When Aspiring and Rational Agents Strive to Coordinate

Roy, Jaideep

Publication date:
2000

Document version
Early version, also known as pre-print

Citation for published version (APA):
Roy, J. (2000). *When Aspiring and Rational Agents Strive to Coordinate*. Department of Economics, University of Copenhagen.

DISCUSSION PAPERS
Department of Economics
University of Copenhagen

00-16

When Aspiring and Rational Agents Strive to Coordinate

Jaideep Roy

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

When aspiring and rational agents strive to coordinate*

Jaideep Roy†
Institute of Economics
University of Copenhagen
Studiestræde 6
DK 1455, Copenhagen K, Denmark.

September 4, 2000.

Abstract

The paper studies a game of common interest played infinitely many times between two players, one being aspiration driven while the other being a myopic optimizer. It is shown that the only two long run stationary outcomes are the two static equilibrium points. Robustness of long run behavior is studied to show that whenever the optimizer is allowed to make small mistakes, players are able to coordinate on the Pareto dominant equilibrium point most of the time in the long run if the speed of evolution of aspirations is sufficiently fast. However, when only the aspiring player is allowed to make small mistakes, achieving coordination is inevitable and independent of the speed at which aspirations evolve.

KEYWORDS: coordination, evolution of aspiration, myopic optimization.

JEL Classification: C72, D83.

---

*This is a part of my Ph.D. thesis at Universidad Carlos III de Madrid. I remain grateful to Sandro Brusco for his support and guidance throughout this work. Special thanks are due to Antonio Cabrales and Fernando Vega-Redondo for many remarks and suggestions which improved the paper a lot. I would also wish to thank Luis Corchon, Laurence Kranich, Carmelo Nuñez and Konstantinos Serfes for helpful comments. The paper also benefitted from seminar presentation at the Indian Institute of Management, Calcutta. Remaining errors and omissions are solely my own.

†Correspondence: tel: +45 35 32 44 22; fax: +45 35 32 30 00; e-mail: jaideep.roy@econ.ku.dk

1 Introduction

The paper studies repeated interactions between two players using different behavior rules in a strategic environment represented by a *common interest game* where there are multiple pure strategy equilibrium points, and the players face the problem of deciding which action to take in each period. Some equilibria Pareto dominate others but playing these equilibrium strategies may nevertheless involve risks arising out of the possibility of *coordination failure*. Environments exhibiting such features are in abundance in economics (and social sciences in particular) and gain special attention in models of *strategic-complementarity* of investment decisions. For example, a low level of economic activity can sometimes be thought of as the outcome of an economy wide coordination failure, in which several investments do not occur because other complementary investments are not made, and these latter investments are not forthcoming simply because the former are missing. This also provides a potential explanation of why similar economies behave very differently, depending upon their past economic ventures.

Even though such efficient outcomes (like the one where all firms invest) are equilibrium points, there are reasons to doubt that they would be chosen. For instance, the risk dominance criterion of Harsanyi and Selten (1988) may select an equilibrium from which there is scope for drastic Pareto improvement. Nevertheless, in a game of common interest, the Pareto dominant equilibrium point seems “more compelling”. Many authors have tried to derive it as the unique prediction for the repeated version of the game.[1]

We essentially study this problem of coordination failure faced by two players, one being a *myopic best-respondent* while the other being *aspiration driven*. In a model of investment, the use of different behavior rules has some interesting applications. For example, we can imagine that the

[1] Balkenborg (1993), Aumann and Sorin (1989), Binmore and Samuelson (1997), to mention a few.

working rules of management units differ across firms. In many instances, the managers of some firms set profit targets and accordingly take their investment decisions. If their decisions meet such targets, they tend to take up similar projects. Otherwise, they explore the possibility of expanding their business in other directions. In many other instances, the management sets short term goals and invests in projects which maximize their immediate profit. Our paper can be used to study the level of aggregate investment in situations where these two different management bodies interact. The aspiration driven player is "naive" and uses the following simple behavior rule. He has a payoff aspiration and takes actions. If an action yields a payoff of at least the aspiration with which that action was implemented, he keeps playing the same action. Otherwise, he switches actions with a positive probability. Furthermore, aspirations are updated by taking a simple weighted average of past aspirations and payoff experience.

Bendor et al. (1995) study repeated games where both players are aspiration driven but aspirations are static. They show that in games of common interest, players can generally attain individually rational pure strategy Pareto efficient payoffs. Also, initial conditions play an important role in the selection of long run outcomes. Karandikar et al. (1998) study 2×2 games repeated infinitely many times between two aspiration driven players with evolving aspirations. They show that for sufficiently slow evolution of aspirations and small trembles in the use of the aspiration updating rules, both players are able to coordinate on the Pareto dominant equilibrium most of the time in the long run. In a significantly different strategic setting, Palomino and Vega-Redondo (1999) propose an aspiration based model of "cooperation" where aspiring players drawn from a large population are randomly matched every period to play a Prisoner's Dilemma game. Their aspirations are updated on the basis of population average of past payoff experiences. They show that any long run outcome of the game displays cooperative behavior by a positive (and always less than 1/2) fraction of the population. Pazgal (1995) studies repeated play of mutual interest games by satisficing decision makers and shows that for a high enough initial aspiration level, there is a high probability of convergence to the Pareto dominant outcome.

On the other hand, it is a well established result that if all players are myopically rational, they will converge with probability one to some pure strategy equilibrium point in any given game. Although this result is interesting, it is somewhat intuitive. Moreover, it fails to add any further insight on the issue of *equilibrium selection*. Aspiration based models clearly have sharper predictions in the sense that they are able to find conditions under which the Pareto dominant equilibrium point is selected most of the time in the long run, as discussed above. Given this, we ask the following question: what happens if a myopically rational and an aspiring player interact? One may suspect that as two myopically rational players converge to an equilibrium point and as two aspiration driven players are able to coordinate in the long run on their commonly preferred outcome, some convex combination of these two behavior rules will demonstrate similar long run behavior. Firstly, this assertion is not at all obvious and *needs a proof*. Secondly, and most importantly, the question remains as to *whether*, and *under what conditions* the players are able to achieve long run coordination on the Pareto dominant equilibrium.

In the environment studied, players are not only "*boundedly rational*" but also their thinking and reacting processes are very different. One player is driven by *disappointment* while the other is driven by short sighted maximization. In this respect, the paper is the first of its kind in the existing literature and demands the following remarks. Whenever one moves away from models of rationality, the direction is somewhat arbitrary although not unrealistic. That is any behavior studied must adhere to some consistent thinking and reacting process. Nonetheless, if the choice of behavior is arbitrary to the above extent, there is no reason to believe that when many players interact, they all follow a common rule of thumb. Matching models with heterogenous populations is motivated exactly by this observation. Such models predict on several occasions that heterogeneity of population can be sustained in the long run. While population games are important, there are numerous instances where a given set of players with different behavior rules interact repeatedly. As another example different from the one of investment, consider a married couple. On all probabilities there are differences in personalities between the man and the woman and some of these differences stay even until death. Each day of their life may be summarized by a game which is played only between themselves. Careful study of such cases cannot be avoided, and is the central motivation of this project. We show that

i) from any level of initial aspiration, the Pareto dominant equilibrium point can indeed be achieved and that if the initial aspiration is sufficiently high, players are more likely to be able to coordinate on the Pareto dominant equilibrium in the long run, while with a relatively low initial aspiration, play may converge to the Pareto dominated equilibrium;

ii) furthermore, there are no other long run outcomes.

