CONVERGENCE OF ASPIRATIONS AND (PARTIAL) COOPERATION IN THE PRISONER'S DILEMMA[*]

Frédéric Palomino & Fernando Vega-Redondo[**]

[**] Frédéric Palomino: Tilburg University
    Fernando Vega-Redondo: University of Alicante and Instituto Valenciano de Investigaciones Económicas.

CONVERGENCE OF ASPIRATIONS AND (PARTIAL) COOPERATION IN THE PRISONER'S DILEMMA

Frédéric Palomino & Fernando Vega-Redondo

ABSTRACT

This paper proposes an aspiration-based model for (anonymous) cooperation where a large population of agents are re-matched every period to play a Prisoner's Dilemma. At each point in time, agents hold a certain common aspiration level which is updated on the basis of population-average experience. On the other hand, those agents who (relative to current aspiration) feel "dissatisfied" switch actions at a rate which is increasing in the magnitude of the dissatisfaction. The induced process is shown to converge in the long run under quite general conditions. Moreover, if agents are responsive enough, the long-run social state is seen to display some extent of cooperation, a constant positive fraction of the population (always less than half) choosing to cooperate in every period.

1 Introduction

The Prisoner's Dilemma has become a paradigmatic benchmark in the study of cooperative behavior within sharply competitive environments. In response to its familiar "paradox," the game-theoretic literature has approached the study of this game from a wide variety of different perspectives (supergames, incomplete information, bounded rationality, evolution, etc.).[1] In everyone of these cases, it has been a crucial consideration in the analysis that players' interaction should display a certain degree of stability over time.[2] Otherwise, the strict dominated nature of cooperation in any one-shot situation seems to lead unavoidably to full defection as the only stable social configuration, at least when the population involved is very large.[3]

This raises the question of whether there might be some considerations inducing cooperation (at least partially) in social situations that are substantially more "diffuse," i.e., when interaction is mostly ephemeral and anonymous within a very large population. Casual observation suggests that in many such situations (e.g., when "strangers" orderly interact in, say, queues and traffic, or when valuable lost items are returned) *indiscriminate* opportunistic behavior does not fully prevail, despite substantial odds to the contrary.

Of course, one possible rationalization of such altruistic behavior is that some people are guided by moral (non-strategic) principles. Here, however, we show that an alternative (possibly complementary) view can also be put forward. Specifically, we propose a very simple model of bounded rationality which leads to a partial (never full) degree of cooperation under those circumstances. The context involves a large set of agents (a continuum) who are randomly matched in pairs every period to play a Prisoner's Dilemma. On the basis of past experience, they hold at every point in time some common aspiration on the payoffs they should get from the game. If, for any particular agent, this aspiration is fulfilled (i.e. the payoff obtained does *not*

---

[1]See, for example, Friedman (1977) for models displaying indefinite repeated interaction, Kreps *et al* (1982) for a context with incomplete information, Neyman (1985) or Rubinstein (1986) for an approach based on bounded rationality, and Fudenberg & Maskin (1990) or Binmore & Samuelson (1992) for an analysis reflecting evolutionary considerations.

[2]Even when, as in many evolutionary models, random re-matching of players is assumed every period – cf. Axelrod (1984) – the game assumed to be played at each "point" in time is in fact a repeated game, i.e., reflects some stable interaction between the two selected opponents.

[3]In a context where the population is *not* too large (relative to the discount rate and the gains from unilateral defection), Kandori (1992) and Ellison (1994) show that cooperation can be sustained in the Prisoner's Dilemma as part of an intertemporal equilibrium even with repeated random matching.

fall below aspiration), this agent is assumed to remain choosing her former action. Otherwise, she changes it with some probability, which will generally depend on the magnitude of the dissatisfaction.

To underscore the social dimension of aspirations (see more on this below), the whole population is postulated to have a common level of aspiration which is continuously updated on the basis of "social experience." Specifically, it is assumed that if the observed (population-wide) average payoff is higher than current aspiration, the latter gradually grows; correspondingly, if average payoff falls below the current aspiration level, the entailed adjustment is carried out downwards.

In the context outlined, the main results of our analysis can be summarized as follows. First, we show that the long-run dynamics of the process is particularly well-behaved. Specifically, the process is always seen to display long-run convergence to some limit state, starting from any initial conditions.

A crucial issue, of course, hinges upon whether or not such long-run behavior will display some degree of cooperation. We characterize the conditions for which each of these possibilities arise. Essentially, it is shown that, if players are sufficiently reactive when "dissatisfied", all robust limit states involve a *positive* fraction of cooperators (never more than half of the population). An interesting feature of these partially-cooperative limit states is that they give rise to a situation which remains in a never-ending flux. For, even though the extent of cooperation they induce of course remains stationary, there is always a constant fraction of the population which feels dissatisfied and thus switches actions.

Finally, the issue of uniqueness is addressed. In this respect, certain (non-pathological) conditions are identified which ensure that the process displays a unique (robust) limit state. In this case, therefore, the model produces a single long-run prediction, whose dependence of the parameters will be illustrated by means of simulations. We shall also rely on further simulations to illustrate the behavior of the model when such uniqueness does not apply, the dynamics exhibiting a multiplicity of equilibria with disjoint basins of attraction.

The most direct precedent of the research reported here can be found in the paper of Karandikar, Mookerjee, Ray & Vega-Redondo (1995) – hereafter KMRV – itself motivated by prior work of Bendor, Mookerjee & Ray (1992). In KMRV, *two* given players interact over time playing a certain Prisoner's Dilemma. Along the process, both players follow an aspiration-based adjustment rule, complemented with infrequent experimentation. On the other hand, each player adjusts her aspiration level on the basis of her *own* experienced payoffs. Under these conditions, these authors show that the unique invariant distribution of the process is concentrated (for a small mutation rate and gradual aspiration adjustment) in the cooperative outcome.

6

Our approach exhibits the following crucial differences with KMRV. Firstly, the aspiration-updating dynamics is postulated to depend on social information. This responds to the intuitive idea that, in *symmetric* social contexts, aspirations should be largely tailored to population-wide experience (assuming information on it is available). A second crucial difference concerns the size of the population and the structure of interaction postulated in each case. In KMRV, the same (two) individuals are taken to interact throughout the process. This permits the kind of inter-agent "feedback effects" which crucially underlie their results.[4] Here, however, such effects are fully absent since the population is infinitely large and continuously re-matched every period. Despite the substantial odds against cooperation induced by such anonymous and large-population context, we find that some positive fraction of the population (in contrast with KMR, never more than half of it) may end up cooperating in the long run.

