=1}^{\infty} p(\lambda^t \alpha)$, which is positive if and only if

$$\sum_{t=1}^{\infty} [1 - p(\lambda^t \alpha)] < \infty,$$

a condition which is satisfied, since the left hand side equals

$$\sum_{t=1}^{\infty} [p(0) - p(\lambda^t \alpha)] \leqslant \frac{M\lambda\alpha}{1 - \lambda}.$$

A similar argument can be given for the possibility of an infinite run on any action pair, and any initial set of aspirations, as detailed in the proof of Proposition 1. Note that no assumption has been made concerning the speed at which aspirations are updated; the result follows solely from our assumptions regarding the nature of inertia. Nevertheless, the formulation of inertia can be substantially weakened without affecting this result (see Section 6).

(ii) Now turn to the *ergodicity of the trembled process*, as described in Proposition 2. This is familiar by now in the literature on stochastic evolution. What is somewhat unusual, however, is that as trembles vanish the limit invariant distribution places weight on more than one pure strategy state. Specifically, it is possible to transit from any pss to any other *with one tremble*, so that one pss cannot become infinitely more likely than another as the tremble probability vanishes. For instance, consult Fig. 2 which depicts the Prisoners Dilemma. Here, a tremble (represented by a dotted path) perturbs the aspiration of player 2 from the cooperative payoff (σ, σ), following which the untrembled process generates a "long aspiration cycle" (shown by the curved arrow) that ultimately converges to the defection payoffs from below. Of course, the reverse transition from the (D, D) pss to the (C, C) pss is also possible, as shown in the same figure by a dotted tremble from (δ, δ) followed by the straight arrow.

(iii) These observations imply that it is no longer possible to characterize long run outcomes by studying the support of the limit invariant distribution: we need to understand the *relative* weights on the different pss's. To get a handle on which pure strategy state is likely to receive the lion's share of probability weight, we must deduce an *explicit formula for the limit invariant distribution*. Proposition 4 specifically states that this limit is the same as the limit distribution of the process QR. The proof of this
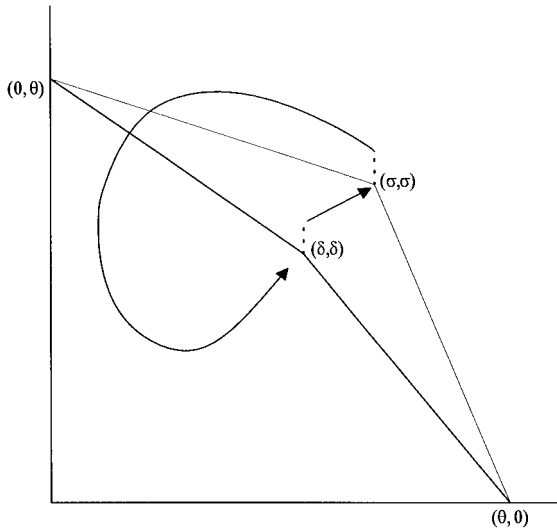
FIG. 2. Possible convergence to (D, D) from perturbed (C, C).

proposition (and that of Proposition 3) relies on Theorem 2, a technical result concerning perturbed Markov processes, which we state and prove in the appendix.

(iv) Given Proposition 4, we need thereafter to focus on the process QR. As discussed following Proposition 4, the unique limit distribution of QR must be concentrated on the four pss's. To prove Theorem 1, therefore, we need to characterize *the relative probabilities of transiting from any pss to any other*, *when the aspiration of one player is trembled once, and the untrembled process operates perpetually thereafter*. Specifically, it suffices to prove that as $\lambda \to 1$, transitions from the (C, C) pss from the other three pss's in the QR process occur with probability approaching 0, while the likelihood of the reverse transitions remains bounded away from 0.

Since the proof of this is somewhat long and involved, it will help to outline the various steps in the argument. *The case of a game of pure coordination* $(\theta = 0)$ is the easiest to consider, and is illustrated in Fig. 3. Note first that if aspirations lie in the region I where player 2 has aspiration lying between δ and σ, then if λ is close enough to one the untrembled process will converge to the mutual cooperation pss with probability close to one. The reason is that (a) if aspirations start in the interior of I, then for λ close enough to one, they will continue to lie within I for a large number of subsequent dates; (b) at any time during this period, the probability that (C, C) will be played at least once within the next two dates is bounded
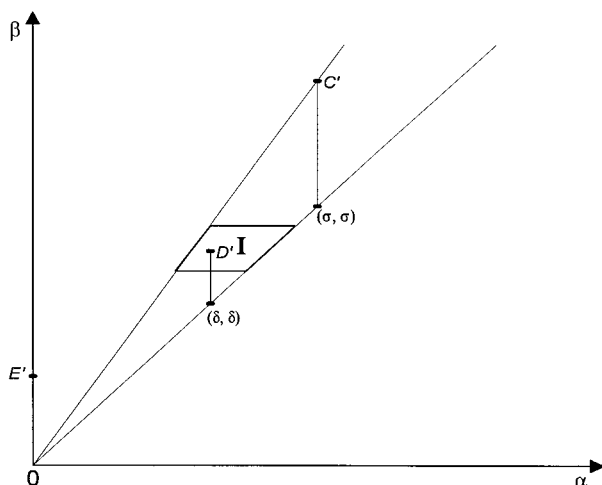
FIG. 3. Cooperation in the game of pure coordination.

away from zero;[7] (c) once (C, C) is played while aspirations are still in the region I, both players are satisfied, and will have no incentive to switch. Thereafter, (C, C) will be repeated forever, causing the untrembled process to converge to the mutual cooperation pss.

Now consider the dynamic of the process QR. If we start at the mutual cooperation pss, and subject one player's aspiration to an upward tremble, say player 2, the new aspiration vector is at point C' in Fig. 3. If the untrembled process runs thereafter, aspirations must lie in the convex hull of the point C', and the pure strategy payoff points. Morever, if aspirations are updated slowly enough, then the process can converge to any pss *not* involving mutual cooperation, only if they pass through the interior of the intermediate region I. But by the argument of the preceding paragraph, if aspirations are ever in the interior of I, they must converge back to the mutual cooperation pss with probability close to one (if λ is close enough to one). In other words, the probability of transiting from the mutual cooperation pss to any other in the QR process converges to 0 as λ approaches 1.

[7] If (D, C) or (C, D) is played today, then both players are disappointed by a discrete amount. The player selecting C today will not switch with some probability (at least $\tilde{p}$) owing to inertia, while the other player will also switch to C with some probability that is independent of λ, resulting in a play of (C, C). And if (D, D) is played today the second player is disappointed by at least a certain amount, causing her to switch to C at the following date, while the first player may continue to play D. Hence with positive probability (D, D) will be followed by (D, C), which in turn may be followed by (C, C) as in the preceding argument.

The reverse transitions, however, continue to remain probable in the QR process. Starting from the mutual defection pss, player 2's aspiration could undergo an upward tremble, taking aspirations to point D', in the interior of the region I. Then the untrembled process would move to the mutual cooperation pss with probability close to one. Similarly, starting from any of the asymmetric pss's where the two players fail to coordinate, the player selecting D could experience an upward tremble to her aspiration, taking aspirations from $(0, 0)$ to the point E' in Fig. 3. This would cause (C, C) to be played, following which the untrembled process would converge to the mutual cooperation pss. This completes the argument for the pure coordination game with $\theta = 0$.

When θ is positive, the argument becomes substantially more complicated, owing to the fact that plays of the asymmetric pss's can then serve to raise one player's aspiration level. This is best illustrated in the alternative case of the Prisoners Dilemma, where $\theta > \sigma$. Starting with the mutual cooperation pss, if player 2 receives a boost in aspiration, it causes her to switch to D. This results in a play of (C, D), which serves to raise 2's aspiration even further. Of course, player 1 will be disappointed and will switch to D, but will be disappointed with the resulting play of (D, D) as well, and tend to switch back to C. If player 2 has stuck to D, then another play of (D, C) results, which again boosts her aspiration. It is possible therefore for a large number of plays of (D, C) to occur, which serve to raise player 2's aspiration, and lower player 1's. It is then possible for the aspirations to wind around, as depicted by the curved arrow in Fig. 2, and converge to the mutual defection pss from below, following one perturbation of the mutual cooperation pss.

Nevertheless, Theorem 1 applies to the Prisoners Dilemma, and extends to values of θ intermediate between σ and δ as well. The main steps in the reasoning (for the case of the Prisoners Dilemma) are as follows. First, as Lemma 3 in Section 7 shows, if a player (1, say) has an aspiration smaller than his maxmin payoff of δ, then his aspiration will drift upward with probability close to 1 (if λ is close enough to 1). The reason for this is simple: either the two players continually cooperate (in which case 1's aspiration must increase steadily), or sooner or later player 1 will select his maxmin action D. In the latter case, 1 will stick to action D, being satisfied with the maxmin payoff δ which exceeds the current aspiration level. His aspirations must then drift upwards. Intuitively, players learn to aspire to payoff levels that are at least as large as their maxmin payoffs.

Lemma 5 in Section 7 then shows that this upward drift in players' aspirations continues to hold when aspiration levels are approximately equal to the maxmin payoff δ. Specifically, if aspirations start within region N in Fig. 4, then player 1's aspiration must continue to increase beyond δ, while player 2's aspirations stay intermediate between δ and σ. The intuition
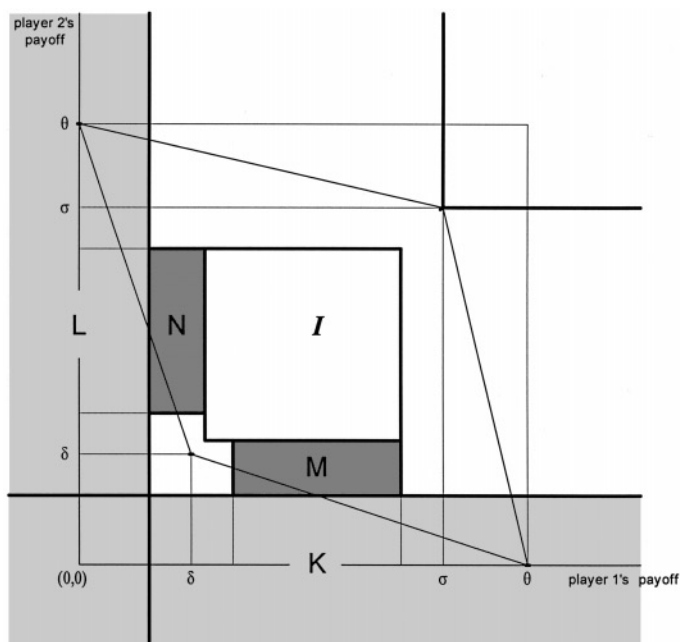
FIG. 4. Steps in the proof of Theorem 1: The Prisoner's Dilemma.

for this is that the dynamics are continuous at an aspiration level of δ for player 1, so the upward drift that occurs when 1's aspiration is less than δ, continues to extend at the point when his aspirations are approximately at δ. However, the proof for this is long and delicate.