Robustness of long run behavior is also studied. We perturb the deterministic system by allowing players to make mistakes with positive probabilities. There are two interesting observations to make.

1) We find that whenever we allow for at least the best-respondent to tremble, as the probabilities with which players make mistakes tend to zero, the commonly desired coordination is achieved most of the time in the long run if the speed of *evolution of aspirations is sufficiently fast*.

2) Secondly, if we only let the aspiration driven player commit mistakes but do not allow for any tremble on part of the rational player, players successfully converge to coordinate[2] most of the time *independent of the*

[2] The term "coordination" will often be used to mean coordination on the Pareto dominant equilibrium point.

speed of evolution of aspirations.

These are in sharp contrast with the result in Karandikar et al., where sufficiently *slow* evolution of aspirations is necessary to obtain coordination. The intuition behind this difference in the conclusions between the two models is roughly as follows. In our case, any experience of high payoffs increases aspirations rapidly. Therefore, any subsequent experience with a low payoff disappoints the aspiration driven player to a greater extent. Thus, every time the players enter a play of the Pareto dominated equilibrium point from any other outcome that gives a relatively high payoff to the aspiring player, he experiences a very high degree of disappointment and is therefore more likely to experiment with the Pareto dominant action. This makes the rational player do the same in order to maximize her expected payoff from the following round of play. Clearly, this enhances the possibility of achieving a very high degree of coordination. On the other hand, when both players are aspiration driven, in any outcome which is not a strategic equilibrium point, whenever one player receives a high payoff, the other receives a low one. Therefore, if we want to achieve a high degree of coordination, we need the speed of aspirations to be sufficiently slow so that when in the subsequent period they enter the Pareto dominated equilibrium, *both players experience a high degree of disappointment* and therefore prefer experimenting with the Pareto dominant action.

The rest of the paper is structured as follows. Section 2 formally describes the environment and specifies the behavior rules which our players are assumed to follow. Section 3 deals with the evolution of the states of the game. Section 4 studies robustness of long run behavior. Finally the paper concludes in section 5.

2 The Model

Consider the following normal form game Γ,

		player 2	
		C	D
player 1	C	β, β	$0, \theta$
	D	$\theta, 0$	δ, δ

where $\beta > \theta > \delta > 0$. Γ is a game of Common Interest and there are two pure strategy equilibrium points (viz. (C, C) and (D, D)) with (C, C) Pareto dominating every other outcome.

Γ is infinitely repeated between players 1 and 2. Player 1 is aspiration driven while player 2 is a myopic best-respondent. Player 1's state at period t is given by his action A_t in $\{C, D\}$, and his aspiration level α_t in $[0, \beta]$. We do not allow aspirations to lie outside the convex hull of possible payoffs of Γ. This is more reasonable although allowing aspirations to lie outside $[0, \beta]$ is not logically impossible.[3] Player 2's state at period t is given solely by her action $B_t \in \{C, D\}$. Let $E = \{C, D\}^2 \times [0, \beta]$ be referred to as the state space of Γ, and denote by $e = (A, B, \alpha)$ in E as a state of Γ. Let $u_t^i : \{C, D\}^2 \to \{0, \delta, \theta, \beta\}$ be the period t payoff function for player $i = 1, 2$, as defined by the above payoff matrix.

2.1 Behavior Rules

Player 1 has a payoff aspiration with which he begins playing the game. If an action yields at least his aspiration, he keeps playing the same action. If however an action disappoints him with a lower payoff, he experiments with other available actions. Formally, player 1 behaves in the following way:

$$\begin{aligned} \text{if } u_t^1 &\geq \alpha_t, \text{ then } A_{t+1} = A_t \\ \text{if } u_t^1 &< \alpha_t, \text{ then } A_{t+1} \neq A_t \text{ with probability } (1-p), \end{aligned}$$

where $p \in (0, 1)$ indicates the *inertia* of the previous decision. We assume $p(\cdot)$ to be strictly decreasing in the extent of *disappointment*, which equals

[3] Karandikar et al. studies the case where aspirations lie in $\mathbb{R}$.

the signed difference between his aspiration and his payoff experience, $\alpha_t - u_t^1$, satisfying the following criteria:

$(A.a)$ *Satisfaction*: $p = 1$ if $\alpha_t \leq u_t^1$;

$(A.b)$ *Bounded Experimentation*: p is in $(\tilde{p}, 1)$ if $\alpha_t > u_t^1$ for some $\tilde{p}$ in $(0, 1)$;

$(A.c)$ *Disappointment Driven*: p is continuous and there exists K finite such that for all x positive, we have $1 - p(x) < Kx$, with $p(x) \geq 0$.

$(A.a)$ says that if player 1 is satisfied with his current action, he plays the current action in the subsequent period with probability 1; $(A.b)$ says that if disappointed with a particular action, he never rules out the disappointing action from his choice set; and $(A.c)$ implies that no matter how disappointed he is, he limits his experiments with new actions although his tendency to experiment with new actions increases with the disappointment.

Evolution of aspirations follows a simple *weighted average rule* over the aspiration level and the payoff experience at the previous play. In particular,

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

where λ in $(0, 1)$ is a parameter that measures the degree of *persistence* of aspirations. For λ close to 1 (the case for slow updating), the current aspiration is not very sensitive to past payoff experiences, with $\lambda = 1$ being the case when the aspiration is static. On the other hand, for λ close to 0 (the case for fast updating), the current aspiration follows the payoff experience very closely, with $\lambda = 0$ being the case when the aspiration exactly equals the previous payoff experience.

Player 2 is a myopic optimizer. She[4] is assumed to know the behavior rule of player 1, and begins by predicting the current action of her opponent in order to play a *short-sighted best response* to it that maximizes her current payoff. Let q_t in $[0, 1]$ be the probability assigned by player 2 to the event

[4]The use of gender is arbitrary.

that player 1 chooses to play C at period t. Let z_t be the probability with which player 2 plays C at period t. Typically, z_t will depend on q_t, given the value of $p(\cdot)$ as evaluated at period t. We rule out the use of mixed strategies by player 2 and take z_t to be in $\{0, 1\}$. This will not affect our results in general and we will work with a simplifying behavioral assumption that whenever player 2 is indifferent between her available actions, she keeps playing her ongoing action with probability one.

In the following section, we will study the stochastic process generated by the behavior rules of players over the state space E, and concentrate on the existence and characteristics of stationary states.