More generally, the approach pursued in this paper is also related to the recent literature on bounded rationality in games which, in Selten's (1991) terminology, displays "strategic reinforcement." Important representatives of it are the Case-Based Theory of Gilboa & Schmeidler (1995),[5] or the evolutionary models of Binmore & Samuelson (1993) and Börgers & Sarin (1994). This literature has an early precursor in the work of Bush & Mostellar (1955) and has been found to enjoy some empirical support in laboratory experiments – see, for example, Selten & Stoecker (1986), Roth & Erev (1993), or Mokherjee & Sopher (1994).

The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 presents and discusses the analytical results. Section 4 illustrates matters by means of some simulations. Finally, Section 5 concludes.

---

[4]For example, when one player in the KMRV context starts defecting after a long phase of cooperation, the other player will immediately respond by defection: the latter's payoff has drastically fallen and, therefore, she is now substantially dissatisfied with cooperation. This makes any unilateral switch to defection a very short-lived situation, thus imposing (dissatisfaction) pressure on both players to regain a cooperative situation.

[5]See also the work by Kim (1995) and Pazgal (1995) who apply Cased-Based Theory to the Prisoner's Dilemma. In order to obtain long-run cooperation they require that players' initial aspirations are sufficiently high relative to the cooperation payoffs.

7

2 The Model

Time $t \in [0, \infty)$ is measured continuously. At every t, the individuals of a certain large population (with the cardinality of the continuum) are randomly matched in pairs to play the following Prisoner's Dilemma:

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

Table 1

with $\theta > \sigma > \delta > 0$ and $\theta/2 < \sigma$. These are the standard payoff requirements on the Prisoner's Dilemma. They make "defection" D a dominant strategy, whereas "cooperation" C becomes the only *symmetric* and efficient outcome (in particular, it dominates any symmetric mixing of cooperation and defection). We shall also find it useful to assume that $\theta/2 > \delta$. This ensures that the average payoff earned by the population is non-decreasing in the fraction of cooperators, which simplifies some of the ensuing arguments.

At every t, the population is assumed to have a *common* aspiration $\alpha(t)$ on what payoff to obtain from playing the game.[6] If the payoff a player receives at t is at least as high as this aspiration, then she is assumed to remain playing the same action. Otherwise, she feels "dissatisfied" and is postulated to switch actions at a rate which is dependent on the dissatisfaction gap.

To formalize these matters in a compact fashion, let $f(\chi)$ denote the rate at which any given agent receiving a payoff of π changes her action when the current aspiration equals $\alpha = \pi + \chi$.[7] Quite generally, this function is simply postulated to be non-decreasing and satisfy:

$$f(\chi) \begin{cases} = 0 & \text{if } \chi \leq 0, \\ > 0 & \text{if } \chi > 0. \end{cases}$$

It is also assumed to be continuously differentiable on the non-negative real line $[0, \infty)$. Note that this only requires $f(\cdot)$ to be right-differentiable at

---

[6]Note that the assumption of a common aspiration level is *not* the key feature of our model. Rather, the crucial requirement is that the aspiration-updating rule be symmetric across players and based on identical population-average information. For, under these conditions, any initial heterogeneity in aspiration levels would disappear in the long run, thus rendering the model proposed directly applicable.

[7]As usual in continuous-time adjustment models, $f(\alpha, \pi)$ can be conceived as the independent probability rate at which each of the agents obtaining a payoff of π switches actions. In fact, since speeds of adjustment do not affect our analysis (see below), these rates are best viewed as just *relative* magnitudes.

$\chi = 0$, thus permitting that marginal increases in dissatisfaction from $\chi = 0$ have positive first-order effects.

Denote by $\mu(t) \in [0,1]$ the frequency of individuals in the population playing strategy C at t. On the basis of our previous discussion, the law of motion for this variable can be written as follows:[8]

$$\dot{\mu} = \kappa_1 \left\{ \begin{array}{c} -\mu^2 f(\alpha - \sigma) - \mu(1 - \mu) f(\alpha - 0) \\ + \mu(1 - \mu) f(\alpha - \theta) + (1 - \mu)^2 f(\alpha - \delta) \end{array} \right\} \tag{1}$$

where $\kappa_1 > 0$ is some arbitrary speed of adjustment. The first two terms of the above expression represent the flow of those individuals formerly playing C who, given the payoffs they obtain after random matching, switch to playing D. Reciprocally, the latter two terms reflect those individuals previously choosing D who turn to playing C.

Note that (1) can be re-written as follows:

$$\dot{\mu} = \kappa_1 \left\{ \left[\begin{array}{c} \mu^2 [1 - f(\alpha - \sigma)] + \mu(1 - \mu) [1 - f(\alpha - 0)] + \\ \mu(1 - \mu) f(\alpha - \theta) + (1 - \mu)^2 f(\alpha - \delta) \end{array} \right] - \mu \right\}$$

which, provided $f(\chi) \leq 1$, suggests the following slight reinterpretation of the dynamics. At every t, an infinitesimal and *unbiased* sample of the population is given the opportunity of adjusting their strategy. From those selected with a dissatisfaction gap equal to χ, only a fraction $f(\chi)$ in fact decides to change strategies.

To complete the dynamics, we need to specify the law of motion for the aspiration level $\alpha(t)$. As explained, we shall focus on a context where aspiration adjustment is a social phenomenon based on population-wide performance. Specifically, we postulate that the current aspiration level gradually adjusts in the direction of realized average payoffs. That is,

$$\dot{\alpha} = \kappa_2 \left\{ \bar{\pi}(\mu) - \alpha \right\} \tag{2}$$

where $\kappa_2 > 0$ can be chosen arbitrarily and

$$\bar{\pi}(\mu) \equiv (1 - \mu) \left(\mu\theta + (1 - \mu)\delta \right) + \mu^2 \sigma \tag{3}$$

denotes the average payoff obtained by the population.

The pair $\omega(t) \equiv (\mu(t), \alpha(t))$ represents the state of the process at t. Its dynamics is governed by the two-dimensional system of differential equations given by (1) and (2), whose behavior is analyzed in the next section.

---

[8]For the sake of notational simplicity, the time variable t is omitted.