The next step is to consider what happens when the aspirations of both players are intermediate, lying between δ and σ. Using a logic similar to the case of a pure coordination game, Lemma 2 shows that if λ is close to one, the untrembled process converges thereafter to the mutual cooperation pss with probability approaching one. Intuitively, any play of (D, D) will leave both players disappointed, inducing both to switch simultaneously to C, thus resulting in a play of (C, C), from which point the process converges to the mutual cooperation pss. And any play of (D, C) or (C, D) will tend to be followed by (D, D), and thereafter by (C, C).

Lemma 6 then uses these results to prove that starting with the mutual cooperation pss, the process QR will move back to this pss with probability approaching one as $\lambda \to 1$. Hence, paths such as the curved arrow in Fig. 2 by which the process moves from the cooperation pss to the defection pss following one tremble, become increasingly improbable. The geometric structure of the argument is described in Fig. 4. In this figure, region I is a rectangle containing aspirations that are intermediate between

δ and σ for both players. Region N satisfies the same property for player 2 as region I does, but only contains aspirations for player 1 that lie in some neighborhood around δ. Finally, region L contains all aspiration vectors such that player 1's aspiration falls short of his lowest aspiration in N. (Regions M and K are defined analogously to N and L respectively.)

The area comprising the union of the regions L, N, I, M and K acts as a reflecting barrier (with high probability) as $\lambda \to 1$. By Lemmas 2 and 5, regions N, M and I cause aspirations to revert back to (σ, σ), whereas by Lemma 3, neither regions L nor K can be penetrated from outside. It follows that the probability of aspirations converging to $(0, \theta)$, $(\theta, 0)$, or (δ, δ), following a single tremble from the (C, C) pss, approaches zero as $\lambda \to 1$. On the other hand, the reverse transition from the (D, D) pss to the (C, C) pss becomes increasingly likely since a single tremble takes aspirations from (δ, δ) into the sets N or M with positive probability, following which Lemma 5 can be applied.[8]

Similarly, starting from any neighborhood of $(0, \theta)$, aspirations must drift "rightwards", and thereafter out of the set L. The likelihood of converging to $(\theta, 0)$ is close to zero, since region K cannot be penetrated from outside. Hence starting from the (D, D) or the (C, C) pss the process QR will almost surely converge to either the (D, D) or (C, C) pss, implying that the weight on both asymmetric pss's in the invariant distribution of QR must approach zero. Combining this with the earlier argument, it follows that almost all the weight of the limit invariant distribution must be on the (C, C) pss.

Intuitively, transitions following a single perturbation of the (C, C) pss to the other pss's must *necessarily* require one or both players to not switch their actions for long stretches of time, even if they are disappointed. As aspirations adjust more and more slowly, the bouts of inertia required become indefinitely large, and therefore increasingly improbable.[9] On the

---

[8] This argument does not apply to the case when θ exactly equals δ. Then starting from the mutual defection pss, a perturbation of exactly one player's aspiration level will cause the process to revert back to the defection pss with probability 1. Hence in this case almost all the weight in the limiting distribution is placed on the mutual defection pss, instead of the mutual cooperation pss. The reason is that if player 2's aspiration goes above δ, for instance, inducing her to switch to C, such a switch will not affect player 1's aspiration at all, since $\theta = \delta$. So player 1 will continue to stick to D. And as long as she does so, she obtains a payoff of δ, no matter what player 2 does. Hence player 1 can never deviate away from the action D, implying that player 2 must return to D as well. However, as our proof shows, this argument is inapplicable as long as θ is even slightly different from δ, for that causes a divergence between payoffs and aspiration for player 1 at some date, motivating a switch away from the action D, provided λ is close enough to 1. The degenerate case also indicates that the basic arguments for the cases $\theta > \delta$ and $\theta < \delta$ are qualitatively different from each other.

[9] If the inertia function p were to be discontinuous at 0, then the probability of converging to any other pss from the cooperative pss following one tremble is obviously zero, so the result is straightforward in that case.

other hand, the transition to the (C, C) pss from the other pss's do not rely at all on inertia, and are thus infinitely more likely relative to the reverse transitions.

Note in conclusion that even for small tremble probabilities and speed of aspiration updating, the process does *not* converge to mutual cooperation in the long run. Cooperation simply becomes statistically dominant. Players perpetually oscillate between different action pairs, as their aspirations are occasionally subjected to trembles in different directions, with effects that last beyond the trembles. In particular, these trembles cause them to experiment with actions different from those used in the recent past. For instance, starting with mutual cooperation, such trembles induce experimentation with defection, which leads to transitory gains from exploiting the cooperative partner. However, the partner is dissatisfied in such situations, and will switch to defection as well, which serves to "punish" the initial defector. A period of mutual defection then ensues, which tends to disappoint *both* players, inducing an eventual return to the cooperative phase. Even though, along this process, the players often play a dominated strategy, this is *not* due to the fact that the learning process *per se* makes them stick to an inferior action (e.g., it is not that they would zero in on "defect" in the Prisoner's Dilemma when facing a fixed strategy on the part of the opponent). Instead, it is the interaction between the learning dynamics of both players which leads to this possibility.

The above is reminiscent of how cooperation may be sustained in a repeated game. It does seem that aspirations-based learning leads to a manner of play in which trigger-based punishments appear to be used. However, the establishment of a precise connection is beyond the scope of this paper. It should be mentioned, nevertheless, that we obtain this similarity despite the fact that players are only "learning" to play the stage game. This is in contrast to some recent literature such as [11], in which convergence to an equilibrium of the *repeated* game is considered.

5. RELATED LITERATURE

The model of adaptive behavior considered in this paper presumes a limited form of rationality, where players need not know the structure of the game, or the opponents' previous actions; nor do they have to be able to solve maximization problems. The notion of "learning", if any, does not reflect the acquisition of any new knowledge *per se*; instead it concerns the adjustment of aspirations or payoff expectations on the basis of past experience, which in turn shapes agent behavior in the face of current experience. In the terminology of [21], such models represent "stimulus"

or "reinforcement" rather than "belief" learning. They originated in the mathematical psychology literature [7], and have received a certain degree of support in laboratory experiment situations involving human subjects [15, 16, 20, 22, 23]. Other recent explorations of such models of learning include [4] and [6], both of which explore the relationship with "replicator dynamics" models, and [10], which develops an axiomatic "case-based decision theory" where players satisfice relative to aspiration levels that are based on past experience.

The structure of interaction between players in our model does not correspond to random matching of pairs selected from a certain population, as in the evolutionary literature (see, for instance, [4, 12, 24]). In our context, a *given* pair of agents plays the game repeatedly over time. This also stands in contrast with some of the well-known models whose concern is to provide an evolutionary basis for the rise of cooperation, e.g., [4, 8, 19]. These papers consider situations in which the *repeated* Prisoner's Dilemma is recurrently being played between pairs of individuals randomly selected from the general population. In a sense, the objective is to select among multiple equilibria of the underlying supergame by embedding it in a wider intertemporal framework. Instead, our approach remains within the scenario of a single indefinite repetition of the stage game, singling out the *stage* outcome which happens to be played most of the time in the long run.

[3], a precursor to this paper, assumed aspiration levels were fixed, and analyzed models of reinforcement learning leading to long run outcomes consistent with the given aspiration levels. In the context of general two player repeated games, it was shown that such an equilibrium concept allowed individually rational Pareto-undominated pure strategy outcomes to be played in the long run. This paper extends this model by providing an explicit process by which aspiration levels as well as chosen actions evolve, but in the context of a specific class of 2×2 games. Aspirations do turn out to converge, and the long run outcome is essentially cooperative, thus complementing the analysis in that paper.

[13] and [18] both apply the Gilboa–Schmeidler case-based theory to games of coordination or the Prisoners Dilemma. They allow aspirations to evolve simultaneously with the strategies selected by players in a context of repeated interaction, and provide conditions under which long run outcomes entail cooperation. These conditions entail initial aspiration levels lying in prespecified ranges: for instance, Pazgal needs to assume that they are sufficiently high relative to the cooperative payoffs for both players, while Kim assumes that they lie above a level which is slightly below the cooperative payoff for both players. Our model in contrast predicts cooperation in the long run, irrespective of initial conditions. Another important difference is that [13] and [18] both assume that aspirations average

maximal experienced payoffs in past plays, whereas we assume they average the actual experienced payoffs. Hence the theory in these papers imparts a certain additional degree of "ambitiousness" to players, which helps in ensuring convergence to cooperative outcomes.

Finally, [9] and [17] consider models in which numerous pairs of players simultaneously play a given bilateral game, their aspirations being formed on the basis of the average payoff earned across the whole population. [9] postulates that the pair of players is repeatedly matched over time, and obtains the result that in the long run all players must cooperate. In [17], it is assumed instead that players are rematched every period, and demonstrate a tendency towards partial cooperation in such a case. In these papers cooperation is attained by virtue of the fact that non-cooperators become dissatisfied owing to the relative success experienced by *other* agents who manage to cooperate. In our paper, such mechanisms of "social learning" are absent.

6. EXTENSIONS AND CONCLUDING COMMENTS

The behavioral dynamic studied in this paper is admittedly stylized, and the context to which it is applied undoubtedly special. Simple behavior rules of the kind we have analyzed here are more plausible in settings where players are involved in games of greater complexity (e.g., where each player has a large number of actions to choose from). Considerations of tractability caused us to initially restrict attention to a simple class of 2×2 games, which should be viewed as the first step of a more general analysis that needs to be explored in future research. We believe, nevertheless, that some of the insights of this paper would extend to more general environments. We describe below possible alternative formulations of the following features of our model.

(a) *Inertia.* While our stated assumptions provide the easiest way to prove Proposition 1, the result also holds under weaker assumptions concerning the inertia function p, providing we restrict aspirations to not be updated too rapidly. For instance, if λ is close to 1, convergence is ensured by an alternative argument which does not rely on the inertia-based infinite runs on ongoing action pairs.[10] This argument works even

---

[10] Consider the Prisoner's Dilemma, for instance, and suppose, contrary to the assertion of the proposition, that the untrembled process does not converge. Then it can be shown that the process must wander infinitely often through the interior of the rectangle I depicted in Fig. 4. Using the argument of Lemma 2, this implies that the process must converge to the mutual cooperation pss.

when p falls at an unbounded rate near the point of zero dissatisfaction, or even if it is discontinuous at that point.

(b) *Experimentation.* A natural variant of our formulation of trembles would be directly with respect to the actions chosen, rather than aspiration levels. Since the effects of trembles in aspirations and action choices are similar, one would expect the main results to extend to such contexts.

(c) *Aspiration Updating.* One would also expect our results to extend to a more general formulation of adaptive aspiration revision rules, for instance, nonlinear ones. More complex is the question of what happens when aspirations are formed on the basis of simple time averages of past payoffs. Then *any* transition becomes progressively more lengthy as time proceeds, which could conceivably lead to non-ergodic behavior (even with trembles). However, the *relative* likelihoods of the different transitions across pss's would be similar to that described in our model.