3 Evolution of activity and long run analysis

The behavior rules of the two players induce state dependent probabilities on the state space E. Let $P(e_{t+1} \mid e_0, ..., e_t)$ be the probability of being in state e_{t+1} in E at period $t+1$, given that in past periods the realized states were $e_0, ..., e_t$. Since $P(e_{t+1} \mid e_0, ..., e_t) = P(e_{t+1} \mid e_t)$, we can define a Markov process over the state space E of Γ. We call this process $M(\Gamma)$. As usual, a *long run outcome* of Γ will be defined as an infinite sequence of states in E which occur once t tends to ∞.

In the following theorem we show that repeated play of the equilibrium points of Γ are the only two long run outcomes.

Theorem 1 *(C, C, β) and (D, D, δ) are the only two* ***stationary states*** *of the Markov process $M(\Gamma)$ and infinite constant sequences of either (C, C) or (D, D) are the only two* ***long run outcomes*** *of Γ.*

The proof will involve the following four lemmas. The proof of the theorem along with those of the lemmas are to be found in the appendix.

Lemma 2 *(i)* $(C, C, [0, \beta])$ *is an ergodic set of* $M(\Gamma)$ *and* (C, C, β) *is a stationary state. (ii)* $(D, D, [0, \delta])$ *is an ergodic set of* $M(\Gamma)$ *and* (D, D, δ) *is a stationary state.*

Denote by $\Pi_t^T(A, B)$ as the probability of a T- period run on the action pair (A, B) starting at period t and ending at period $t + T - 1$.

Lemma 3 *(i) Assume that at some period* $t \geq 0$ *we have* $(A_t, B_t) = (D, C)$*. Then, for any* α_t *in* $[0, \beta]$*, we have*

$$\lim_{T \to \infty} \Pi_t^T(D, C) = 0.$$

(ii) Assume that at some period $t \geq 0$ *we have* $(A_t, B_t) = (C, D)$*. Then, for any* α_t *in* $[0, \beta]$*, we have*

$$\lim_{T \to \infty} \Pi_t^T(C, D) = 0.$$

Lemma 4 *There exists an* $\varepsilon > 0$ *such that for any period* $t \geq 0$ *with* α_t *in* $[0, \delta + \varepsilon]$ *and* $(A_t, B_t) = (D, D)$*, we have*

$$\lim_{T \to \infty} \Pi_t^T(D, D) > 0.$$

For any two subsets E_1, E_2 of the state space E, denote by $\Pi_t^T(E_1 \mapsto E_2)$ as the probability that from any state e in E_1 at time $t > 0$, the process moves for the first time to a state e' in E_2 at time $T > t$.

Lemma 5 *For any period* $t \geq 0$*, we have*

$$\lim_{T \to \infty} \Pi_t^T\left((C, D, [0, \beta]) \cup (D, C, [0, \beta]) \mapsto \{(C, C, \beta), (D, D, \delta)\}\right) = 1.$$

The intuition behind Theorem 1 is as follows. From lemma 2 we see that for any ongoing value of the aspiration, once we enter a state where both players play C, players get stuck in a repeated play of C and over time the aspiration of player 1 moves up until it converges to β, the ongoing payoff experience. We therefore converge to (C, C, β) which is a stationary state of $M(\Gamma)$. This is because starting from any initial aspiration that is at most β by assumption and getting a payoff more than this aspiration makes the aspiration driven player stick to his strategy of C. Although his aspiration evolves upwards, since $\lambda \in (0, 1)$, it cannot exceed β at any period, and therefore he never gets disappointed by playing C. The best response of player 2 remains fixed at C.

It may be of interest to point out here that when both players are aspiration driven, lemma 1 remains true. However in that case it can be shown that any state where the current payoffs equal current aspirations is a stationary state. Therefore it follows that all action profiles are potential stationary states. In our environment this is not true. To see this, we do not need the theorem as it follows trivially from the following observation. In any normal form game such that for any non-equilibrium outcome, every player has an incentive to deviate (notice that Γ satisfies this property), non-equilibrium outcomes clearly cannot be stationary states in the presence of an optimizer. However, in general it is not true that every stationary state is an equilibrium point. For example, consider a game with an outcome from where only the aspiring player has an incentive to deviate. It may well be the case that in a period such an outcome is observed, the payoff aspiration is satisfied. Then this outcome, which is not an equilibrium point by construction, is indeed a stationary state. Nevertheless, what is true in our model is that every equilibrium point is indeed a stationary state. If this was enough, we could do away with lemmas 3, 4, and 5. But how do we know that play converges in the long run to one of these stationary states and that there are no other forms of long run outcomes (such as limit cycles)? The last three lemmas help us prove that constant sequences exhibiting play of either (C, C) or (D, D) are the only two long run outcomes.

We begin with lemma 3. Suppose play begins with player 1 playing D and player 2 playing C. If player 2 knows that a payoff of θ satisfies player 1 (that is $\alpha < \theta$), she is sure that in the subsequent period, player 1 will keep playing D. In that case, player 2 plays D with probability 1 and we move to the action pair (D, D). In the case player 2 knows that player 1 will be disappointed with the payoff of θ (that is $\alpha > \theta$), she is not sure as to what will be the subsequent action of player 1. Since player 2 is an expected payoff maximizer, she attaches subjective probabilities to the subsequent actions of player 1 and plays a best response to her prediction. Furthermore, since the disappointment of player 1 decreases over time as he keeps experiencing payoffs equal to θ, the probability that player 1 will switch and play C decreases to a value such that it becomes optimal for player 2 to play D with probability 1, and we move out of (D, C). Similarly, when play begins in (C, D), for any level of aspiration, player 1 is disappointed except in the case in which his aspiration is zero. When this is the case, player 2 knows that player 1 will keep playing C and therefore deviates with probability 1 to play C and we are immediately out of (C, D). If not, player 1 experiments with all his available actions. In this experiment, as long as he attaches a very high probability (the case when he is extremely disappointed with his current action C) of playing D, it remains optimal for player 2 in expected terms to keep playing D. However, repeated experience of a zero payoff (the only interesting case here) also reduces aspirations drastically. This implies that a time comes when the probability with which player 1 plays C is sufficiently high for player 2 to deviate and play C with probability 1.

Thus, lemma 3 guarantees that if the ongoing action pair is either (C, D) or (D, C), we eventually move out. This however does not imply that (C, D) and (D, C) cannot be observed in any long run outcome because there still remains the possibility of returning to play of (D, C) or (C, D) infinitely often. It only establishes the non-ergodicity of the sets $(D, C, [0, \beta])$ and $(C, D, [0, \beta])$.

We would like to have a similar result for the situation where both players play D. However, lemma 4 guarantees that both players playing (D, D) forever has a positive probability of occurrence if aspirations are sufficiently low at the time (D, D) is observed for the first time, but surprisingly *not necessarily lower than* δ. Since infinite repetitions of (D, D) bring the aspiration closer and closer to δ, we eventually converge to the stationary state (D, D, δ). To see this observe that when the initial action profile is (D, D), if the aspiration is sufficiently low, player 1 attaches only a small probability of switching to C as a low aspiration implies a low disappointment. Furthermore, since the inertia with which player 1 keeps playing D is bounded away from zero and the degree of experimentation is bounded from above, the rate at which the probability of playing C drops is fast enough. This implies that we need only a few repetitions of (D, D) for player 1 to start playing D with near certainty. As in all such cases, it remains optimal for player 2 to keep playing D anyway, the result follows.