3 Analysis

It will facilitate our discussion to focus on

$$\Omega = [0,1] \times [\delta,\sigma]\,,$$

as the (compact) state space of the system. Thus, we shall concern ourselves with aspiration levels that lie between the defection and cooperation payoffs, δ and σ. Under our assumptions on payoffs, we have:

$$\mu \in (0,1) \Rightarrow \delta < \bar{\pi}(\mu) < \sigma. \qquad (4)$$

Therefore, it is clear that the dynamical system (1)-(2) is well defined within Ω since, in particular, the corresponding vector field "points inwards" in the aspiration dimension at non-stationary points in the boundary of this set. Furthermore, the fact that the vector field is Lipschitzian guarantees (by virtue of standard results in the Theory of Differential Equations) the existence of a unique solution path:

$$\omega(\cdot \mid \omega_0) : [0,\infty) \to [0,1] \times \mathbb{R}\,,$$

given any initial conditions $\omega_0 \in \Omega$.

Obviously, both $(0,\delta)$ and $(1,\sigma)$ are rest point of the dynamics. Our first result establishes that, except if the system starts in the latter full-cooperation state, every limit point of the system involves no more than half of the population cooperating.

Proposition 1 *For any initial conditions $\omega_0 = (\mu_0,\alpha_0)$ with $\mu_0 \neq 1$, all ensuing limit points $\omega^* = (\mu^*,\alpha^*)$ have $\mu^* \leq 1/2$.*

Proof. Rewrite (1) compactly as

$$\dot{\mu} = \kappa_1\, G_1(\mu,\alpha)\,, \qquad (5)$$

Since, for all t, $\alpha(t) \leq \sigma < \theta$, we have

$$f(\alpha(t) - \theta) = f(\alpha(t) - \sigma) = 0$$

and therefore we can write:

$$\begin{aligned} G_1(\mu(t),\alpha(t)) &= -\mu(t)(1-\mu(t))\,f(\alpha(t)) + (1-\mu(t))^2\,f(\alpha(t) - \delta) \\ &= -(1-\mu(t))\,[\mu(t)\,f(\alpha(t)) - (1-\mu(t))\,f(\alpha(t) - \delta)]\,. \end{aligned}$$

Assume that $\mu(t) > 1/2$. Then, since $f(\alpha(t)) > 0$ and

$$f(\alpha(t)) \geq f(\alpha(t) - \delta) \geq 0,$$

10

it follows that

$$\dot{\mu} = \kappa_1\, G_1(\mu, \alpha) < 0,$$

provided $\mu(t) < 1$. The latter is implied by any initial $\mu_0 < 1$, which completes the proof of the Proposition. ∎

In fact, one can build upon the previous Proposition and strengthen substantially its conclusion. The next result establishes that all orbits of the system converge to a *unique* limit point with the above described characteristics.[9]

Theorem 1 *Given any initial conditions $\omega_0 = (\mu_0, \alpha_0)$, the ensuing orbit of the system (1)-(2) satisfies $\lim_{t\to\infty} \omega(t \mid \omega_0) = \omega^*$ for some rest point $\omega^* = (\mu^*, \alpha^*)$. Furthermore, if $\mu_0 \neq 1$, then $\mu^* \leq 1/2$.*

Proof. As established by the well-known Poincaré-Bendixon Theorem (see, for example, Hirsch & Smale (1974, p. 248)), the limit sets of the orbits induced by a two-dimensional dynamical system either include a rest point or are closed orbits. Thus, in order to ensure that the orbits of our system are convergent, it is enough to rule out closed orbits and saddle rest points towards which orbits return infinitely often.

To do so, we shall rely on the following result in the Theory of Ordinary Differential Equations (see, for example, Arnold (1973, p. 198)):[10]

Liouville's Theorem: Let $\dot{x}(t) = H(x(t))$ be a dynamical system defined on a certain open subset $U \subseteq I\!\!R^n$, where $H(\cdot)$ is a differentiable vector field. Then, if $A \subseteq U$ has a volume $V \equiv \int_A dx$, then the volume $V(t)$ of the set $A(t) = \{y = x(t) : x(0) \in A\}$ satisfies:

$$\dot{V}(t) = \int_{A(t)} \mathbf{div}\ H(x)\ dx,$$

where the divergence of the vector field $H(\cdot)$ is defined as follows:

$$\mathbf{div}\ H(x) \equiv \sum_{i=1}^{n} \frac{\partial H_i(x)}{\partial x_i},$$

i.e., the trace of the Jacobian of $H(\cdot)$.

---

[9]By virtue of the Poincaré-Bendixon Theorem, what this rules out is the existence of limit cycles. Of course, Theorem 1 makes Proposition 1 redundant. They are both included to clarify the line of reasoning involved.

[10]Other recent applications of this result within economic and game-theoretic environments include Keenan & Rader (1985) or Corchón & Mas Colell (1996). The latter apply it to guarantee convergence of adjustment paths in bilateral games with one-dimensional and compact strategy spaces. Liouville's Theorem has also been widely used in the evolutionary literature – see e.g. Hofbauer and Sigmund (1988) or Vega-Redondo (1996).

Let $v : \Omega \to I\!\!R^2$ be the vector field leading to (1)-(2). We show below that in the relevant subset of the state space (a region which is absorbing in finite time), it displays negative divergence. To see that this implies the claimed orbit convergence, suppose instead (the argument here follows Corchón & Mas-Collell (1996)) that the system were to induce some limit cycle or have a limit rest point which is a saddle. In the first case, the set encircled by the closed orbit would be obviously invariant. So would happen in the second case for the region enclosed by a corresponding chain of orbits joining rest points. (Such a region must exist if orbits are to visit any neighborhood of the saddle infinitely often.) In both cases, region invariance implies constancy of (positive) volume, thus being incompatible (in view of Liouville's Theorem) with a vector field displaying negative divergence.

As advanced, we shall not show negative divergence of the vector field everywhere in the state space. This is only confirmed for the open subspace

$$\Gamma_\epsilon = \{\omega = (\mu, \alpha) \in \Omega : \mu < \frac{1}{2} + \epsilon, \ \delta < \alpha < \sigma\}.$$

for some sufficiently small $\epsilon > 0$, as chosen below. Given any such ϵ, we know from Proposition 1 and 4) that the system will become eventually absorbed (in finite time) by this set unless it starts at $(1, \sigma)$, a rest point. Thus, in establishing orbit convergence, we may restrict without loss of generality to initial conditions $\omega_0 \in \Gamma_\epsilon$.