(d) *Matching Rules.* A different direction for extension would be a finite-population context where players are randomly matched every period to play the Prisoner's Dilemma, say. We conjecture that the cooperative outcome (i.e., all players choosing C) would still be the unique long-run outcome for λ close to one. The reason is that, in this context, the destabilizing effect of one tremble would still be the same as before, once we are allowed to specify the particular outcome (or chain of outcomes) of the matching mechanism. Since every matching outcome has positive probability, so does any finite chain of matches required for the perturbation to operate in the desired direction.[11]

(e) *Behavior Rules.* Finally, one might allow players to select from a richer class of behavior rules, for instance, those permitting current action choices to be conditioned on the history of recent plays. Payoff experience from the use of different rules could be used to discriminate between them. Such a framework may shed light on the difficult issue of whether sophisticated rules may either arise or, at least, play a crucial role in facilitating the emergence of simple cooperative behavior.

7. PROOFS

Proof of Proposition 1. Given initial aspirations (α_0, β_0), aspirations at all later dates will be contained in the convex hull of (α_0, β_0) and the four (pure) payoff points of the game. Let this convex hull be denoted $\mathcal{H}$ and

---

[11] On the other hand, random matching within an infinite population may cause perpetual defection to be the unique long run outcome, as the probability of the appropriate matches shrinks to zero as the population grows indefinitely.

let the maximum aspirations for the two players in $\mathscr{H}$ be denoted by $\bar{\alpha}$ and $\bar{\beta}$ respectively.

Consider any state (A, α, B, β) with $(\alpha, \beta) \in \mathscr{H}$. Let the payoffs generated by the action pair (A, B) be denoted (π^1, π^2). Consider the probability of an infinite run on this action pair, which is given by

$$h(\alpha, \beta) \equiv \prod_{t=1}^{\infty} p(\lambda^t(\alpha - \pi^1)) \, p(\lambda^t(\beta - \pi^2)).$$

We claim that (i) $h(\alpha, \beta) > 0$ for every $(\alpha, \beta) \in \mathscr{H}$, and (ii) h is nonincreasing in each argument.

To prove (i), it suffices to check that:

$$\sum_{t=1}^{\infty} \left[1 - p(\lambda^t(\alpha - \pi^1)) \, p(\lambda^t(\beta - \pi^2)) \right] < \infty.$$

This condition is satisfied because the left hand side is bounded above by

$$M \frac{\lambda}{1 - \lambda} \left\{ |\alpha - \pi^1| + |\beta - \pi^2| \right\},$$

thus establishing (i). Claim (ii) follows directly from the fact that p is nonincreasing.

Given (i) and (ii), we see that for every $(\alpha, \beta) \in \mathscr{H}$,

$$h(\alpha, \beta) \geqslant h(\bar{\alpha}, \bar{\beta}) > 0.$$

Let $\varepsilon > 0$ denote the minimum value of $h(\bar{\alpha}, \bar{\beta})$ across all possible initial action pairs. It follows that at every date the probability of converging to the pure strategy state corresponding to the ongoing action pair is at least ε, thereby completing the proof. ∎

We now prepare for the statement and proof of Theorem 2. Let $\Delta(E)$ denote the set of probability measures on a compact metric state space E, endowed with the Borel σ-algebra. For any transition probability Q on E and any measure $\mu \in \Delta(E)$, define a measure μQ by $\mu Q(A) \equiv \int_E Q(x, A) \, \mu(dx)$ for any Borel set A. In this way, two transition probabilities P and Q naturally induce a third PQ, where $PQ(x, .) \equiv P(x, .) \, Q$. This permits us to define m-step (for $m \geqslant 2$) transition probabilities iteratively: $P^m \equiv P^{m-1}P$, where $P^1 \equiv P$.

Given a real-valued measurable function f on E and a transition probability Q, define the function $Qf(x) \equiv \int_E f(y) \, Q(x, dy)$. Q is said to have the *strong Feller* property if Qf is continuous for every bounded measurable function f.

A measure μ on E is *invariant with respect to P* if $\mu P = \mu$.

For each $\eta \in (0, 1)$, let Q_*^η be a transition probability. Let $\phi(\eta)$ and $\psi(\eta)$ be functions, with $0 < \phi(\eta) < 1$ and $0 \leqslant \psi(\eta) < 1$, and with $(\phi(\eta), \psi(\eta)) \to (0, 0)$ as $\eta \to 0$. Define

$$Q^\eta \equiv (1 - \psi(\eta))\, Q + \psi(\eta)\, Q_*^\eta,$$

and now perturb some given transition probability P in the following manner:

$$P^\eta \equiv (1 - \phi(\eta))\, P + \phi(\eta)\, Q^\eta$$

for all $\eta \in (0, 1)$.

The following lemma will be useful.

LEMMA 1. *Let Q be a strong Feller transition probability function on a compact metric space E. Let g^η be a family of measurable functions, uniformly bounded by K, such that for each $x \in E$, $g^\eta(x) \to g(x)$ as η tends to 0.*

Then

$$\lim_{\eta \to 0} \sup_{x \in E} |Qg^\eta(x) - Qg(x)| \to 0. \tag{2}$$

Proof. Suppose (2) is not true. Then there exists $\varepsilon > 0$, η_n converging to 0 and x_n in E such that

$$|Qg^{\eta_n}(x_n) - Qg(x_n)| \geqslant \varepsilon. \tag{3}$$

Since E is compact, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ converging to some x in E. For simplicity of notation, write $f_k = g^{\eta_{n_k}}$, $\mu_k = Q(x_{n_k}, .)$, $f = g$ and $\mu(.) = Q(x, .)$. Let

$$B_m \equiv \bigcup_{k=m}^\infty \left\{ y \in E : |f_k(y) - f(y)| > \frac{\varepsilon}{2} \right\}.$$

Then B_m decreases to the null set. Now $\mu_k(B)$ converges to $\mu(B)$ for every measurable set B in E (by the strong Feller property). Hence by the Vitali–Hahn–Saks theorem (see, e.g., [2, p. 43]),

$$\sup_j \left[\mu_j(B_m) \right] \to 0 \qquad \text{as} \quad m \to \infty. \tag{4}$$

Hence

$$|Qg^{\eta_{n_k}}(x_{n_k}) - Qg(x_{n_k})| \leqslant \int |f_k - f| \, d\mu_k$$

$$\leqslant \int |f_k - f| \, I_{\{|f_k - f| \leqslant \varepsilon/2\}} \, d\mu_k$$

$$+ \int |f_k - f| \, I_{\{|f_k - f| > \varepsilon/2\}} \, d\mu_k$$

$$\leqslant \frac{\varepsilon}{2} + K\mu_k \left(\left\{ |f_k - f| > \frac{\varepsilon}{2} \right\} \right)$$

$$\leqslant \frac{\varepsilon}{2} + K \sup_m \left[\mu_m(B_k) \right]$$

Using (4) we see that

$$\limsup_k |Qg^{\eta_{n_k}}(x_{n_k}) - Qg(x_{n_k})| \leqslant \frac{\varepsilon}{2}$$

contradicting (3). So (2) must hold. ∎

THEOREM 2. *Assume that*

[a] *For each $x \in E$, $(1/(T+1)) \sum_{t=0}^T P^t(x, .)$ converges weakly to $R(x, .)$ as $T \to \infty$ (so that R is a transition probability).*

[b] *Q has the strong Feller property.*

[c] *Q is open set irreducible, i.e., for all open sets U and all $x \in E$, $\sum_{n=1}^\infty Q^n(x, U) > 0$; and*

[d] *QR has a unique invariant measure μ^*.*

Then P^η has a unique invariant measure μ^η, which converges weakly to μ^* as $\eta \downarrow 0$.

Proof. Given any $\eta > 0$, properties [b] and [c] imply that P^η is a T-chain. Applying Theorem 16.2.5 in [14], P^η is uniformly ergodic, and has a unique invariant measure μ^η. Then

$$\mu^\eta[(1 - \phi(\eta)) \, P + \phi(\eta) \, Q^\eta] = \mu^\eta,$$

implying that

$$\mu^\eta[\phi(\eta) \, Q^\eta] = \mu^\eta - (1 - \phi(\eta)) \, \mu^\eta P.$$

For any bounded continuous function f, apply the probability measures in the equation above to $P^t f$:

$$\mu^\eta [\phi(\eta) \, Q^\eta] \, P^t f = \mu^\eta P^t f - (1 - \phi(\eta)) \, \mu^\eta P^{t+1} f.$$

Multiplying by $(1 - \phi(\eta))^t$ and summing over $t = 0, ..., T$ we obtain

$$\phi(\eta) \, \mu^\eta Q^\eta \left\{ \sum_{t=0}^{T} (1 - \phi(\eta))^t \, P^t f \right\} = \mu^\eta f - [1 - \phi(\eta)]^{T+1} \, \mu^\eta P^{T+1} f.$$

Taking $T \to \infty$, and using $\sup_x |\mu^\eta P^{T+1} f(x)| \leqslant \sup_x |f(x)|$, we get

$$\mu^\eta Q^\eta g^\eta = \mu^\eta f, \tag{5}$$

where

$$g^\eta(x) \equiv \phi(\eta) \sum_{t=0}^{\infty} [1 - \phi(\eta)]^t \, P^t f(x). \tag{6}$$

Note that $\lim_{T \to \infty} (1/(T+1)) \sum_{t=0}^{T} P^t f = Rf$ implies $g^\eta(x) \to Rf(x)$, for each $x \in E$, as $\eta \downarrow 0$. Moreover, $|g^\eta(x)| \leqslant \sup_y |f(y)| \equiv M < \infty$ for all η. We may therefore apply Lemma 1 to conclude that

$$\lim_{\eta \to 0} \sup_x |Qg^\eta(x) - QRf(x)| \to 0.$$

Consequently, if η_k is any sequence, with $\eta_k \downarrow 0$ and with $\mu^{\eta_k} \to \hat{\mu}$ in the topology of weak convergence,

$$\mu^{\eta_k} Q g^{\eta_k} \to \hat{\mu} Q R f. \tag{7}$$

At the same time,

$$\mu^\eta Q^\eta g^\eta = (1 - \psi(\eta)) \, \mu^\eta Q g^\eta + \psi(\eta) \, \mu^\eta Q_*^\eta g^\eta, \tag{8}$$

and $\psi(\eta) \to 0$ as $\eta \to 0$, while $\mu^\eta Q_*^\eta g^\eta$ is uniformly bounded. So we may combine (7) and (8) to conclude that

$$\lim_n \mu^{\eta_n} Q^{\eta_n} g^{\eta_n} = \hat{\mu} Q R f. \tag{9}$$

Combining (5) and (9),

$$\hat{\mu} Q R f = \hat{\mu} f.$$

Because f was an arbitrary bounded continuous function, it then follows from [d] that $\hat{\mu} = \mu^*$, which completes the proof. ∎

Proof of Propositions 3 *and* 4.   Define

$$\phi(\eta) \equiv \eta^2 + 2\eta(1-\eta),$$

$$\psi(\eta) \equiv \frac{\eta^2}{\eta^2 + 2\eta(1-\eta)}.$$

Let Q_* denote the transition probability when both players are subjected simultaneously to a tremble. It is then evident that

$$P^\eta = [1 - \phi(\eta)] P + \phi(\eta) Q^\eta$$

where

$$Q^\eta \equiv [1 - \psi(\eta)] Q + \psi(\eta) Q_*,$$

Assumption [a] of Theorem 2 is established by Proposition 1, while assumptions [b] and [c] are valid by construction. Hence it remains to check assumption [d], i.e., that the process QR has a unique invariant distribution. We know from Proposition 1 that every invariant distribution of QR must be concentrated on the four pss's. Hence it suffices to show that there exists a pss (the mutual defection pss) which can be reached with positive probability from every other pss in the QR process.