From lemmas 2, 3, and 4 we conclude that the Markov process $M(\Gamma)$ has the following properties:

(i) there are two ergodic sets, $(C, C, [0, \beta])$ and $(D, D, [0, \delta])$;

(ii) if α is in $\left[0, \delta + p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)\right]$, then the probability of an infinite run on the action profile (D, D) is positive; and

(iii) for any α in $[0, \beta]$, infinite runs on action profiles (D, C) or (C, D) is impossible.

Properties (i), (ii), and (iii) however do not guarantee that states where the action pairs are either (D, C) or (C, D) cannot occur infinitely often and we would like to rule out the cases where this happens. In this spirit, lemma 5 is needed to guarantee that the two stationary states (C, C, β) and (D, D, δ) are also the only two *global attractors* of $M(\Gamma)$. Notice that as $(C, C, [0, \beta])$ and $(D, D, [0, \delta])$ are ergodic sets, if there is any long run play which does not include play of (C, C) and (D, D), then also (D, C) cannot be included as otherwise (C, C) must be given a positive probability of occurrence. But since we are already in the long run, doing this would make us enter $(C, C, [0, \beta])$ with probability one. This observation is very helpful as it implies that the only remaining possibilities are infinite sequences of play fluctuating between (C, D) and (D, D). However, such fluctuations cannot go on forever as sooner or later aspirations drop towards δ and we enter a play of (D, D) with $\alpha \leq \delta$.

3.1 When player 2 is forward looking

It may be intuitively easy to see that if player 2 is *forward looking and patient*, then she can always induce the aspiration of player 1 in a way that he starts playing C and that doing so is actually better for player 2. Thus if player 2 is sufficiently patient, we can eliminate the "bad" ergodic set $(D, D, [0, \delta])$. As an example, suppose we are in the set $(D, D, [0, \delta])$. If player 2 is forward looking and sufficiently patient, she realizes that if she can manage to increase the aspiration above δ, player 1 will eventually start experimenting with action C. Since she knows that player 1 will never play C as long as his aspiration is below δ, she starts playing C herself. With probability one, this yields her a payoff of zero and a payoff of θ to player 1. She knows that repeated experience of θ eventually brings aspirations above δ. When this happens, she might as well play D and disappoint player 1, thereby inducing him to play C with a positive probability. Will player 2 actually do this? Yes, if she is forward looking and sufficiently patient and θ is sufficiently high relative to δ, because then her expected discounted lifetime payoff is higher than getting δ forever.

4 Robustness of Long Run Outcomes

The stochastic process studied in section 3 converges in the long run to play of either (C, C) forever or (D, D) forever. One may suspect that these long run outcomes may not be *robust* to either perturbations of the process or to alternative specifications. Furthermore, although Γ is *payoff symmetric* between players, since the behavior rules used by the players are different, robustness of these long run outcomes may depend crucially upon the sources of such perturbations.

4.1 Trembling Rationality

Assume all the hypotheses of section 2 and suppose that at each period player 2 takes actions according to her best response function with probability $(1-\eta)$, while with probability η she commits a mistake and chooses the non-optimal action. This specification trembles the stochastic process to what we call $M^\eta(\Gamma)$. Since the state space E is compact, by theorem 16.2.4 in Meyn and Tweedie (1993), for every $\eta > 0$, the trembled process $M^\eta(\Gamma)$ *converges strongly* to a unique limit distribution f^η, irrespective of the initial state. Thus, introduction of mistakes guarantees a *unique probabilistic long run outcome*. We will typically be interested in such limit distributions when the probability of mistakes is close to zero (i.e. $\eta \to 0$). To do this, we use the following trick.[5] Let us denote by Q^η as the one-step transition probability of the stochastic process conditional on the fact that the rational player trembles with probability $\eta > 0$, and let P^∞ denote the infinite-step transition rule in the deterministic process $M(\Gamma)$. Let $Q^\eta \circ P^\infty$ denote the composition of Q^η and P^∞. The process $\lim_{\eta\to 0} Q^\eta \circ P^\infty$ is then "equivalent" to $\lim_{\eta\to 0} M^\eta(\Gamma)$, but subjecting the rational player to tremble only once, followed by the untrembled process thereafter. This in itself captures the idea of the trembling probability being close to zero. The following proposition will be helpful.

Proposition 6 *The net $\{f^\eta\}_\eta$ converges **weakly** to a distribution f on E as $\eta \to 0$ and is the **unique invariant distribution** of the process $Q^\eta \circ P^\infty$.*

Proof. Application of proposition 3, proposition 4 and theorem 2 in Karandikar et al.. ■

[5] See Karandikar et al. for a similar use.

Since from section 3 we know that the untrembled process converges to either (D, D, δ) or (C, C, β), it follows that the invariant distribution f of the process $Q^{\eta} \circ P^{\infty}$ must be concentrated around these two states as well. *What probability weights will f attach to these states*? In the following theorem, we show that the two players are able to coordinate their actions and achieve play of the Pareto dominant equilibrium point most of the time in the long run if the aspiration follows the path of payoff experiences sufficiently closely, implying a relatively fast evolution of aspirations.

Theorem 7 *Let $f(C, C, \beta)$ be the probability weight assigned on the state (C, C, β) by the unique invariant distribution f of the process $Q^{\eta} \circ P^{\infty}$ for $\eta \to 0$. Then, $f(C, C, \beta) = 1$ if λ lies in the interval $\left(0, 1 - \frac{p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)}{\theta-\delta}\right)$.*

Proof. Denote $\Pi_t^T((A, B, \alpha) \mapsto (A', B', \alpha'))$ to be the probability that the system starts at period t in some state (A, B, α) in E and transits at period $T > t + 1$ to some other state (A', B', α'), when the rational player commits a mistake only once at period $t + 1$ with probability η. Since by theorem 1, (C, C, β) and (D, D, δ) are the only two stationary states of the untrembled process P^{∞}, it suffices to show that

$$\Pi_t^{\infty}((D, D, \delta) \mapsto (C, C, \beta)) > \Pi_t^{\infty}((C, C, \beta) \mapsto (D, D, \delta)).$$

Take any period $t > 0$ such that $(A_t, B_t, \alpha_t) = (D, D, \delta)$. Then, $A_{t+1} = D$ and $z_{t+1} = \eta > 0$ by construction. Let $B_{t+1} = C$. Since $\theta > \delta$, we have $\alpha_{t+1} - u_{t+1}^1 < 0$, implying that $A_{t+2} = D$ and $z_{t+2} = 0$. Since this further implies that $(A_{t+2}, B_{t+2}) = (D, D)$, we have

$$\alpha_{t+2} = \lambda\delta + (1 - \lambda)\theta > \delta, \text{ for all } \lambda \text{ in } (0, 1).$$

Suppose $\lambda < 1 - \frac{p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)}{\theta-\delta}$. Then, invoking the proof of lemma 4, it can be shown that α_{t+2} lies outside the interval $\left[0, \delta + p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)\right]$ and therefore, $\lim_{T\to\infty} \Pi_{t+3}^T(D, D) = 0$. Therefore, the interesting case is when