We now compute the divergence of $\mathbf{v} = (v_1, v_2)$ for states $\omega \in \Gamma_\epsilon$. This is simply given by the trace of the Jacobian

$$D\mathbf{v}(\mu, \alpha) = \begin{pmatrix} \frac{\partial v_1}{\partial \mu}(\omega) & \frac{\partial v_1}{\partial \alpha}(\omega) \\ \frac{\partial v_2}{\partial \mu}(\omega) & \frac{\partial v_2}{\partial \alpha}(\omega) \end{pmatrix} \qquad (6)$$

with

$$\frac{\partial v_1}{\partial \mu}(\omega) = \kappa_1 \{-f(\alpha) + f(\alpha - \theta) - 2f(\alpha - \delta)$$
$$+ 2\mu [f(\alpha - \delta) - f(\alpha - \theta) + f(\alpha) - f(\alpha - \sigma)]\}$$

$$\frac{\partial v_1}{\partial \alpha}(\omega) = \kappa_1 \{-\mu^2 f'(\alpha - \sigma) - \mu(1 - \mu) [f'(\alpha) - f'(\alpha - \theta)] + (1 - \mu)^2 f'(\alpha - \delta)\}$$

$$\frac{\partial v_2}{\partial \mu}(\omega) = \kappa_2 [\theta - 2\delta + 2\mu(\sigma + \delta - \theta)]$$

$$\frac{\partial v_2}{\partial \alpha}(\omega) = -\kappa_2$$

where note that the vector field is continuously differentiable for all $\omega \in \Gamma_\epsilon$ (in particular the function $f(\cdot)$) since $\delta < \alpha < \sigma$.

First, we obviously have:

$$\frac{\partial v_2}{\partial \alpha}(\omega) = -\kappa_2 < 0.$$

12

As for the other term in the trace of the Jacobian, it can be re-written as:

$$\frac{\partial v_1}{\partial \mu}(\omega) = \kappa_1 \left\{ -f(\alpha) - 2f(\alpha - \delta) + 2\mu \left[f(\alpha - \delta) + f(\alpha) \right] \right\}$$

since $f(\alpha - \theta) = f(\alpha - \sigma) = 0$ for all $\omega \in \Gamma_\epsilon$. Suppose we had $\mu \leq 1/2$. Then,

$$\frac{\partial v_1}{\partial \mu}(\omega) \leq -\kappa_1 f(\alpha - \delta) < 0.$$

Obviously, the negative sign of $\frac{\partial v_1}{\partial \mu}(\omega)$ still applies if $\mu < \frac{1}{2} + \epsilon$, for some sufficiently small $\epsilon > 0$. Thus, for such a choice of ϵ, the trace of $D\ \mathbf{v}(\mu, \alpha)$ (i.e., the divergence of the vector field) can be ensured negative in Γ_ϵ, independently of the parameters κ_1 and κ_2. In view of our former discussion, this completes the proof of the Theorem. ∎

The previous result indicates that, given any initial conditions, there is always a uniquely associated long-run state of the process. This conclusion naturally leads us to what is the central question addressed in this paper: When can it be ensured that *some* extent of cooperation will *robustly* prevail in such long-run state?

Generically (in the underlying characteristics of the environment), an answer to this question is contained in Theorems 2 and 3 below. The key consideration turns out to be whether the switching-probability function $f(\cdot)$ is sufficiently responsive at $\chi = 0$, i.e., at situations of "exact satisfaction." If it is (as reflected by the magnitude of the right-hand gradient), then some partial cooperation will consolidate in the long run unless the system starts exactly at the full-defection state $\hat\omega = (0, \delta)$. If, on the contrary, the function $f(\cdot)$ is not responsive enough, then $\hat\omega = (0, \delta)$ is asymptotically stable, thus being "robust," at least in the face of small enough perturbations.

Theorem 2 *Assume* $\frac{df}{d\chi^+}(0) > \frac{f(\delta)}{\theta - 2\delta}$. *Then, for all initial conditions* $\omega_0 \neq \hat\omega = (0, \delta)$, *the unique limit state* $\omega^* = (\mu^*, \alpha^*) = \lim_{t \to \infty} \omega(t \mid \omega_0)$ *has* $\mu^* > 0$ *and, therefore,* $\alpha^* = \bar\pi(\mu^*) > \delta$.

Theorem 3 *Assume* $\frac{df}{d\chi^+}(0) < \frac{f(\delta)}{\theta - 2\delta}$. *Then, the rest point* $\hat\omega = (0, \delta)$ *is asymptotically stable.*[11]

Proof of Theorem 2. First, we have to address the issue raised by the non-differentiability of the function $f(\cdot)$ at $\chi = 0$ and its implications on the differentiability of

$$v_1(\mu, \alpha) = \kappa_1 \left\{ \begin{array}{c} -\mu^2 f(\alpha - \sigma) - \mu(1 - \mu) f(\alpha - 0) \\ +\mu(1 - \mu) f(\alpha - \theta) + (1 - \mu)^2 f(\alpha - \delta) \end{array} \right\}.$$

---

[11] As standard, the rest point $\hat\omega$ is said to be asymptotically stable if: (a) it is Liapunov Stable (roughly, the system remains always within any pre-specified neighborhood of this point if it starts sufficiently close), and (b) there is some neighborhood of $\hat\omega$ such that every orbit whose initial conditions lie in it converges to this point.

On the one hand, note that for all $\omega = (\mu, \alpha) \in \Omega$, since $\delta \leq \alpha \leq \sigma$, the first and third terms above can be ignored since they are identically zero. To clarify the considerations pertaining to the other two terms, it is useful to define the functions $h_\pi : [\delta, \sigma] \to I\!\!R$, for each $\pi \in \{0, \delta\}$ as follows:

$$h_\pi(\alpha) = f(\alpha - \pi).$$

Then, we have

$$h_0'(\alpha) = f'(\alpha)$$

and we make:

$$h_\delta'(\alpha) = \frac{df}{d\chi^+}(\alpha - \delta)$$

for all $\alpha \in [\delta, \sigma]$. With these specifications, the above functions are differentiable on $[\delta, \sigma]$. Re-writing (1) as follows:

$$\dot{\mu} = \kappa_1 \left\{ -\mu(1 - \mu) h_0(\alpha) + (1 - \mu)^2 h_\delta(\alpha) \right\} \tag{7}$$

we may now view $v_1 \; (\cdot)$ – and, therefore, the vector field $\mathbf{v} = (v_1, v_2)$ – as continuously differentiable on Ω.