For this it suffices to check the claim that the action pair (D, D) will be played with positive probability at some date, when we start at any pss and subject the aspiration of one player to a tremble. The reason is that once this happens, the argument of Proposition 1 implies that with positive probability there will be an infinite run on (D, D) thereafter, causing convergence to the (D, D) pss.

If we start with the (D, D) pss then the claim is obvious, owing to inertia. And if we start with any other pss, then one small upward tremble to one player's aspiration will cause (C, D) or (D, C) followed by (D, D) to be played with some probability.   ∎

For the proof of Theorem 1, we need some additional notation. If π_s denotes the payoff to player 1 in any period s, then

$$\alpha_{s+1} = \lambda\alpha_s + (1-\lambda)\pi_s,$$

so that

$$|\alpha_s - \alpha_{s+1}| = (1-\lambda)|\alpha_s - \pi_s| \leqslant (1-\lambda) W,$$

where W is the maximum conceivable divergence between aspirations and payoffs (i.e., the width of the compact interval Λ to which aspirations belong).

Let T be the minimum number of periods that need to elapse before aspirations at time $T+2$ are different from aspirations at time 0 by an amount not less than a. Then it is clear that $T \geq T(\lambda, a)$, where $T(\lambda, a)$ is the smallest integer such that

$$(1 - \lambda)[T(\lambda, a) + 2] W \geq a.$$

It follows that

$$(1 - \lambda) T(\lambda, a) \geq \frac{a}{W} - 2(1 - \lambda). \tag{10}$$

Note that if $a > 0$, then $T(\lambda, a) > 0$ for all λ sufficiently close to unity, and indeed, that $T(\lambda, a) \to \infty$ as $\lambda \to 1$. This construct will be used at various stages below.

For any $(\hat{a}, \hat{b}) \gg 0$ such that $\delta + \hat{a} < \sigma - \hat{a}$ and $\delta + \hat{b} < \sigma - \hat{b}$, let $I(\hat{a}, \hat{b})$ be the rectangle defined by $[\delta + \hat{a}, \sigma - \hat{a}] \times [\delta + \hat{b}, \sigma - \hat{b}]$.

LEMMA 2.   *Consider any $(\hat{a}, \hat{b}) \gg 0$ such that $\delta + \hat{a} < \sigma - \hat{a}$ and $\delta + \hat{b} < \sigma - \hat{b}$. Then given any $\varepsilon > 0$, there exists $\lambda_1 \in (0, 1)$ such that*

$$\text{Prob}^\lambda(s_t \to (\sigma, C, \sigma, C) \mid J) > 1 - \varepsilon$$

for all T, all events J that are subsets of $\{(\alpha_T, \beta_T) \in I(\hat{a}, \hat{b})\}$ and measurable up to date T, and all $\lambda \in (\lambda_1, 1)$.

Proof. Fix any T and some conditioning event J as described in the statement of the lemma. Then $(\alpha_T, \beta_T) \in I(\hat{a}, \hat{b}))$. Let $T^*(\lambda)$ be the minimum number of periods t such that $(\alpha_{T+t+2}, \beta_{T+t+2}) \notin I(\hat{a}/2, \hat{b}/2)$. Observe that if $a \equiv \frac{1}{2}\min\{\hat{a}, \hat{b}\}$, then $T^*(\lambda) \geq T(\lambda, a)$ (see (10)). It follows that a lower bound for $T^*(\lambda)$ can be found that is independent of T, the initial date.

We observe, next, that there is $\zeta > 0$ (independent of T and λ) such that $s_t = (C, C)$ with probability at least ζ, for each $t = T+2$, $T+3$, ..., $T(\lambda, a) + T$. To see this, suppose first that (C, C) is played in period $t-2$. In that case (C, C) is played in period t with probability one. If (D, D) is played in period $t-2$, then the probability is easily seen to be at least $p(\hat{a}/2) \, p(\hat{b}/2)$. If at date $t-2$, (D, C) is played, then the conditional probability of playing (D, D) at $t-1$ is at least $\tilde{p}[1 - p(\delta)]$. This is because with probability $\tilde{p}$ player 1 will stick to D, while player 2 will switch with probability at least $[1 - p(\delta)]$. Thereafter, a switch to (C, C) in period t occurs with probability at least $p(\hat{a}/2) \, p(\hat{b}/2)$. The conditional probability in this case is therefore at least $\tilde{p}[1 - p(\delta)] \, p(\hat{a}/2) \, p(\hat{b}/2)$. Finally, the argument for (C, D) is symmetric.

Thus in all cases, the conditional probability of playing (C, C) at date t is bounded below by the positive number $\zeta \equiv \tilde{p}[1 - p(\delta)] \, p(\hat{a}/2) \, p(\hat{b}/2)$. The probability that (C, C) will *never* be played between t and $t + T(\lambda, a)$ is thus bounded above by $(1 - \zeta)^{T(\lambda, a) - 2}$. Select λ_1 such that

$$1 - (1 - \zeta)^{T(\lambda_1, a) - 2} \geqslant 1 - \varepsilon,$$

to ensure that (C, C) is played at least once between t and $t + T(\lambda, a)$ with probability at least $1 - \varepsilon$. And if (C, C) is played during one of these dates, s say, $(\alpha_s, \beta_s) \leqslant (\sigma - \hat{a}/2, \, \sigma - \hat{b}/2)$ by the construction of $T(\lambda, a)$. It follows that once (C, C) is played, the state will thereafter converge to the (C, C) pss with probability one. ∎

We now establish the result concerning upward drift in aspirations whenever a player has an aspiration below the maxmin payoff δ.

LEMMA 3. *Suppose $\psi \equiv \min\{\delta, \theta\} > 0$. Let u and v be positive numbers such that $u < \min\{\psi, v\}$. For any $\varepsilon > 0$ and for any date T, there exists $\lambda_2 \in (0, 1)$ such that*

$$\mathrm{Prob}^\lambda[\alpha_t < u \text{ for some } t \geqslant T \mid J] < \varepsilon \tag{11}$$

for any $\lambda \in (\lambda_2, 1)$ and any event J which is a subset of the event that $\{\alpha_T \geqslant v\}$ (measurable with respect to the process up to date T).

Proof. Pick $\lambda_2 \in (0, 1)$ to satisfy $[p(u)]^{[(w - u)/w(1 - \lambda_2)] - 1} < \varepsilon$ and $T(\lambda_2, w - u) > 1$, where $w \equiv \min\{\psi, v\} > u$. Take any $\lambda \in (\lambda_2, 1)$. Define intervening dates (random variables) l and m as follows: m is the *first* date k when $\alpha_k < u$, and l is the *last* date k before m such that $\alpha_k \geqslant w$. Clearly, $u \leqslant \alpha_k < w$ for all intervening dates.

We first claim that at no intervening date k can player 1 play D. The reason is that if he does, he attains a payoff of at least $\delta > w > \alpha_k$, his aspiration level, and will continue to play D for all dates between l and m. This contradicts the supposition that α drops below u during this time interval.

Thus between date l and $m - 1$, only (C, D) and (C, C) could be played. Moreover, for 1's aspiration to drop from w to u, (C, D) must be played at least $(w - u)/w(1 - \lambda)$ times, since 1's aspiration drops by a maximum of $(1 - \lambda) w$ each time (C, D) is played, while it increases whenever (C, C) is played. Each play of (C, D) must create disappointment of at least u for player 1; despite this he must not switch to D. The probability of this is bounded above by

$$[p(u)]^{((w - u)/w(1 - \lambda)) - 1},$$

and by our choice of $\lambda > \lambda_2$, this must be less than ε, and the lemma follows. ∎

We now turn to the dynamics consequent on aspirations falling within the region N depicted in Fig. 4. In analyzing this, the following lemma will be useful.

LEMMA 4. *Let $T(\lambda)$ be a sequence of positive integers such that $T(\lambda) \to \infty$ as $\lambda \to 1$. For each λ, let $\{X_t^\lambda, Z_t^\lambda\}$ be a finite horizon stochastic process with terminal date $T(\lambda)$, such that X_t^λ takes values only in $\{0, 1\}$ (there is no restriction on Z_t^λ).*

Suppose that for each λ, X_0^λ is equal to a constant $i(\lambda)$ (which can take the values either 0 or 1). Use the notation h_t to denote t-histories for each $t \geq 1$, and the notation $L(h_t)$ to denote the value of X_{t-1}^λ for every t-history.

Suppose that for every $\lambda \in (0, 1)$,

$$\text{Prob}(X_t^\lambda = 1 \mid h_t) \leq u < 1,$$

whenever $L(h_t) = 0$, while

$$\text{Prob}(X_t^\lambda = 1 \mid h_t) \leq v < 1,$$

if $L(h_t) = 1$. Then for every $v > 0$,

$$\text{Prob} \left\{ \frac{1}{T(\lambda)+1} \sum_{t=0}^{T(\lambda)} X_t^\lambda \leq \frac{u+v}{1-v+u} \right\} \to 1 \qquad as \quad \lambda \to 1. \tag{12}$$

Remark. $\{Z_t^\lambda\}$ appears to play no role in the statement of the lemma, but it does, for its realizations enter into the histories $\{h_t\}$.

Proof. It will be sufficient to establish the lemma for the case where $i(\lambda)$ is a constant i, independent of λ. (The general case then follows easily from a subsequence argument.) To do so, we use a coupling result, the proof of which is available on request.

Claim. Suppose $\{U_0, U_1, U_2, ..., U_T\}$ is a finite sequence of random variables that assume values in $\{0, 1\}$. Suppose that $U_0 = i$ (where i is either 0 or 1), and that for $t \geq 1$,

$$\text{Prob}(U_t = 1 \mid h_t) = p_t(h_t) \leq u < 1,$$

if $L(h_t) = 0$, and

$$\text{Prob}(U_t = 1 \mid h_t) = p_t(h_t) \leq v < 1,$$

if $L(h_t) = 1$.