(A_{t+3}, B_{t+3}) is in the set $\{(C,C),(D,C)\}$. If $(A_{t+3}, B_{t+3}) = (C,C)$, by lemma 2, we have $(A_{t+3+n}, B_{t+3+n}) = (C,C)$ for every $n \geq 1$. On the other hand, if $(A_{t+3}, B_{t+3}) = (D,C)$, we have $\alpha_{t+3} > \alpha_{t+1}$. Following the above reasoning, the only interesting case is when (A_{t+4}, B_{t+4}) is once again in the set $\{(C,C),(D,C)\}$. We can now apply lemma 3 and claim 3 of lemma 5 to show that

$$\Pi_t^{\infty}\left((D,D,\delta) \mapsto (C,C,\beta)\right) = \eta.$$

What remains to be shown is that $\Pi_t^{\infty}\left((C,C,\beta) \mapsto (D,D,\delta)\right) < \eta$. Let us choose any period t such that $(A_t, B_t, \alpha_t) = (C,C,\beta)$. Then, $A_{t+1} = C$ while $z_{t+1} = 1 - \eta > 0$ by construction. Let us assume that $B_{t+1} = D$. Since $\beta > 0$, we have $\alpha_{t+1} - u_{t+1}^1 > 0$ and $A_{t+2} = C$ with probability $p(\lambda\beta)$ while $A_{t+2} = D$ with probability $1 - p(\lambda\beta)$. Suppose $p(\lambda\beta) > \frac{\delta}{\beta+\delta-\theta}$. Then, $\Pr((A_{t+2}, B_{t+2}) = (C,C)) = p(\lambda\beta) > 0$ and the result follows as then $\Pi_t^{\infty}\left((C,C,\beta) \mapsto (D,D,\delta)\right) < 1$. Suppose therefore that $p(\lambda\beta) \leq \frac{\delta}{\beta+\delta-\theta}$. Then, $\Pr((A_{t+2}, B_{t+2}) = (C,D)) = p(\lambda\beta) > 0$. Let us assume that $(A_{t+2}, B_{t+2}) = (C,D)$. Then, $\alpha_{t+2} = \lambda^2\beta < \lambda\beta$. Let $(A_{t+n}, B_{t+n}) = (C,D)$ for every $n = 2, ...N$, for some $N > 0$. Then, $\alpha_{t+N} = \lambda^N\beta$, and $p(\lambda^N\beta) > p(\lambda^{N-1}\beta)$. By continuity of $p(\cdot)$ in N, there exists an $N^* < \infty$ such that $p(\lambda^{N^*}\beta) > \frac{\delta}{\beta+\delta-\theta}$. Consider the probability of a run on the action pair (C,D) from period $t+1$ to period N^*. This is given by $\prod_{\tau=t+1}^{N^*} p\left(\lambda^{\tau}\beta\right)$ and is clearly positive. But this has the implication that $\Pr((A_{N^*}, B_{N^*}) = (C,C)) > 0$ as well. Therefore,

$$\Pi_t^{\infty}\left((C,C,\beta) \mapsto (D,D,\delta)\right) < \eta.$$

■

When player 2 commits a mistake with positive probability while the current state is (C,C,β), the probability of re-entering play of (C,C) is still positive for any value of the speed of evolution of aspirations. All is required is once play of (C,D) starts, it stays there for some periods until player 2 again plays C. However, if the speed of updating of aspirations is sufficiently fast and player 2 commits a mistake and plays C while in (D, D, δ), player 1 keeps playing D and the system re-enters play of (D, D) with player 1 experiencing a relatively large disappointment. This makes player 1 play C which only enforces repeated play of C by player 2 in every alternate period. This leads to play of (C, C) at some period with probability one. Thus if the speed of evolution of aspirations is sufficiently fast, a small tremble rules out the Pareto inefficient outcome in the long run. Put in another way, if λ is sufficiently low then aspirations change fast. If (D, D, δ) is the current state and the rational player trembles to (D, C), then the aspiration increases a lot. So, the aspiration driven player will be highly disappointed by playing D whenever the system re-enters play of (D, D) and will therefore deviate to C with very high probability. To that, the rational player will "best-respond" by playing C and the new stationary state is reached. On the other hand, if the initial state is (C, C, β) and there is a tremble, then there is still a positive probability of playing (C, C) again. That will not disappoint the aspiration driven player, so that the probability of moving to play of (D, D) is less than η. Mention must be made here of the fact that in case of both players being aspiration driven, such perturbations lead to play of (C, C) in the long run with probability one only if λ is very close to 1, that is when aspirations evolve very slowly.

4.2 Mistaken Aspiration Updating

Let us now perturb the model by allowing only the aspiration driven player to make mistakes in forming current aspiration levels with probability $\gamma > 0$. Thus, with probability $(1 - \gamma)$, aspiration levels are formed according to the deterministic rule as in Eq.(1), while with probability γ, the updated deterministic aspiration level is perturbed according to some density function which depends on the current value of the aspiration. We deal with the case where at no period can the aspiration lie outside $[0, \beta]$. Denote by

$M^{\gamma}(\Gamma)$ as the corresponding perturbed process. Invoking the discussion at the beginning of subsection 4.1, it will be enough to study the case where the aspiration driven player is made to make a mistake only once with probability γ and then the process is left unperturbed thereafter. We will call such a process $Q^{\gamma} \circ P^{\infty}$. The following result is relatively easy to see but nevertheless, surprising. It shows that the players achieve coordination on the Pareto dominant equilibrium point most of the time in the long run irrespective of the speed of evolution of aspirations.

Theorem 8 *Let $g(C, C, \beta)$ be the probability weight assigned to the state (C, C, β) by the unique invariant distribution g of the process $Q^{\gamma} \circ P^{\infty}$ for $\gamma \to 0$. Then $g(C, C, \beta) = 1$ for any λ in the interval $(0, 1)$.*

Proof. We use the notation as developed in subsection 4.1. It is once again sufficient to check that

$$\Pi_t^{\infty}((D, D, \delta) \mapsto (C, C, \beta)) < \Pi_t^{\infty}((C, C, \beta) \mapsto (D, D, \delta)).$$

However, since while at (C, C, β), aspirations can only fall below the current payoff experience, $\Pi_t^{\infty}((C, C, \beta) \mapsto (D, D, \delta)) = 0$ and therefore it suffices to show that $\Pi_t^{\infty}((D, D, \delta) \mapsto (C, C, \beta)) > 0$ which is trivial. ■

Since the aspiration driven player cannot by mistake choose a level of aspiration which exceeds the maximum payoff he can obtain in the given game, if both players coordinate, no matter what is the level of aspiration (mistaken or otherwise), the aspiration driven player continues to play C. Given this, the rational player never deviates to play D. Thus once the system enters the play of (C, C), it never leaves it. However, while at the stationary state (D, D, δ) of the untrembled process, the aspiration driven player is allowed to increase his aspiration with positive probability. In this case a time comes when (C, C) is played with positive probability.[6]

[6] Our conjecture is that if we allowed aspirations to lie in $\mathbb{R}$, as in Karandikar et al., we would obtain a similar result as in theorem 7.