In view of these considerations, the asymptotic stability of $\hat{\omega} = (0, \delta)$ can be evaluated in terms of the eigenvalues of the Jacobian of $\mathbf{v}(\cdot)$ – as the vector field has been reformulated – evaluated at it. We compute:

$$D\mathbf{v}(0, \delta) = \begin{pmatrix} -k_1 \, f(\delta) & k_1 \frac{df}{d\chi^+}(0) \\ k_2 \, (\theta - 2\delta) & -k_2 \end{pmatrix}. \tag{8}$$

The two eigenvalues of the above matrix are real. If $\frac{df}{d\chi^+}(0) > \frac{f(\delta)}{\theta - 2\delta}$, one (and only one) of them is positive. In this case, $\hat{\omega} = (0, \delta)$ is a saddle point (and, therefore, locally unstable). Since the eigenvector associated to this latter eigenvalue points outwards of the state space, it follows that, in fact, *no* initial conditions $\omega_0 \neq \hat{\omega}$ will lead the system to converge to $\hat{\omega}$. (See Figure 1 for an illustration of the local dynamics of the system around $\hat{\omega}$.) In view of Theorem 1, we may then conclude that the orbits must then converge to some state $\omega^* = (\mu^*, \alpha^*)$ with $\mu^* > 0$. This completes the proof of the Theorem. ∎

Proof of Theorem 3. From the considerations explained in the proof of Theorem 3, the desired conclusion simply follows from the fact that if $\frac{df}{d\chi^+}(0) < \frac{f(\delta)}{\theta - 2\delta}$, the determinant of the Jacobian (8) is positive. Therefore, since its trace is always negative, all of its eigenvalues are real negative numbers, thus inducing the asymptotic stability of $\hat{\omega}$. ∎

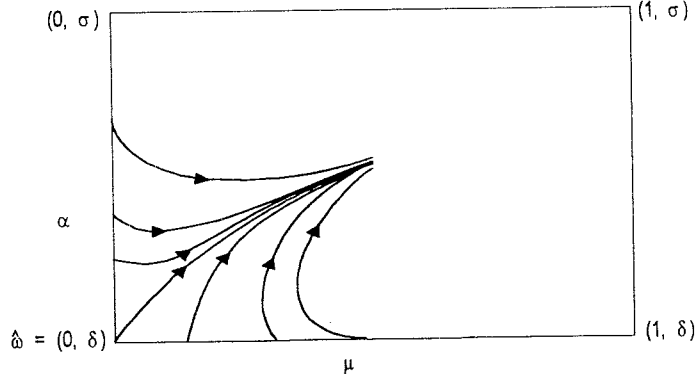Two further points are worth making in connection with the former results.

14

Figure 1: Local behavior around $\hat{\omega}$ in the unstable case.

Firstly, note that an obvious but interesting feature of any long-run state displaying partial cooperation (cf. Theorem 1) is that, even though the aggregate (or average) social configuration remains stationary, the underlying process is in perpetual flux. In particular, there is a continuous process of strategy adjustment carried out by a certain fraction of players which become dissatisfied every period due to unlucky matching outcomes.

Secondly, it is important to understand that even if the conditions contemplated by Theorem 3 are met, this does *not* imply that the full-defection state $\hat{\omega}$ is globally absorbing. Far from this: the simulations conducted in the Section 4 illustrate that, even under those circumstances, long-run states exhibiting partial cooperation will generally have a basin of attraction with positive measure.

Finally, we address an issue on which Theorem 2 remains silent: Is its prediction of partial cooperation unique? In general, it is not difficult to see that this will not be the case, unless one contemplates specific requirements on the different components of the model; specifically, on the function $f(\cdot)$ and the payoff structure. To end our analytical discussion of the model, we identify some illustrative sufficient conditions which guarantee such uniqueness.

Proposition 2 *Assume* $\theta - \sigma - \delta > 0$ *and let the function* $f(\cdot)$ *be strictly increasing and concave with* $f'''(\cdot) \geq 0$. *Then, there exists at most one rest point* $\omega^* = (\mu^*, \alpha^*)$ *with* $\mu^* > 0$.

Proof. Let the functions $G_1(\cdot)$ and $\bar{\pi}(\cdot)$ be as in (5) and (3), respectively. From (1)-(2), it is clear that any rest point of the dynamics (μ^*, α^*) must satisfy

$$G_1(\mu^*, \bar{\pi}(\mu^*)) = 0.$$

15

Define the function $\psi : (0, 1) \to I\!R$ as follows:

$$\psi(\mu) \equiv \frac{1}{1 - \mu} G_1(\mu, \bar{\pi}(\mu))$$

Then,

$$\psi(\mu) = -\frac{\mu^2}{1 - \mu} f(\bar{\pi}(\mu) - \sigma) - \mu f(\bar{\pi}(\mu)) + (1 - \mu) f(\bar{\pi}(\mu) - \delta).$$

Since $\bar{\pi}(\mu) \leq \sigma$, $f(\bar{\pi}(\mu) - \sigma) = 0$, and the above expression can be simplified and re-arranged to read as follows:

$$\psi(\mu) = f(\bar{\pi}(\mu) - \delta) - \mu \left[f(\bar{\pi}(\mu) - \delta) + f(\bar{\pi}(\mu)) \right].$$

To confirm the desired conclusion, it is enough to verify that the function $\psi(\cdot)$ is strictly concave in the interval $(0, \frac{1}{2})$ – recall Theorem 1. We differentiate:

$$\begin{aligned} \psi'(\mu) &= f'(\bar{\pi}(\mu) - \delta)\,\bar{\pi}'(\mu) - f(\bar{\pi}(\mu) - \delta) - f(\bar{\pi}(\mu)) \\ &\quad -\mu \cdot \bar{\pi}'(\mu) \left[f'(\bar{\pi}(\mu) - \delta) + f'(\bar{\pi}(\mu)) \right]. \end{aligned}$$