Construct a (finite-horizon) Markov Chain $\{V_1, V_2, ..., V_T\}$ with values in 0, 1 such that $V_0 = i$ and for all $t \geqslant 0$,

$$\text{Prob}(V_{t+1} = 1 \mid V_t = 0) = u \qquad \text{Prob}(V_{t+1} = 0 | V_t = 0) = 1 - u$$

$$\text{Prob}(V_{t+1} = 1 \mid V_t = 1) = v \qquad \text{Prob}(V_{t+1} = 0 | V_t = 0) = 1 - v.$$

Then for all x,

$$\text{Prob}(U_1 + U_2 + \cdots + U_T \leqslant x) \geqslant \text{Prob}(V_1 + V_2 + \cdots + V_T \leqslant x) \qquad (13)$$

Let $\{V_t\}$ be the Markov chain constructed in the Claim. Then interpreting X_t^λ as U_t, (13) implies immediately that

$$\text{Prob}\left(\frac{1}{T(\lambda)+1} \sum_{t=0}^{T(\lambda)} X_t^\lambda \leqslant \frac{u+v}{1-v+u}\right) \geqslant \text{Prob}\left(\frac{1}{T(\lambda)+1} \sum_{t=0}^{T(\lambda)} V_t \leqslant \frac{u+v}{1-v+u}\right).$$

By the strong law of large numbers for Markov chains, the RHS above converges to 1, and we are done. ∎

In the arguments below, we shall initially consider the case $\theta > \delta$.

LEMMA 5. *Assume $\theta > \delta$. Use $\hat\sigma$ to denote $\min\{\theta, \sigma\}$. Then for any $b > 0$ such that $0 < \delta - 2b < \delta + 2b < \hat\sigma - 2b$, there is $a \in (0, b)$ with the property that for any $\varepsilon > 0$, there exists $\lambda_3 \in (0, 1)$ such that*

$$\text{Prob}\{s_t \to (\sigma, C, \sigma, C) \mid J\} \geqslant 1 - \varepsilon \qquad (14)$$

for every T, every $\lambda \in (\lambda_3, 1)$, and any event J that is a subset of $\{(\alpha_T, \beta_T) \in N\}$ and measurable up to T, where $N \equiv [\delta - a, \delta + a] \times [\delta + 2b, \hat\sigma - 2b]$.

Proof. We begin the proof by showing how given any b satisfying the requirements of the lemma, we select the number a, as well as by defining several other variables used in the argument.

Define the rectangle $M_1 \equiv \{(\alpha, \beta) \mid \alpha \in [\delta - b, \delta + b], \beta \in [\delta + 2b, \hat\sigma - 2b]\}$. Consider the process commencing from some aspiration vector $(\alpha_T, \beta_T) \in M_1$ at date T, and any given pair of actions. For each λ, note that a lower bound on the the minimum number of periods after which $(\alpha_{T+t+2}, \beta_{T+t+2})$ fails to lie within the larger rectangle, $M_2 \equiv \{(\alpha, \beta) \mid \alpha \in [\delta - 2b, \delta + 2b], \beta \in [\delta + b, \hat\sigma - b]\}$ is given by $T(\lambda, b)$ (see (10)). Clearly, there is an interval $(\bar\lambda, 1)$ such that for every λ in this interval, $T(\lambda, b) > 0$. *For the rest of the argument, λ will be taken to lie in this interval.*

Define $K \equiv (b/W) - 2(1 - \bar\lambda) > 0$. It follows then that for all $\lambda \in (\bar\lambda, 1)$,

$$(1 - \lambda) T(\lambda, b) \geqslant K > 0. \qquad (15)$$

Next, for any $a \in (0, b]$, define the following quantities:

$$d' \equiv 1 - p(a) > 0$$
$$e' \equiv p(2b)[1 - p(b)] > 0 \qquad (16)$$
$$i' \equiv p(\delta - 2b) < 1.$$

Observe that the inequalities above hold because $p(x) \in (0, 1)$ whenever $x > 0$.

We will impose a restriction on the choice of a. It must satisfy:

$$[\theta - (\delta + a)] \left(e'\chi - \frac{d'}{1 - i' + d'} \right) - a(1 - e'\chi) - \frac{(\delta + a) d'}{1 - i' + d'} > \frac{2a}{K} \qquad (17)$$

where χ is a strictly positive number given by

$$\chi = \min\{ p(2b) \, p(\hat{\sigma} - \delta - b), 1 - p(\delta + b) \} - \hat{\varepsilon} \qquad (18)$$

for some small but positive $\hat{\varepsilon}$.

Let us check to see that this can be done. Certainly χ can be chosen as in (18). Having done so, note that the RHS of (17) goes to 0 as $a \to 0$, while the LHS of (17) converges to $[\theta - \delta] \, e'\chi > 0$. By the continuity of both sides of (17) in a at $a = 0$, it follows that there exists a small but strictly positive such that both $\delta + a < \hat{\sigma} - a$ and (17) hold.

We complete our construction by noting that there exists $\varepsilon' > 0$ such that if we define

$$d \equiv 1 - p(a) + \varepsilon'$$
$$e \equiv p(2b)[1 - p(b)] - \varepsilon' \qquad (19)$$
$$i \equiv p(\delta - 2b) - \varepsilon',$$

then d, e and i all lie strictly between 0 and 1, and moreover,

$$[\theta - (\delta + a)] \left(e\chi - \frac{d}{1 - i + d} \right) - a(1 - e\chi) - \frac{(\delta + a) d}{1 - i + d} \geqslant \frac{2a}{K}. \qquad (20)$$

In what follows, a, b, d, e, i, $\hat{\varepsilon}$, and ε' are fixed by these considerations, irrespective of the value of $\lambda \in [\bar{\lambda}, 1)$. Now fix any conditioning event J satisfying the description in the statement of the lemma.

Let $W_1 \equiv \{(C, C)$ is played for some t such that: (i) $T \leqslant t \leqslant T(\lambda, b) + T$, and (ii) $\alpha_k \leqslant \delta + a$ for all k such that $T \leqslant k \leqslant t\}$. Note by the construction

of $T(\lambda, b)$ that $(\alpha_t, \beta_t) \leqslant (\sigma - b, \sigma - b)$ for all $t \in \{0, 1, ..., T(\lambda, b)\}$. Consequently, if W_1 occurs, (C, C) will be played repeatedly thereafter, and $s_t \rightarrow (\sigma, C, \sigma, C))$ for sure. Thus

$$\text{Prob}^\lambda \{s_t \rightarrow (\sigma, C, \sigma, C) \mid W_1\} = 1. \tag{21}$$

In what follows, then, we consider cases in which the event W_1 does not occur.

Denote by S^λ the event $\{(\alpha_t, \beta_t) \in I(a, b)$ for some $T \leqslant t \leqslant T(\lambda, b) + T\}$.

CLAIM. *There exists a function $g(\lambda)$ on $(\bar{\lambda}, 1)$, independent of T, with $g(\lambda) \rightarrow 1$ as $\lambda \rightarrow 1$, such that*

$$\text{Prob}^\lambda(S^\lambda | (\alpha_T, \beta_T) \in N, \sim W_1) \geqslant g(\lambda). \tag{22}$$

To establish this claim, let W_2 be the event $\{\alpha_t \geqslant \delta + a$ for some $t = T, T+1, ..., T(\lambda, b) + T - 1\}$. Note that if W_2 first occurs at some date t, then $(\alpha_t, \beta_t) \in I(a, b)$. Thus there exists $\lambda^* \in (0, 1)$ such that $\lambda > \lambda^*$ implies

$$\text{Prob}^\lambda(S^\lambda \mid (\alpha_T, \beta_T) \in N, W_2, \sim W_1) = 1. \tag{23}$$

In what follows, we shall suppose that $\lambda > \lambda^*$. It remains to evaluate the conditional probability $\text{Prob}^\lambda(S^\lambda \mid (\alpha_T, \beta_T) \in N, \sim W_2, \sim W_1)$ for each T.

It will be useful in what follows to concentrate on the action plays in each of the periods $T, T+1, ..., T(\lambda, b) + T$. One way of doing this is to note that the overall stochastic process, conditional on some initial action s_T, $(\alpha_T, \beta_T) \in N$ and the event $\sim (W_1 \cup W_2)$, defines a stochastic process (which in general will be non-Markovian) on the *actions* s_t played in periods $t = T, T+1, ..., T(\lambda, b) + T$. This process can be described, given the initial conditions, by a sequence of functions (one for each date), describing the probability of each action pair at date t conditional on the entire history of actions h_t up to that date. At t, let $P^\lambda(., h_t)$ denote this function. With slight abuse of notation, we will use P^λ to denote the probability of various events as well, conditional or otherwise.

Let μ denote the fraction of occurrences of (D, D), and γ the fraction of occurrences of (C, D), between dates T and $T(\lambda, b) + T$. Of course, μ and γ are random variables for each λ. Recalling the definitions in (19) and the definition of χ in (18), consider the event Z described by

$$\mu \leqslant 1 - e\chi \tag{24}$$

and

$$\gamma \leqslant \frac{d}{1 - i + d}. \tag{25}$$

Subclaim.   There exists a function $\tilde{g}(\lambda)$ on $(\bar{\lambda}, 1)$, independent of T, with $\tilde{g}(\lambda) \to 1$ as $\lambda \to 1$ such that

$$\text{Prob}^\lambda(Z \mid (\alpha_T, \beta_T) \in N, \sim W_2, \sim W_1) \geqslant \tilde{g}(\lambda). \tag{26}$$

The argument up to the third paragraph following (30) is concerned with the proof of this Subclaim.

Begin by computing a lower bound to $P^\lambda((D, D), h_t)$ for each λ, $t \geqslant 1$ and h_t. Let $s(h_t)$ denote the last action vector, i.e., the action at date $t - 1$, under the t-history h_t. If $s(h_t) = (D, D)$, then the probability that both players continue with D is at least $p(2b) \, p(\hat{\sigma} - \delta - b)$ (since player 1's aspiration cannot exceed $\delta + 2b$ and player 2's aspiration cannot exceed $\hat{\sigma} - b$, by the construction of $T(\lambda, b)$). If $s(h_t) = (C, D)$, then by a similar argument, the probability of moving to (D, D) the next period is at least $1 - p(\delta - 2b)$. If $s(h_t) = (D, C)$, the probability of moving to (D, D) is at least $1 - p(\delta + b)$. By invoking (45), we see, therefore, that

$$P^\lambda((D, D), h_t) \geqslant \chi + \hat{\varepsilon}$$

for all h_t and all λ.

Let E^λ be the event

$$\left\{ \frac{1}{T(\lambda, b) + 1} \sum_{t = T}^{T(\lambda, b) + T} 1_{(D, D)} \geqslant \chi \right\},$$

where the notation 1_s denotes the indicator function of the action vector s. Note that χ is independent of T. Now apply Lemma 4, with $X_t = 0$ whenever (D, D) is played, $X_t = 1$ if anything else is played, with Z_t set equal to some constant, with $u = v = 1 - \chi$, and $v = \hat{\varepsilon}$. We may deduce that there exists a function $g_1(\lambda)$ (with $g_1(\lambda) \to 1$ as $\lambda \to 1$) such that

[I] $P^\lambda(E^\lambda) \geqslant g_1(\lambda)$.