4.3 Simultaneous Mistakes

Assume finally that both players make mistakes with positive probabilities. The description of mistakes is identical to those in subsections 4.1 and 4.2. We introduce the following notation. Given the definition of $\Pi_t^T((A,B,\alpha)\mapsto(A',B',\alpha'))$ as before, define $P_t^T((A,B,\alpha)\mapsto(A',B',\alpha'))$ as the probability that the system in the untrembled process starts at period t at some state (A,B,α) and transits to possibly some other state (A',B',α') for the first time at some future period $T>t$. Therefore, $P_t^T(\cdot)$ is the probability of a $T-$ period transition path for the untrembled process. As before, the symbols $\Pi_t^\infty(\cdot)$ and $P_t^\infty(\cdot)$ will denote $\Pi_t^T(\cdot)$ and $P_t^T(\cdot)$ respectively for $T\to\infty$. Using the following theorem, we establish an interesting observation: *the condition for successful coordination, whenever the rational player trembles, is independent of whether the aspiring player is allowed to commit mistakes or not.*

Theorem 9 *If λ lies in the interval $\left(0,1-p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)/(\theta-\delta)\right)$, then in the long run both players play C most of the time.*

Proof. As before, it is sufficient to check that

$$\Pi_t^\infty((C,C,\beta)\mapsto(D,D,\delta))<\Pi_t^\infty((D,D,\delta)\mapsto(C,C,\beta)).$$

Suppose we are at some period t such that $(A_t,B_t,\alpha_t)=(C,C,\beta)$. Then,

$$\begin{aligned}\Pi_t^\infty((C,C,\beta)\mapsto(D,D,\delta)) &= (1-\gamma)\,\eta P_t^\infty((C,D,\beta)\mapsto(D,D,\delta))\\ &\quad+\eta\gamma P_t^\infty((C,D,[0,\beta))\mapsto(D,D,\delta)).\end{aligned}$$

Invoking theorem 7, it can be shown that

$$P_t^\infty((C,D,\beta)\mapsto(D,D,\delta))<1\ .$$

Furthermore since

$$P_t^\infty((C,D,[0,\beta))\mapsto(D,D,\delta))\leq 1,$$

we have

$$\begin{aligned}\Pi_t^{\infty}\left((C,C,\beta)\mapsto(D,D,\delta)\right) &< (1-\gamma)\,\eta+\eta\gamma \\ &= \eta.\end{aligned}$$

Hence, it suffices to show that

$$\Pi_t^{\infty}\left((D,D,\delta)\mapsto(C,C,\beta)\right)>\eta.$$

Now,

$$\begin{aligned}\Pi_t^{\infty}\left((D,D,\delta)\mapsto(C,C,\beta)\right) &= \eta\gamma P_t^{\infty}\left((D,C,(\delta,\beta])\mapsto(C,C,\beta)\right) \\ &\quad +(1-\gamma)\,\eta P_t^{\infty}\left((D,C,\delta)\mapsto(C,C,\beta)\right) \\ &\quad +(1-\eta)\,\gamma\Pi_t^{\infty}\left((D,D,(\delta,\beta])\mapsto(C,C,\beta)\right).\end{aligned}$$

Again, by theorem 7, if λ lies in the interval $\left(0,1-p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)/\left(\theta-\delta\right)\right)$, we have

$$P_t^{\infty}\left((D,C,(\delta,\beta])\mapsto(C,C,\beta)\right)=P_t^{\infty}\left((D,C,\delta)\mapsto(C,C,\beta)\right)=1,$$

and since by lemma 4, it can be shown that

$$P_t^{\infty}\left((D,D,(\delta,\beta])\mapsto(C,C,\beta)\right)>0,$$

we have

$$\begin{aligned}\Pi_t^{\infty}\left((D,D,\delta)\mapsto(C,C,\beta)\right) &= \eta+(1-\eta)\,\gamma P_t^{\infty}\left((D,D,(\delta,\beta])\mapsto(C,C,\beta)\right) \\ &> \eta.\end{aligned}$$

■

Start at a situation where both players play C and the aspiration level is exactly equal to β. Suppose only the rational player makes a mistake and we are out of (C,C). With this mistake probability, the system may enter play of (C,D) with aspirations still equal to $\beta>0$. The other possibility is when both players make mistakes. The rational player mistakenly plays D while the aspiration driven player mistakenly reduces his level of current aspiration below β. It has already been discussed why the probabilities of re-entering (C, C) for both the cases remain strictly positive. Any other possible type of mistake keeps the system in the ergodic set $(C, C, [0, \beta])$ and thus becomes irrelevant in the proof. On the other hand, if we start at a situation where both players play D and the aspiration is exactly equal to δ, the aspiration driven as well as the rational player can make mistakes. The intuition behind why $P_t^{\infty}((D, C, (\delta, \beta]) \mapsto (C, C, \beta))$ and $P_t^{\infty}((D, C, \delta) \mapsto (C, C, \beta))$ are both equal to one when the speed of evolution of aspirations is sufficiently fast is explained in the discussion following theorem 7.

5 Concluding remarks

It follows from theorems 7, 8, and 9 that if only the aspiration driven player makes mistakes, long run coordination on the Pareto dominant equilibrium point is achieved independent of the speed of evolution of aspirations. However, this speed comes into play only when the optimizer is allowed to tremble irrespective of whether the aspiring player makes mistakes or not. In such cases note that the upper bound on λ (that is the minimum speed) required to obtain long run coordination depends crucially upon the exact cardinal values of the payoff matrix and the intertia function $p(\cdot)$. To see this, consider the following two examples.

Example 10

Γ_1 :

		player 2 C	D
player 1	C	4, 4	0, 3
	D	3, 0	1, 1

.

Let λ_1^* denote the upper bound on λ required to obtain long run coordination in this game. It is easy to see that

$$\lambda_1^* = 1 - \frac{p^{-1}(1/2)}{2}.$$

Example 11

$$\Gamma_2 : \qquad \begin{array}{ccc|c|c|} & & & \multicolumn{2}{c}{\textit{player 2}} \\ & & & C & D \\ \cline{4-5} & \textit{player 1} & C & 10,10 & 0,3 \\ \cline{4-5} & & D & 3,0 & 1,1 \\ \cline{4-5} \end{array}.$$

Let λ_2^* denote the upper bound on λ required to obtain long run coordination in this game. Then,

$$\lambda_2^* = 1 - \frac{p^{-1}(7/8)}{2}.$$

If the same aspiration driven player plays Γ_1 and Γ_2 with a best respondent, will the players succeed in achieving coordination in both the games in some distant future? Notice that for any given inertia function $p(\cdot)$, we have $\lambda_2^* > \lambda_1^*$. From what we have established in this paper, we can conclude that if λ lies in the interval $(0, \lambda_2^*)$, the players will certainly succeed in achieving coordination in both Γ_1 and Γ_2. On the other extreme, if λ lies in the interval $(\lambda_2^*, 1)$, they may fail to do so in both occasions. However, it may well be the case that λ falls in the interval $(\lambda_1^*, \lambda_2^*)$. Then, the players will certainly achieve coordination in Γ_2, while they may not be able to do so in Γ_1. To summarize, it seems intuitively clear that rational and aspiring players tend to achieve coordination on Pareto dominant equilibrium points in games where the "opportunity costs" of not doing so are extremely high. More on these issues are reserved for future research.