And differentiating again:

$$\begin{aligned} \psi''(\mu) &= f''(\bar{\pi}(\mu) - \delta) \left[\bar{\pi}'(\mu) \right]^2 + f'(\bar{\pi}(\mu) - \delta)\,\bar{\pi}''(\mu) \\ &\quad - \left[f'(\bar{\pi}(\mu) - \delta) + f'(\bar{\pi}(\mu)) \right] \bar{\pi}'(\mu) - \left[f'(\bar{\pi}(\mu) - \delta) + f'(\bar{\pi}(\mu)) \right] \bar{\pi}'(\mu) \\ &\quad -\mu \cdot \bar{\pi}''(\mu) \left[f'(\bar{\pi}(\mu) - \delta) + f'(\bar{\pi}(\mu)) \right] - \mu \left[\bar{\pi}'(\mu) \right]^2 \left[f''(\bar{\pi}(\mu) - \delta) + f''(\bar{\pi}(\mu)) \right]. \end{aligned}$$

which can be re-written as follows

$$\begin{aligned} \psi''(\mu) &= \left[\bar{\pi}'(\mu) \right]^2 \left\{ f''(\bar{\pi}(\mu) - \delta) - \mu \left[f''(\bar{\pi}(\mu) - \delta) + f''(\bar{\pi}(\mu)) \right] \right\} \quad (9) \\ &\quad - f'(\bar{\pi}(\mu)) \left[2\bar{\pi}'(\mu) + \mu\bar{\pi}''(\mu) \right] + (1 - \mu) f'(\bar{\pi}(\mu) - \delta)\,\bar{\pi}''(\mu) \\ &\quad - 2 f'(\bar{\pi}(\mu) - \delta)\,\bar{\pi}'(\mu) \end{aligned}$$

Since $f(\cdot)$ is assumed concave and $f'''(\cdot) \geq 0$, we have that

$$0 \geq f''(\bar{\pi}(\mu)) \geq f''(\bar{\pi}(\mu) - \delta),$$

and

$$f''(\bar{\pi}(\mu) - \delta) - \mu \left[f''(\bar{\pi}(\mu) - \delta) + f''(\bar{\pi}(\mu)) \right] \leq f''(\bar{\pi}(\mu) - \delta) - 2\,\mu\, f''(\bar{\pi}(\mu) - \delta)$$

which implies that the first term of (9) is non-positive for all $\mu \in (0, \frac{1}{2})$. On the other hand, concerning the second term of this expression, we have:

$$\begin{aligned} 2\bar{\pi}'(\mu) + \mu\bar{\pi}''(\mu) &= 2(\theta - 2\delta) - 6\mu (\sigma + \delta - \theta) \\ &= (\sigma - \delta) + (2\sigma - \delta) \end{aligned}$$

16

which, in view of the assumptions that $f(\cdot)$ is strictly increasing, implies that the second term of (9) is negative (again, for $\mu < \frac{1}{2}$) since $\sigma > \theta/2$. Furthermore, its third term is also negative since

$$\bar{\pi}''(\mu) = 2(\delta + \sigma - \theta) < 0,$$

again from the restrictions contemplated on payoffs. Finally, the fact that the last term of (9) is obviously negative, implies that $\psi''(\mu) < 0$ for all $\mu \in (0, \frac{1}{2})$, as desired. ∎

A scenario which satisfies the conditions contemplated by Proposition 2 is partly the focus of the simulations carried out in the next section (cf. Scenario A below).

4 Simulations

Our former analysis has characterized (generically) those conditions under which we should expect some extent of cooperation to prevail in the long run. Even though, as established by Proposition 1, such long-run level of cooperation can never involve more than half of the population, it is obviously interesting to explore in more detail its particular magnitude as well as its dependence on the different parameters of the model. To do this analytically, one would need to impose stringent conditions on the different components of the model – most crucially, on the switching function $f(\cdot)$. Instead, we choose to resort to numerical simulations since, for the present illustrative purposes, they would seem to allow a more intuitive and transparent discussion.

All our simulations will be conducted for a family of switching functions of the following form:

$$f(\chi) = \frac{\chi^q}{v + \chi^q}.$$

This family is parametrized by q and v, where $q \in \mathbb{N}$ and $v \in \mathbb{R}_{++}$. Clearly, given any such q and v, $f(\chi) \in (0,1)$ for all $\chi > 0$. When $q = 1$, the function $f(\cdot)$ is a concave function whose right-hand slope at $\chi = 0$ becomes arbitrarily large as $v \to 0$. On the other hand, if $q > 1$, the function is of sigmoidal type, with a zero slope at $\chi = 0$, and approaching a step function (with a single unit step at zero) as $q \to \infty$.

It follows from Theorems 2 and 3 that the behavior of the system must be qualitatively different if $q = 1$ or $q > 1$. Therefore, we organize our simulations along these lines in what follows.

4.1 Scenario A: $q = 1$

In this scenario, the fact that $f(\cdot)$ is strictly increasing and concave with $f'''(\cdot) > 0$ would allow us to rely on the uniqueness of interior rest points established by Proposition 2, provided that the payoffs of the game are chosen to satisfy the required conditions. In order to have these conditions satisfied, we shall consider the following payoff table throughout our simulations:

	C	D
C	$7,7$	$0,13$
D	$13,0$	$4,4$

Table 2

This choice of payoffs also implies that the hypothesis of Theorem 2 is satisfied. That is,

$$\frac{df}{d\chi^+}(0) > \frac{f(\delta)}{\theta - 2\delta}$$

or equivalently:

$$v > \frac{\delta(\theta - 2\delta)}{3\delta - \theta} \tag{10}$$

for all values of v, since the RHS of the above expression is negative.

Consequently, the state $\hat{\omega} = (0, \delta)$ is *not* locally stable and every (robust) long-run state must display a positive fraction of cooperators. (In fact, such a state is unique, as explained above, by virtue of Proposition 2.)

If, instead, the strict inequality converse to (10) applied for some v, then $\hat{\omega}$ would be asymptotically stable for this case and, in fact, uniquely so.[12] Thus, given the convergence established by Theorem 1, we would have that, generically, all trajectories would converge to $\hat{\omega}$, thus leading to trivially uninteresting dynamics. It is this motivates focusing on the case where (10) holds.

For this case, Figures 2(a)-2(f) show how the dynamics of the process responds to increases in v (i.e. a lower "responsiveness to dissatisfaction").

[Figures 2(a)-2(f): see the end of the text]

First, we find that there is always a single (robust) limit outcome. In view of our former discussion, this is as expected. Secondly, we also observe that the frequency of cooperators in the corresponding limit states decreases monotonically with v. Again, this behavior should come to no surprise since it is fully in accordance with an intuitive understanding of the model. Specifically, note from the law of motion for μ:

$$\dot{\mu}(t) = \kappa_1(1 - \mu)\left\{(1 - \mu)\left[f(\alpha) - \delta\right] - \mu f(\alpha)\right\}$$

that, for any (μ, α), the ratio between the flows in and out of cooperation is:

$$\frac{(1 - \mu)\,f(\alpha - \delta)}{\mu\,f(\alpha)},$$

which is a decreasing function of v. As a consequence, it is intuitively clear that the long-run frequency of cooperators should decrease as v increases.