Next, consider the probability $P^\lambda((D, C), h_t)$ for histories such that $s(h_t) = (D, D)$. For (D, C) to follow (D, D), player 1 must stay at D while player 2 switches. Because $\alpha_t \leqslant \delta + 2b$, player 1 stays with probability at least $p(2b)$, while because $\beta_t \geqslant \delta + b$, player 2 switches with probability at least $[1 - p(b)]$. Consequently, recalling (19), we see that

$$P^\lambda((D, C), h_t) \geqslant e + \varepsilon' \tag{27}$$

for all $t \geqslant 1$, all λ, and all h_t with $s(h_t) = (D, D)$.

Now, define the event

$$F^\lambda \equiv \left\{ \frac{1}{T(\lambda, b) + 1} \sum_{t = T}^{T(\lambda, b) + T} 1_{(D, C)} \geqslant \chi \right\},$$

and consider the conditional probability $P^\lambda(F^\lambda \mid E^\lambda)$. Note that we are conditioning on the event E^λ where (D, D) occurs at least $\chi T(\lambda, b)$ times. Use this information to construct a stochastic process as follows. Each occurrence of (D, D) is to be treated as a "date". The number of dates will be taken to be the greatest integer not exceeding $\chi T(\lambda, b)$: this is $T(\lambda)$ in Lemma 4. Throw away all information after this date. Let X_t be the following random variable that describes the action vector immediately following the tth realization of (D, D): $X = 0$ if (D, C) occurs, and $X = 1$ otherwise. Let Z_t be a list of all the action vectors that follow the tth occurrence of (D, D), up to the $(t + 1)$ th occurrence of (D, D). [If (D, D) is immediately followed by another (D, D), then set Z equal to some arbitrary constant.] This process fits all the conditions of Lemma 4, if both u and v are identified with $1 - e - \varepsilon'$. Applying the lemma, we may conclude that there exists a function $g_2(\lambda)$ with $g_2(\lambda) \to 1$ as $\lambda \to 1$ such that

[II] $P^\lambda(F^\lambda \mid E^\lambda) \geqslant g_2(\lambda)$.

Again, note that g_2 is independent of T, since e is.

Next, consider the probability $P^\lambda((C, D), h_t)$. First suppose that $s(h_t) = (D, D)$. For (C, D) to follow (D, D), player 1 must switch to C while player 2 stays at D. Because we are conditioning on $\sim W_2$, we have $\alpha_t < \delta + a$. It follows that the transition occurs with probability *no more* than $[1 - p(a)]$. Therefore (recalling (19)),

$$P^\lambda((C, D), h_t) \leqslant 1 - p(a) = d - \varepsilon' \tag{28}$$

for all $t \geqslant 1$, all λ, and all h_t with $s(h_t) = (D, D)$.

Next, suppose that $s(h_t) \neq (D, D)$. In this case, $P^\lambda((C, D), h_t)$ can only be positive when (C, D) itself was played last (remember that we are conditioning on $\sim(W_1 \cup W_2)$). This requires that player 1 stick to his previous action, which will occur with probability at most $p(\delta - 2b)$. Using (19), it follows that

$$P^\lambda((C, D), h_t) < i + \varepsilon' \tag{29}$$

for all $t \geqslant 1$, all λ, and all h_t with $L(h_t) \neq (D, D)$.

Consider, then, the event

$$G^\lambda \equiv \left\{ \frac{1}{T(\lambda, b) + 1} \sum_{t = T}^{T(\lambda, b) + T} 1_{(C, D)} \leqslant \frac{d}{1 - i + d} \right\}.$$

We claim that there exists a function $g_3(\lambda)$, independent of T, with $g_3(\lambda) \to 1$ as $\lambda \to 1$ such that

[III] $P^\lambda(G^\lambda) \geqslant g_3(\lambda)$.

Proving this claim requires the application of Lemma 4 yet again. Start with the (unconditional) event G^λ. Define a stochastic process as in Lemma 4 with $T(\lambda) = T(\lambda, b)$, with $X_t = 1$ if the action vector at time $t + T$ is (C, D) (and 0 otherwise), with $u = d - \varepsilon'$ and $v = i + \varepsilon'$, and with $v = \varepsilon'$. Take Z_t to be some constant for all t. Now the lemma applies, so we see that there exists a function $g_3(\lambda)$ with the required properties.

We may now combine observations [I]–[III] to infer that

$$P^\lambda(G^\lambda \cap F^\lambda \cap E^\lambda) \geqslant \tilde{g}(\lambda), \tag{30}$$

for some function $\tilde{g}(\lambda)$ that is independent of T and with the property that $\tilde{g}(\lambda) \to 1$ as $\lambda \to 1$.

To complete the proof of (26), we unravel what the event $G^\lambda \cap F^\lambda \cap E^\lambda$ implies for the values of μ and γ, which, it will be recalled, are the fractions of (D, D)'s and (C, D)'s occurring respectively between the dates T and $T(\lambda, b) + T$.

Let κ denote the fraction of (D, C)'s during this period. Note, first, that $\mu + \kappa \leqslant 1$, while under the event F^λ, $\kappa \geqslant e\chi$. Combining these two observations, we may conclude that $\mu \leqslant 1 - e\chi$. This shows that (24) must hold under the events E^λ and F^λ.

Next, note that under the event G^λ, $\gamma \leqslant d/(1 - i + d)$, which is, of course, (25).

The observations in the last two paragraphs, coupled with (30), establish (26), and the proof of the Subclaim is complete.

Suppose, then, that the conditional event described by (26) does in fact occur. Let us find a lower bound on the change in player 1's aspirations as a result of this event. Recall that $\alpha_{t+1} = \lambda\alpha_t + (1 - \lambda)\pi_t$, where π_t is the payoff at date t, so that

$$\alpha_{t+1} - \alpha_t = (1 - \lambda)(\pi_t - \alpha_t) \tag{31}$$

Recalling that $\alpha_t < \delta + a$ for $t = 0, 1, \ldots T(\lambda, b) - 1$ (i.e., the event W_2 does not occur), we see that the RHS of (31) is bounded below by $(1 - \lambda)[\theta - (\delta + a)]$ when the action (D, C) is played, by $-(1 - \lambda)a$ when the action (D, D) is played, and by $-(1 - \lambda)(\delta + a)$ when the action (C, D) is played. Using (24) and (25), we may therefore provide a lower bound to the *total* rightward drift over $T(\lambda, b)$ periods by

$$T(\lambda, b)(1 - \lambda)\{[\theta - (\delta + a)][1 - \gamma - \mu] - \mu a - \gamma(\delta + a)\}$$

$$\geqslant T(\lambda, b)(1 - \lambda)\left\{[\theta - (\delta + a)]\left(e\chi - \frac{d}{1 - i + d}\right) - a(1 - e\chi) - \frac{(\delta + a)d}{1 - i + d}\right\}$$

$$\geqslant T(\lambda, b)(1 - \lambda)\frac{2a}{K} \geqslant 2a,$$

using (15) and (20). We conclude, then, that under the event $J \cap Z$,

$$\alpha_{T + T(\lambda, b)} \geqslant \delta + a, \tag{32}$$

while, by the construction of $T(\lambda, b)$,

$$\beta_{T + T(\lambda, b)} \geqslant \delta + b. \tag{33}$$

From (32) and (33), it follows that

$$\text{Prob}^\lambda(S^\lambda \mid (\alpha_T, \beta_T) \in N, \sim W_2, \sim W_1) \geqslant \tilde{g}(\lambda), \tag{34}$$

where $\tilde{g}(\lambda)$ was introduced in (30).

By combining (23) and (34), and defining $g(\lambda) \equiv \tilde{g}(\lambda)$, we obtain (22), which completes the proof of the claim.

Combining this claim with Lemma 2, the proof of Lemma 5 is complete. ∎

LEMMA 6.   *Assume $\theta > \delta$. For any $\varepsilon^* > 0$, there exists $\lambda_4 \in (0, 1)$ such that*

$$\text{Prob}\{s_t \to (\delta, D, \delta, D) \mid (\alpha_0, \beta_0) \geqslant (\sigma, \sigma)\} < \varepsilon^* \tag{35}$$

for all $\lambda \in (\lambda_4, 1)$.

Proof.   Denote by F the event $\{s_t \to (\delta, D, \delta, D)\}$, and by G the conditioning event $\{(\alpha_0, \beta_0) \geqslant (\sigma, \sigma)\}$.

Because $(\alpha_0, \beta_0) \geqslant (\sigma, \sigma)$ and $(\alpha_t, \beta_t) \to (\delta, \delta)$, the intervening region must be traversed almost "continuously" for λ close to unity. We note one implication of this formally. Fix any a, b and N satisfying the conditions of Lemma 5. Then there exists $\tilde{\lambda} \in (0, 1)$ such that for any $\lambda \in (\tilde{\lambda}, 1)$, the event F (conditional on G) must be accompanied by at least one of the following events:

[1] $\alpha_t < \delta - a$ for some t,

[2] $(\alpha_t, \beta_t) \in N$ for some t,

[3] $(\alpha_t, \beta_t) \in I(a, 2b)$ for some t, or

[4]   One or more of events [1]–[3], with players 1 and 2 permuted.

Figure 5 explains the simple geometry behind this assertion. The thick lines illustrate the zones implicit in [1]–[3]. The lighter lines simply indicate the mirror regions on permutation of the players.