An interesting extension of this paper would be to study population dynamics where players are randomly matched from a heterogenous population consisting of aspiring and rational agents to play a common interest game. Will the aspiring population survive and possibly grow in size, or will they be made extinct by the competitive optimizers?

Appendix

Proof of Theorem 1:

Proof of Lemma 2.

(i) Take any α_t in $[0, \beta]$ and any period t such that $(A_t, B_t) = (C, C)$, implying that $u_t^1(A_t, B_t) = \beta \geq \alpha_t$. Then, $p(\alpha_t - u_t^1(A_t, B_t)) = 1$. Since $z_{t+1}(1) = 1$, and for any period $\tau > t$ such that $(A_\tau, B_\tau) = (C, C)$, with $\lambda \in (0, 1)$ we have $p(\alpha_\tau - u_\tau^1(A_\tau, B_\tau) = 1$, ergodicity of $(C, C, [0, \beta])$ is established. Furthermore, $\lim_{t \to \infty}(C, C, \alpha_t) = (C, C, \beta)$. Therefore, (C, C, β) is a stationary state.

(ii) Take any α_t in $[0, \delta]$ and any period t such that $(A_t, B_t) = (D, D)$, implying that $u_t^1(A_t, B_t) = \delta \geq \alpha_t$. Then, $p(\alpha_t - u_t^1(A_t, B_t)) = 1$. Since then, $z_{t+1}(1) = 0$, and for any period $\tau > t$ such that $(A_\tau, B_\tau) = (D, D)$, with $\lambda \in (0, 1)$, we have $p_\tau(\alpha_\tau - u_\tau^1(A_\tau, B_\tau) = 0$, ergodicity of $(D, D, [0, \delta])$ is established. Furthermore, $\lim_{t \to \infty}(D, D, \alpha_t \in [0, \delta]) = (D, D, \delta)$. Therefore (D, D, δ) is a stationary state.

■

Proof of Lemma 3.

(i) Suppose $\alpha_t \leq \theta$. Then the result is immediate as for any period $\tau > t$, $z_\tau = 0$ while $p\left(\alpha_\tau - u_\tau^1\right) = 1$. So suppose $\alpha_t > \theta$. Then for any period $\tau \geq t$,

$$p(\alpha_\tau - \theta) > (\beta - \theta) \, / \, (\beta + \delta - \theta)$$

implying that $z_\tau = 0$. Given any $T > t$, if we want to set $z_\tau = 1$ for every τ in $[t, T-1]$, we need

$$p(\alpha_\tau - \theta) \leq (\beta - \theta) \, / \, (\beta + \delta - \theta) \, .$$

This in turn would imply that

$$\prod_t^T p(\alpha_\tau - \theta) \leq ((\beta - \theta) \, / \, (\beta + \delta - \theta))^T \, .$$

Since

$$\lim_{T\to\infty} ((\beta-\theta)\,/\,(\beta+\delta-\theta))^T = 0,$$

it is easy to see that $\lim\limits_{T\to\infty} \Pi_t^T(D,C) = 0$.

(ii) Define the variable $x_t = \alpha_t - u_t^1$. Then, for any x_t and any period $t \geq 0$ such that $p(x_t) > \delta/(\beta+\delta-\theta)$, if $(A_{t-1}, B_{t-1}) = (C, D)$, we have $z_t = 1$. Furthermore, since for any period $t \geq 0$ such that $(A_t, B_t) = (C, D)$, we have $u_t^1 = 0$, it follows that $x_t = \alpha_t$. Thus for any period $T > t$, we obtain $\alpha_T = \lambda^T \alpha_t$. Since $\lambda \in (0,1)$, it follows that $\lim\limits_{T\to\infty} \alpha_T = 0$. By continuity of $\lambda^T \alpha_t$ in T and the fact that $p^{-1}(\delta/(\beta+\delta-\theta)) > 0$, there will exist a period $T^* < \infty$ such that for any period $T > T^*$, we have $\alpha_T < p^{-1}(\delta/(\beta+\delta-\theta))$ implying that $p_T(x_T) > \delta/(\beta+\delta-\theta)$.

■

Proof of Lemma 4.

For any $\alpha_t \leq \delta$, the result is immediate as in part (i) of lemma 2. So suppose $\alpha_t > \delta$. Then for any period $\tau > t$, we have $z_\tau = 0$ if and only if

$$p(\alpha_\tau - \delta) \geq \frac{\beta-\theta}{\beta+\delta-\theta}.$$

Observe that for any α_τ in $[0, \beta]$, if

$$p(\alpha_\tau - \delta) \geq \frac{\beta-\theta}{\beta+\delta-\theta},$$

then for any $\alpha'_\tau < \alpha_\tau$, we have

$$p\left(\alpha'_\tau - \delta\right) > (\beta-\theta)\,/\,(\beta+\delta-\theta)\,.$$

We need to construct the set $[0, \delta+\varepsilon]$, for some $\varepsilon > 0$, such that for any α in $[0, \delta+\varepsilon]$, we have

$$p(\alpha - \delta) \geq \frac{\beta-\theta}{\beta+\delta-\theta}.$$

But this implies that

$$\begin{aligned} \alpha - \delta &\leq p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right) \\ &\Rightarrow \alpha \leq \delta + p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right). \end{aligned}$$

Set $\varepsilon = p^{-1}\left(\frac{\beta-\theta}{\beta+\delta-\theta}\right)$. Then for any $\alpha_\tau < \delta + \varepsilon$, we have $p(\alpha_\tau - \delta) > p(\varepsilon)$. Since $\alpha_{\tau+1} < \alpha_\tau$, we have $(A_\tau, B_\tau) = (D, D)$, and therefore we only consider the set $[0, \delta + \varepsilon) \subseteq [0, \beta]$. By A.c, we know

$$p(\lambda^{\tau-1}(\alpha_\tau - \delta)) > 1 - K\lambda^{\tau-1}(\alpha_\tau - \delta) \text{ for any } \tau > t.$$

Consider the function $f(y) = \log(1-y)$. By Taylor's expansion, if $y > 0$ and sufficiently small,

$$f(y) = -y + y^2/2 + \{......\} > -2y.$$

Thus,

$$\log(1 - K\lambda^{\tau-1}(\alpha_\tau - \delta)) > -2K\lambda^{\tau-1}(\alpha_\tau - \delta).$$

This implies that

$$\begin{aligned} \log p(\lambda^{\tau-1}(\alpha_\tau - \delta)) &> \log(1 - K\lambda^{\tau-1}(\alpha_\tau - \delta)) \\ &> -2K\lambda^{\tau-1}(\alpha_\tau - \delta). \end{aligned}$$