In a different respect, it seems also of some interest to explore the dependence of dynamics of the process on alternative payoff structures. Since, for the reasons explained above, we would still want to preserve both (10) and the conditions contemplated by Proposition 2, a natural way to proceed

---

[12]This follows from the fact that if there were also any interior and asymptotically stable rest point, there would also have to be another (unstable) one, thus contradicting Proposition 2.

19

is to consider a family of games which represent a "scaled version" of the one considered above. That is, games with payoff tables of the form

	C	D
C	$7y, 7y$	$0, 13y$
D	$13y, 0$	$4y, 4y$

Table 3

with $y \in I\!R_{++}$. Obviously, these payoff tables satisfy the desired conditions for all y. Focusing on such parametrized family of games, we now argue that the same effect induced on the dynamics by an increase in v is achieved by a similar proportional *decrease* in y.

To confirm this claim, denote by $f_v(\chi)$ the switching function associated with some given parameter v. Then, we have that, for any real v, y, and v:

$$f_{y \cdot v}(y \cdot \chi) \equiv f_v(\chi).$$

Based on this identity, it is easy to see that, as long as v/y remains constant, the paths of the system associated to two different values y' and y'' are trivially isomorphic. Specifically, to any path $[\mu'(\cdot), \alpha'(\cdot)]$ corresponding to y' one may associate the path $[\mu''(\cdot), \alpha''(\cdot)]$ corresponding to y'' with

$$\mu''(\cdot) = \mu'(\cdot); \quad \alpha''(\cdot) = \frac{y''}{y'} \alpha'(\cdot).$$

Therefore, the role of y and v in the model is fully reciprocal and symmetric.

4.2 Scenario B: $q > 1$

When $q > 1$, the slope of the switching function $f(\cdot)$ at $\chi = 0$ vanishes. Therefore, in view of Theorem 3, we conclude that the state $\hat{\omega} = (0, \delta)$ is asymptotically stable. However, since the function $f(\cdot)$ is no longer concave (as explained it has a sigmoidal shape), Proposition 2 no longer applies and, therefore, there could be other robust (i.e., asymptotically stable) rest points lying in the interior of the state space. In fact, this is shown to be the case in the simulations presented in Figures 3(a)-3(f) and 4(a)-4(f) for two different values of q ($q = 2, 5$).[13]

[Figures 3(a)-3(f) & 4(a)-4(f): see the end of the text]

---

[13]The simulations are conducted for a Prisoner's Dilemma with payoffs as displayed in Table 2, i.e. with $y = 1$. The same considerations as before can be applied to changes in y, the magnitude which now has to remain constant being v/y^q.

In these two cases, we find that, for small enough values of v, the interior of the state space is partitioned (except for a zero-measure separating boundary) into two different basins of attraction. One of them corresponds to the no-cooperation state $\hat{\omega} = (0, \delta)$. The other basin includes all those trajectories which converge to a unique interior rest point exhibiting a positive degree of cooperation.

The maximum value of v for which the simulations show the existence of two different basins of attraction is $v = 2.5$ for $q = 2$ and $v = 6$ for $q = 5$. For larger values of v, all interior trajectories of the system converge to state $\hat{\omega}$. Thus, in the simulations conducted, higher values of q strengthen the tendency towards cooperation. This reflects the fact that higher values of q increase "agents' responsiveness", thus reinforcing the key mechanism underlying the rise of cooperation.

Finally, we also observe that, provided a robust interior rest point exists (i.e. the value of v does not exceed the corresponding upper bound), increases in v have the twin effect of decreasing the level of cooperation attained, as well as shrinking its corresponding basin of attraction. This simply extends to the present context ($q > 1$) an intuition analogous to that already encountered for $q = 1$.

5 Summary and concluding comments

In this paper, a stylized aspiration-based model of bounded rationality has been applied to a large-population context where players are matched to play a Prisoner's Dilemma. Under certain conditions on payoffs, every path of adjustment and learning has been shown to converge to some given social configuration. In this sense, therefore, the model has been found to induce a clear-cut long-run prediction from every initial condition.

The central concern of the paper has been to characterize those conditions under which a positive fraction of cooperators is obtained in the long run as the (unique) limit outcome of social adjustment paths. This has been seen to hinge upon the local behavior of the switching-probability function at points of zero dissatisfaction. If it is sufficiently responsive (i.e., players react with a steep enough gradient to small degrees of dissatisfaction), then all adjustment paths starting in the interior of the state space lead to a positive fraction of long-run cooperation (never more than half). Otherwise, the state where nobody cooperates is asymptotically stable, thus displaying some local robustness in the face of small perturbations.

The focus of the paper has been on the Prisoner's Dilemma since this game represents a natural benchmark for which the implications of our approach can be usefully compared with other alternative models of bounded rationality and social learning found in the literature. Next in our research agenda is the study of other types of games for which aspiration-based learning may lead to interesting conclusions. As a first step, we are presently studying simple coordination games where the main issue involved becomes one of equilibrium selection. Preliminary analysis suggests that considerations very different from the usual dichotomy of efficiency *vs.* risk dominance (e.g. a certain notion of "spite" built upon *relative* payoff comparisons) drive equilibrium selection in this context.