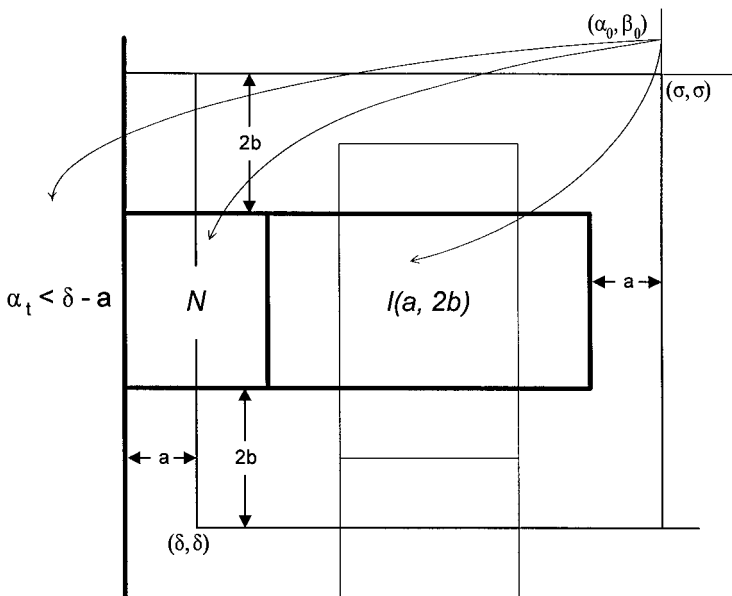By the symmetry of the problem, we may therefore assert that for $\lambda \in (\tilde{\lambda}, 1)$,

FIG. 5. Illustration of [1]–[4] in the proof of Lemma 6.

$$\tfrac{1}{2} \operatorname{Prob}^{\lambda}(F \mid G) \leqslant \operatorname{Prob}^{\lambda}(F \mid \{\alpha_t < \delta - a \text{ for some } t\} \cap G)$$

$$\times (\operatorname{Prob}^{\lambda}(\{\alpha_t < \delta - a \text{ for some } t\} \mid G)$$

$$+ \sum_{T=0}^{\infty} \operatorname{Prob}^{\lambda}(F \mid \{(\alpha_T, \beta_T) \in N \text{ first at } T\} \cap G)$$

$$\times \operatorname{Prob}^{\lambda}(\{(\alpha_T, \beta_T) \in N \text{ first at } T\} \mid G)$$

$$+ \sum_{T=0}^{\infty} \operatorname{Prob}^{\lambda}(F \mid \{(\alpha_T, \beta_T) \in I(a, 2b) \text{ first at } T\} \cap G)$$

$$\times \operatorname{Prob}^{\lambda}(\{(\alpha_T, \beta_T) \in I(a, 2b) \text{ first at } T\} \mid G). \qquad (36)$$

Take $\varepsilon = \varepsilon^*/6$ and $\lambda_4 = \max\{\tilde{\lambda}, \lambda_1, \lambda_2, \lambda_3\}$, where λ_1 is given by Lemma 2 for this ε and $(\hat{a}, \hat{b}) = (a, b)$, where λ_3 is given in Lemma 5, corresponding to this ε as well, and λ_2 is given in Lemma 3, corresponding to this ε, $u = \delta - a$ and $v = \sigma$.

The following arguments are made for $\lambda \in (\lambda_4, 1)$. The first term on the RHS of (36) is bounded above by

$$\operatorname{Prob}^{\lambda}(\{\alpha_t < \delta - a \text{ for some } t\} \mid G),$$

and Lemma 3 implies that this term is less than $\varepsilon = \varepsilon^*/6$.

Using Lemma 5, the second term on the RHS of (36) is bounded above by

$$\varepsilon \sum_{T=0}^{\infty} \text{Prob}^{\lambda}(\{(\alpha_T, \beta_T) \in N \text{ first at } T\} \mid \cap G) \leqslant \varepsilon = \frac{\varepsilon^*}{6},$$

because $(\{(\alpha_T, \beta_T) \in N \text{ first at } T\} \mid G)$ qualifies as an event of the form J in that lemma. Finally, using Lemma 2, the third term is bounded above by

$$\varepsilon \sum_{T=0}^{\infty} \text{Prob}^{\lambda}(\{(\alpha_T, \beta_T) \in I(a, 2b) \text{ first at } T\} \mid G) \leqslant \varepsilon = \frac{\varepsilon^*}{6},$$

because $\{(\alpha_T, \beta_T) \in I(a, 2b) \text{ first at } T\} \cap G$ qualifies as an event of the form J in that lemma.

Combine these three observations with (36) to complete the proof. ∎

We are finally in a position to complete the

Proof of Theorem 1 *when* $\theta > \delta$. For each λ, let R^{λ} denote the infinite-step transition in the untrembled process. We know that the limiting distribution of states (as trembles vanish) is given, for each λ, by QR^{λ}. By Proposition 1, limit weight is only placed on some pss, so that QR^{λ} may be identified with a 4×4 matrix, each cell corresponding to a pair of pss's. It clearly suffices to prove the following:

[A] $\lim_{\lambda \to 1} QR^{\lambda}(s^{**} \mid (C, C) \text{ pss}) = 0$ for all pss $s^{**} \neq (C, C)$ pss.

[B] $\lim_{\lambda \to 1} QR^{\lambda}((C, D) \text{ pss} \mid (C, D) \text{ pss}) < 1$ and $\lim_{\lambda \to 1} QR^{\lambda}((D, C)$ pss $\mid (C, D) \text{ pss}) < 1$, and the same is true starting from the (D, C) pss.

[C] $\lim_{\lambda \to 1} QR^{\lambda}((C, C) \text{ pss} \mid (D, D) \text{ pss}) > 0$.

Let s^* and s^{**} be two pss's. Then the corresponding entry of QR^{λ} is

$$QR^{\lambda}(s^{**} \mid s^*) = \int_s R^{\lambda}(s^{**} \mid s) \, dq(s \mid s^*), \tag{37}$$

where $q(. \mid s^*)$ is the measure induced by the perturbation of aspirations from the pss s^*.

By a standard argument using (37) and the dominated convergence theorem, we see that to establish [A]–[C] it suffices to prove the following steps:

[1] For any state s in the support of $q(. \mid (C, C) \text{ pss})$,

$$R^{\lambda}(s^{**} \mid s) \to 0 \qquad \text{as} \quad \lambda \to 1$$

for any other pss s^{**}. This will prove [A].

To prove [1] (and thus [A]), begin with any tremble of aspirations (of a single player) from the (C, C) pss leading to a state of the form $s = (C, \sigma + x, C, \sigma + y)$, where only one of x and y is nonzero. Wlog $y = 0$ (the argument is symmetric when $y \neq 0$). If $x < 0$, the untrembled process reverts to the (C, C) pss with probability one. So let $x > 0$. By Lemma 6, $R^\lambda((D, D)$ pss $| s) \to 0$ as $\lambda \to 1$. Convergence to the (C, D) pss is certainly included in the event described in Lemma 3, for some $u \in (0, \delta)$ and $v = \sigma$, and therefore the probability of such convergence tends to zero as well. An analogue of Lemma 3 applied to player 2 similarly takes care of the (D, C) pss, and the proof of [1] is complete.

[2] There exists a positive measure of states s under $q(. \mid (C, D)$ pss$)$ under which

$$\lim_{\lambda \to 1} R^\lambda((C, D) \text{ pss} \mid s) = 0,$$

and

$$\lim_{\lambda \to 1} R^\lambda((D, C) \text{ pss} \mid s) = 0,$$

and the same is true starting from the (D, C) pss. This will prove [B].

To prove [2] (and thus [B]), recall that the tremble has $g(\alpha' \mid \alpha) > 0$ for all aspirations α' in some nondegenerate interval around α (relative to Λ). This translates into the statement that a tremble to states s of the form $(C, x; D, \theta)$ has positive probability under $q(. \mid (C, D)$ pss$)$, for x in some nondegenerate closed interval of strictly positive numbers. For each such x, define $v = x$ and $u = \min\{x, \delta\}/2$. Now for these values of (u, v), re-convergence to the (C, D) pss is contained in the event described in Lemma 3, so this establishes the first part of [2].

For the second part, we simply apply the analogue of Lemma 3 to player 2. This time, take u to be any positive number in $(0, \delta)$ and $v = \theta$. Replacing α by β and applying the lemma, we are done.

[3] There exists a positive measure of states s under $q(. \mid (D, D)$ pss$)$ under which

$$\lim_{\lambda \to 1} R^\lambda((C, C) \text{ pss} \mid s) > 0.$$

This will prove [C].

To prove [3] (and thus [C]), note (as in part [2]) that a tremble to states s of the form $(D, \delta, D, \delta + x)$ has positive probability under $q(. \mid (D, D)$ pss$)$, for x in some nondegenerate closed interval of strictly positive numbers. Pick b satisfying the conditions of Lemma 5 such that it is less than the minimum value in this interval, any positive a satisfying

the conditions of Lemma 5 (given b), and define N as in that lemma. Then following this class of trembles, the aspiration vector $(\alpha, \beta) \in N$. By Lemma 5, we are done following any of these trembles.

The proof of the theorem in the case $\theta > \delta$ is now complete. ∎

We now turn to the case $\theta < \delta$. We start with an analogue of Lemma 2.

LEMMA 7.   *Assume $\delta > \theta \geqslant 0$. Then for any a, b satisfying $0 < a < b$, $\sigma - a > \theta + a$ and $\sigma - b > \theta + b$,*

$$\lim_{\lambda \to 1} \text{Prob}^\lambda[s_t \to (\sigma, C, \sigma, C) \mid (\alpha_T, \beta_T)$$

$$\in (\theta + a, \sigma - a) \times (\delta + b, \sigma - b) \text{ for some } T] = 1.$$

This is proved analogously to Lemma 2, by using the observation that conditional on (α_t, β_t) staying in $(\theta + a/2, \ \sigma - a/2) \times (\delta + b/2, \ \sigma - b/2)$ for at least T^* periods starting from date T, the probability of playing (C, C) at any given date between $T + 2$ and $T + T^*$ is bounded below by $\psi \equiv \tilde{p}^2[1 - p(c/2)]^2$, where $c \equiv \min\{a, b\}$. This is because a play of (D, C) or (C, D) at any date will be followed by a play of (C, C) at the next date with probability at least $\tilde{p}[1 - p(c/2)]$, as the player selecting C repeats it owing to inertia, while the player selecting D is disappointed and switches to C. Morever, a play of (D, D) at any date will cause player 2 to be disappointed by at least $c/2$. Hence with probability at least $\tilde{p}[1 - p(c/2)]$, this will be followed by a play of (D, C) at the next date. The probability that (C, C) will be played two periods later is thus bounded below by ψ. Then as $\lambda \to 1$ and $T^* \to \infty$, the probability that (C, C) will be played at least once between $T + 2$ and $T + T^*$ converges to one. With T^* constructed suitably relative to the value of λ, the result of the lemma follows.

Next, we turn to the following analogue of Lemma 5:

LEMMA 8.   *Suppose $0 < \theta < \delta$. Then there exists $z \in (0, \theta)$ such that for any $b \in (0, z)$, there is $a \in (0, b)$ with $a < \delta - \theta$, such that*

$$\lim_{\lambda \to 1} \text{Prob}^\lambda[s_t \to (\sigma, C, \sigma, C) \mid J^*] = 1 \qquad (38)$$

for any date T and event J^* a subset of $\{(\alpha_T, \beta_T) \in N^*\}$ and measurable up to T, where $N^* \equiv [\theta - a, \theta + a] \times [\delta + 2b, \sigma - 2b]$.

Proof.   For any $b \in (0, \theta/2)$, define the rectangles $M_1^* \equiv [\theta - b, \theta + b] \times [\delta + 2b, \ \sigma - 2b]$ and $M_2^* \equiv [\theta - 2b, \ \theta + 2b] \times [\delta + b, \ \sigma - b]$. Given any such b we can find $\bar{\lambda}(b) < 1$ such that $\lambda \geqslant \bar{\lambda}(b)$ implies that $T(\lambda, b)$, the minimum number of periods it takes for aspirations to exit M_2^* if they start in

M_1^*, is positive. Let $K(b) \equiv (b/W) - 2(1 - \bar{\lambda}(b))$. Then for all $\lambda \in (\bar{\lambda}(b), 1)$ we have $(1 - \lambda) T(\lambda, b) \geqslant K(b) > 0$.