Since by $A.c$,

$$\sum_{\tau}^{\infty} K\lambda^{\tau-1}(\alpha_\tau - \delta) = K(\alpha_\tau - \delta)/(1-\lambda) < \infty,$$

we have $\sum\limits_{\tau}^{\infty} -2K\lambda^{\tau-1}(\alpha_\tau - \delta) > -\infty$. Thus,

$$\sum_{\tau}^{\infty} \log p(\lambda^{\tau-1}(\alpha_\tau - \delta)) > \sum_{\tau}^{\infty} -2K\lambda^{\tau-1}(\alpha_\tau - \delta) > -\infty.$$

But

$$\sum_{t}^{\infty} \log p(\lambda^{\tau-1}(\alpha_\tau - \delta)) = \log \prod_{t}^{\infty} p(\lambda^{\tau-1}(\alpha_\tau - \delta)),$$

implying that $\log \prod_{\tau}^{\infty} p(\lambda^{\tau-1}(\alpha_\tau - \delta)) > -\infty$ and therefore $\prod_{\tau}^{\infty} p(\lambda^{\tau-1}(\alpha_\tau - \delta))$ > 0. This implies that $\lim_{T\to\infty} \Pi_t^T(D, D) > 0$.

■

Proof of Lemma 5.

The Lemma is proved by contradiction. Denote by $(A, B)^*$ to be the set of all possible sequences of action profiles realized in the long run. Thus, $(A, B)^*$ is the set of all possible long run outcomes of Γ. Define

$$\mathcal{L}\{(C, D), (D, C), (D, D), (C, C)\}$$

as the subset of long run outcomes in $(A, B)^*$ such that each one of the four pairs of action profiles appears infinitely often. By lemma 2, if there exists a long run outcome ℓ in $(A, B)^*$ such that the action pairs (C, D) and (D, C) are both in ℓ, then (C, C) cannot be in ℓ and therefore we conclude that $\mathcal{L}\{(C, D), (D, C), (D, D), (C, C)\} = \emptyset$. Let

$$\mathcal{L}\{(C, D), (D, C), (D, D)\} \subset (A, B)^*$$

be all possible elements of $(A, B)^*$ containing each (C, D), (D, C) and (D, D) infinitely often and no other element.

Claim 1: $\mathcal{L}\{(C, D), (D, C), (D, D)\} = \emptyset$. To see claim 1, suppose there exists an ℓ in $\mathcal{L}\{(C, D), (D, C), (D, D)\}$. For any period $t \geq 0$ such that $(A_t, B_t) = (D, C)$, since we are in ℓ, the aspiration level cannot lie outside $[0, \theta]$. Thus, $p\left(\alpha_t - u_t^1\right) = 1$. Let $B_\tau = C$ for the periods $\tau = t{+}1, t{+}2, ..., t{+}k$, for some k. It is easy to check that there exists T such that at any period $\tau = t + T$, we have $B_{t+T} = D$. Since ℓ is in $\mathcal{L}\{(C, D), (D, C), (D, D)\}$, there exists T' such that for all $T > T'$, we have $\alpha_{\tau+T}$ in $[0, \theta]$. There are two possibilities at this stage. If $\theta \leq \delta + \varepsilon$, then by lemma 4, we have $\lim_{T\to\infty} \Pi^{T}_{\tau+K}(D,D) > 0$. But this, by lemma 2, implies that (C,D) and (D,C) does not belong to ℓ. On the other hand, if $\theta > \delta + \varepsilon$, then $z_{\tau+T} = 1$ and $p_{\tau+T}(\alpha_{\tau+T} - \delta) < 1$, implying that $\Pr[(A_{\tau+T}, B_{\tau+T}) = (C,C)] > 0$, thus contradicting (by lemma 2) with the fact that (C,C) is not included in ℓ.

Let

$$\mathcal{L}\{(C,D),(D,D)\} \subset (A,B)^*$$

be all possible elements of $(A,B)^*$ containing each (C,D) and (D,D) and no other element.

Claim 2: $\mathcal{L}\{(C,D),(D,D)\} = \emptyset$. To see this, suppose by passing to contradictions, we find an ℓ in $\mathcal{L}\{(C,D),(D,D)\}$. Then, for every period $t \geq 0$ such that (A_t, B_t) occurs in ℓ, there exists $\epsilon(t) > 0$ with $\alpha_t < \delta + \epsilon(t)$, and there exists a period t^* such that $\alpha_{t^*} < \delta$. Since for every $t > t^*$, we have $\alpha_t < \delta$, there exists a $T^* > t^*$ such that

$$z_{T^*+1}(p(\alpha_{T^*} - \delta \mid (A_{T^*}, B_{T^*}) = (C,D)) = 1,$$

where the expression $p(\alpha_{T^*} - \delta \mid (A_{T^*}, B_{T^*}) = (C,D))$ denotes the value of $p(\cdot)$ at period T^*+1, given that at period T^* the realized action profile is (C,D). Since ℓ is in $\mathcal{L}\{(C,D),(D,C),(D,D)\}$, for every period $T > T^*$, we have (A_T, B_T, α_T) to belong to $(D,D,[0,\delta])$. By lemma 2, this contradicts with (C,D) being in ℓ.

Define similarly the set $\mathcal{L}\{(D,C),(D,D)\} \subset (A,B)^*$.

Claim 3: $\mathcal{L}\{(D,C),(D,D)\} = \emptyset$. The proof of this claim follows directly from the proof of claim 1.

Finally by lemma 3, $(A,B)^* = \mathcal{L}\{(C,C),(D,D)\}$ and the result follows by lemma 2 and lemma 4.

■

This completes the proof of the theorem. ■■

References

[1] Aumann R. J. and S. Sorin, 1986. Cooperation and Bounded Recall, Games and Economic Behavior, 1, 5-39.

[2] Balkenborg D., 1993. Strictness, Evolutionary Stability and Repeated Games with Common Interests, mimeograph, University of Pennsylvania.

[3] Bendor J., D. Mookherjee, and D. Ray, 1995. Aspirations, Adaptive Learning and Cooperation in Repeated Games, Discussion Paper, Planning Unit, Indian Statistical Institute, New Delhi.

[4] Binmore K. and L. Samuelson, 1997. Muddling Through: Noisy Equilibrium Selection, mimeograph, University College London.

[5] Harsanyi J. and R. Selten, 1990. A General Theory of Equilibrium Selection in Games, Cambridge, MIT Press.

[6] Karandikar R., D. Mookherjee, D. Ray, and F. Vega-Redondo, 1998. Evolving Aspirations and Cooperation, Journal of Economic Theory, 80, 292-331.

[7] Meyn S. P. and R. L. Tweedie, 1993. Markov Chains and Stochastic Stability, Springer-Verlag London Ltd..

[8] Palomino F. and F. Vega-Redondo, 1999. Convergence of Aspirations and (Partial) Cooperation in the Prisoner's Dilemma, International Journal of Game Theory, 28, 465-488.

[9] Pazgal A., 1995. Satisficing leads to Cooperation in Mutual Interest Games, mimeograph, Northwestern University.