Claim 9.1. There exists $z \in (0, \theta/2)$ and $\varepsilon > 0$ such that for any $b \in (0, z)$, there is $a \in (0, b)$ with $a < \delta - \theta$, such that letting

$$\chi^* \equiv \min\{[1 - p(\delta + b)] p(a), p(\sigma - b - \theta)[1 - p(\theta - b)], p(\sigma - b - \delta)\} + \varepsilon\},$$

$$i^* \equiv p(\theta - a) p(\delta + b - \theta) - \varepsilon, \qquad \text{and}$$

$$d^* \equiv 1 - p(b) + \varepsilon,$$

we have

$$[\delta - (\theta + a)] \chi^* - [1 - \chi^*] a - \frac{d^*}{1 - i^*} (\theta + a) \geqslant \frac{2a}{K(b)}, \qquad (39)$$

while χ^*, i^*, d^* all lie in $(0, 1)$.

To prove this claim, note that when evaluated at $a = b = \varepsilon = 0$,

$$\chi^* = \bar{\chi} \equiv \min\{[1 - p(\delta)], p(\sigma - \theta)[1 - p(\theta)] p(\sigma - \delta)\} \in (0, 1), \quad \text{and}$$

$$i^* = \bar{i} \equiv p(\theta) p(\delta - \theta) \in (0, 1),$$

while $d^* = 0$. Since the LHS of (39) is continuous in (a, b, ε) at $(0, 0, 0)$, and takes the value $(\delta - \theta) \bar{\chi} > 0$ at this point, it follows that there exist numbers $z \in (0, \theta/2)$ and $\bar{K} > 0$ such that $a \in (0, z)$, $b \in (0, z)$, $\varepsilon \in (0, z)$ implies that the LHS of (39) is at least $\bar{K}$, while $\chi^* > 0$, $i^* > 0$, for all $(a, b, \varepsilon) \in (0, z)^3$. Fix any $b \in (0, z)$, and then we can select a small enough so that the RHS of (39) is smaller than $\bar{K}$, and the claim is proven.

In what follows, we fix a and b as specified by Claim 9.1.

Claim 9.2. Given date T, define event Z^* by the requirement that between the dates T and $T + T(\lambda, b)$, the fraction of occurences of (D, D) is at least χ^*, and of (C, D) is at most $d^*/(1 - i^*)$. Then

$$\lim_{\lambda \to 1} \text{Prob}^\lambda[Z^* \mid (\alpha_T, \beta_T) \in N^*, \sim W_1^*, \sim W_2^*] = 1 \qquad (40)$$

where W_1^* denotes the event that (C, C) is played at least once between T and $T + T(\lambda, b)$, and W_2^* the event that $\alpha_t \geqslant \theta + a$ at any $t \in \{T, T + 1, ..., T + T(\lambda, b)\}$.

To prove this claim, note that given any date $t \in \{T, T + 1, ..., T + T(\lambda, b)\}$, and any history h_{t-1} upto $t - 1$, we have

$$\text{Prob}^\lambda[(D, D) \text{ is played at } t \mid h_{t-1}, \sim W_1^*, \sim W_2^*] \geqslant \chi^* - \varepsilon.$$

Morever,

$$\text{Prob}^\lambda[(C, D) \text{ is played at } t \mid h_{t-1}, \sim W_1^*, \sim W_2^*] \leqslant i^* + \varepsilon.$$

Applying an argument analogous to that used to prove the subclaim in Lemma 5, Claim 9.2 follows.

Claim 9.3. Let S^* denote the event that $(\alpha_t, \beta_t) \in (\theta + a, \sigma - a) \times (\delta + b, \sigma - b)$ for some date t between T and $T + T(\lambda, b)$. Then

$$\lim_{\lambda \to 1} \text{Prob}^\lambda[S^* \mid (\alpha_T, \beta_T) \in N^*, \sim W_1^*, \sim W_2^*] = 1 \tag{41}$$

To see this, suppose that event Z^* occurs. Then the total change in player 1's aspiration between T and $T + T(\lambda, b)$ is bounded below by the LHS of (39). Using the result of Claim 9.1, it follows that the total rightward drift of α is at least $2a$. Hence if event Z^* occurs, then $\alpha_{T+T(\lambda, b)} > \theta + a$. Morever, by construction, $\beta_{T+T(\lambda, b)} \in (\delta + b, \sigma - b)$, and $\alpha_{T+T(\lambda, b)} < \sigma - a$, thus establishing Claim 9.3.

Finally, combining Lemma 7 and Claim 9.3, the result of Lemma 8 follows. ∎

LEMMA 9. *Suppose* $\theta < \delta$. *Then* $\lim_{\lambda \to 1} QR^\lambda((D, D) \text{ pss} \mid (C, C) \text{ pss}) = 0$.

Proof. If $0 < \theta < \delta$, we utilise an argument analogous to that used in Lemma 6 and the part of the proof of Theorem 1 already completed: for λ close enough to 1, at least one of the following events must occur (conditional on $(\alpha_0, \beta_0) \geqslant (\sigma, \sigma)$, and the event $s_t \to (\delta, D, \delta, D)$): [1'] $\alpha_t < \theta - a$ for some t; [2'] $(\alpha_t, \beta_t) \in N^*$ for some t; [3'] $(\alpha_t, \beta_t) \in (\theta + a, \sigma - a) \times (\delta + 2b, \sigma - 2b)$; [4'] one or more of the above events, with players 1 and 2 permuted. An application of Lemmas 3, 7 and 8 establishes this.

For $\theta = 0$, the argument needs to be modified as follows. For small positive numbers a, b and λ close enough to 1, for the state to converge to the (D, D) pss following one perturbation to the (C, C) pss, it must be the case that (with the two players permuted if necessary) [3''] holds, where [3''] denotes the event that at some date t, aspirations $((\alpha_t, \beta_t))$ lie in the rectangle $(\delta - a, \sigma + a) \times (\delta + b, \sigma - b)$. This follows from the fact that given the aspirations $(\sigma, \sigma + x)$ resulting from one tremble to player 2's aspiration when starting from the (C, C) pss, the aspirations resulting in the untrembled process thereafter must be confined to the convex hull of $(\sigma, \sigma + x)$ and the four pure strategy payoff points of the game. Hence selecting any a such that $0 < a < \delta[1 - (\sigma/W)]$, where W is the maximum aspiration resulting from one tremble of aspirations from σ, it follows that $\beta_t \geqslant (\delta - a)(W/\sigma)$ implies $\alpha_t \geqslant \delta - a$. Hence event [3''] defined by a and

$b = \delta[(W/\sigma) - 1] - a(W/\sigma) > 0$ must hold at some date. The proof of Lemma 9 then follows from Lemma 7. ∎

To complete the proof of Theorem 1 for the case $0 \leqslant \theta < \delta$, note that the inequality

$$\lim_{\lambda \to 1} QR^{\lambda}((C, C) \text{ pss} \mid (D, D) \text{ pss}) > 0.$$

follows from applying Lemma 8 consequent on a perturbation of aspirations from the (D, D) pss to $(\delta, \delta + b)$. Morever, Lemma 3 implies that when $0 < \theta < \delta$, the probability of converging to any of the asymmetric pss's when starting at any of them and applying one tremble, goes to zero as $\lambda \to 1$. This completes the proof of Theorem 1 for the case $0 < \theta < \delta$. When $\theta = 0$, note that starting at any of the asymmetric pss's and applying one tremble (smaller than σ) to the aspiration of the player selecting D, (C, C) will be played at the next date with positive probability, following which the state will converge to the (C, C) pss. Hence, the probability of transiting from an asymmetric pss to the (C, C) pss is positive in the QR process, for all λ.

This completes the proof of Theorem 1.

REFERENCES

1. B. Arthur, On designing economic agents that behave like human agents, *J. Evolutionary Econ.* **3** (1993), 1–22.
2. R. B. Ash, "Real Analysis and Probability," Academic Press, New York, 1972.
3. J. Bendor, D. Mookherjee, and D. Ray, "Aspirations, Adaptive Learning and Cooperation in Repeated Games," discussion paper, Planning Unit, Indian Statistical Institute, New Delhi, 1992; revised June 1995.
4. K. Binmore and L. Samuelson, Evolutionary stability in games played by finite automata, *J. Econ. Theory* **57** (1992), 278–305.
5. K. Binmore and L. Samuelson, Muddling through: Noisy equilibrium selection, mimeo, University College, London (1993).
6. T. Börgers and R. Sarin, Learning through reinforcement and replicator dynamics, mimeo, University College London and Texas A & M University (1994).
7. R. Bush and F. Mosteller, "Stochastic Models of Learning," Wiley, New York, 1955.
8. D. Fudenberg and E. Maskin, Evolution and cooperation in noisy repeated games, *Amer. Econ. Rev.* **80** (1990), 274–279.
9. H. D. Dixon, Keeping up with the Joneses: Aspirations and learning with a known average, mimeo, University of York (1996).
10. I. Gilboa and D. Schmeidler, Case-based decision theory, mimeo, Kellogg School of Management, Northwestern University (1992).
11. E. Kalai and E. Lehrer, Rational learning leads to Nash equilibrium, *Econometrica* **61** (1993), 1019–1045.
12. M. Kandori, G. Mailath, and R. Rob, Learning, mutation and long run equilibria in games, *Econometrica* **61** (1993), 29–56.

13. Y. Kim, Satisficing, cooperation and coordination, mimeo, Department of Economics, Queen Mary and Westfield College, University of London (1995).
14. S. P. Meyn and R. L. Tweedie, "Markov Chains and Stochastic Stability," Springer-Verlag, London/New York, 1993.
15. D. Mookherjee and B. Sopher, Learning behavior in an experimental matching pennies game, *Games Econ. Behav.* **7** (1994), 62–91.
16. D. Mookherjee and B. Sopher, Learning and decision costs in experimental constant sum games, *Games Econ. Behav.* **19** (1997), 97–132.
17. F. Palomino and F. Vega-Redondo, Convergence of aspirations and (partial) cooperation in the Prisoner's Dilemma, mimeo, University of Alicante (1996).
18. A. Pazgal, Satisficing leads to cooperation in mutual interest games, mimeo, Kellogg School of Management, Northwestern University (1995).
19. A. Robson, Efficiency in evolutionary games: Darwin, Nash, and the secret handshake, *J. Theoretical Biol.* **144** (1990), 379–96.
20. A. Roth and I. Erev, Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term, *Games Econ. Behav.* **8** (1995), 164–212.
21. R. Selten, Evolution, learning and economic behavior, *Games Econ. Behav.* **3** (1991), 3–24.
22. R. Selten and R. Stoecker, End behavior in sequences of finite Prisoners' Dilemma super-games, *J. Econ. Behav. Organ.* **7** (1986), 47–70.
23. P. Suppes and R. Atkinson, "Markov Learning Models for Multiperson Interactions," Stanford Univ. Press, Stanford, 1960.
24. H. P. Young, The evolution of conventions, *Econometrica* **61** (1993), 57–84.