Figure 2.a q=1 v=0.1
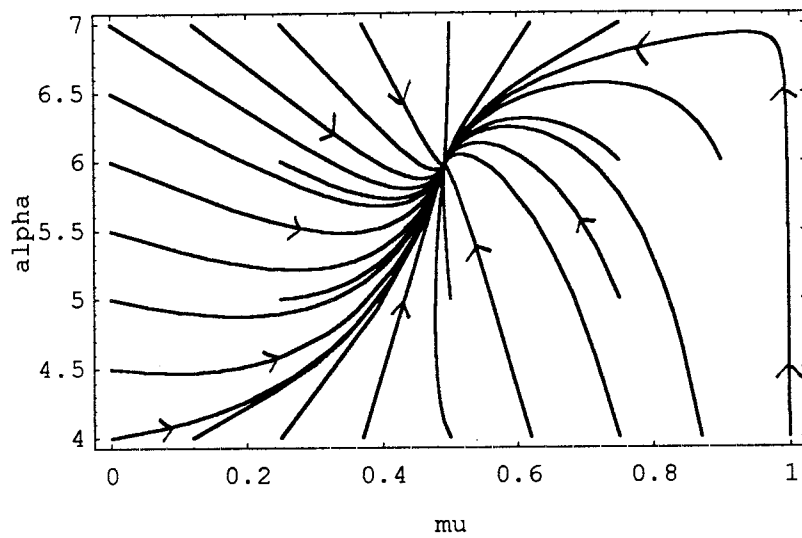
Figure 2.b q=1 v=0.5
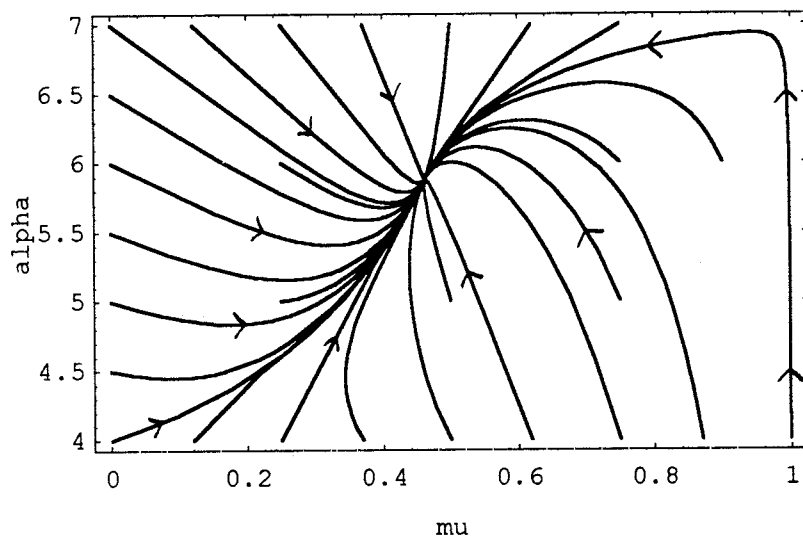
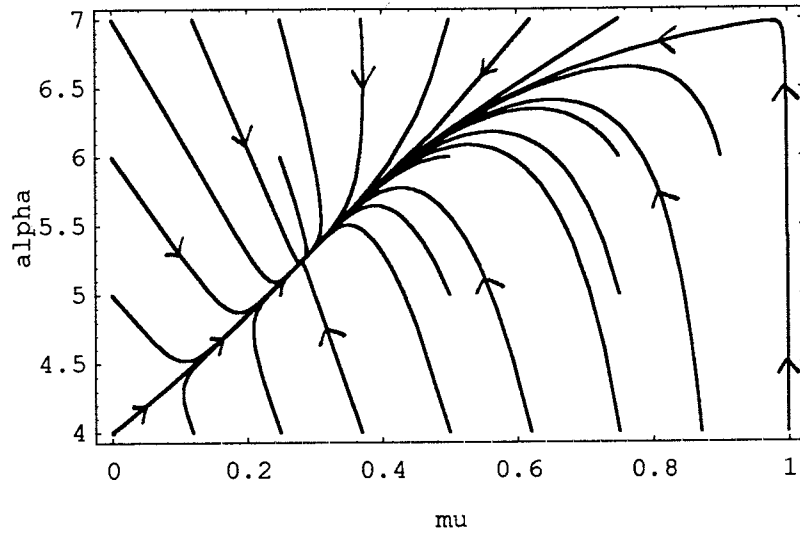Figure 2.c q=1 v=1
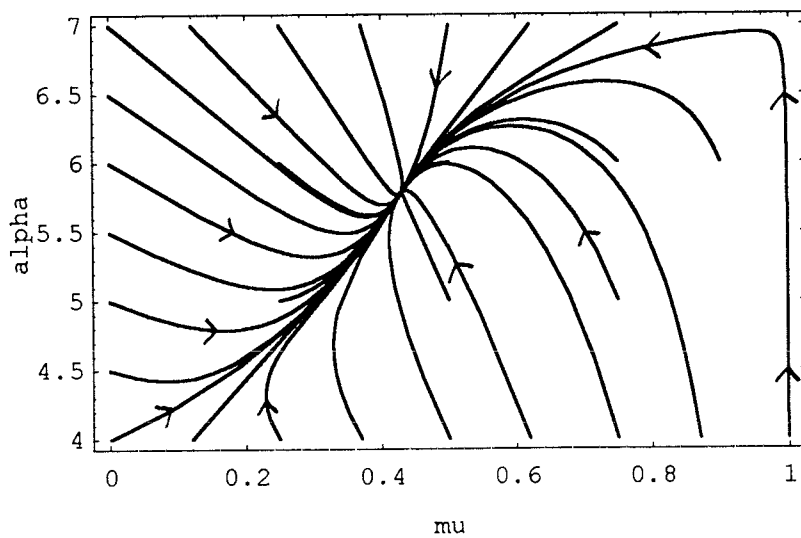
Figure 2.d q=1 v=5
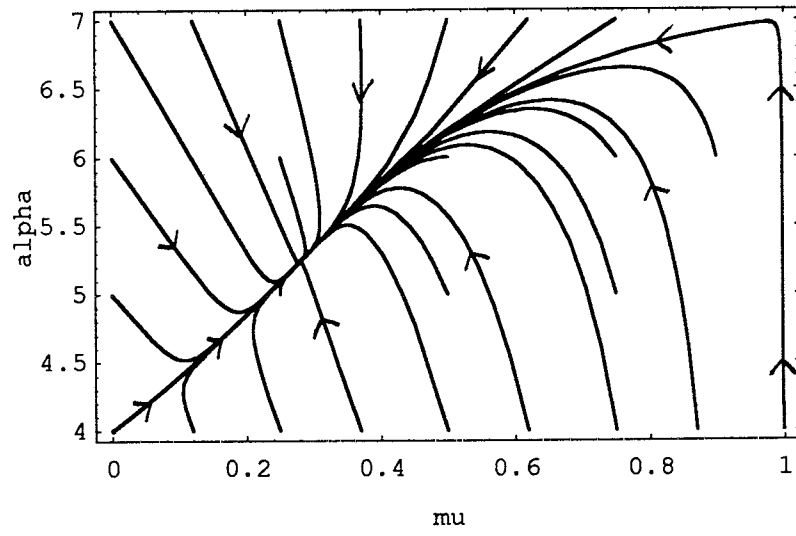
Figure 2.e q=1 v=10

Figure 2.f q=1 v=20

24

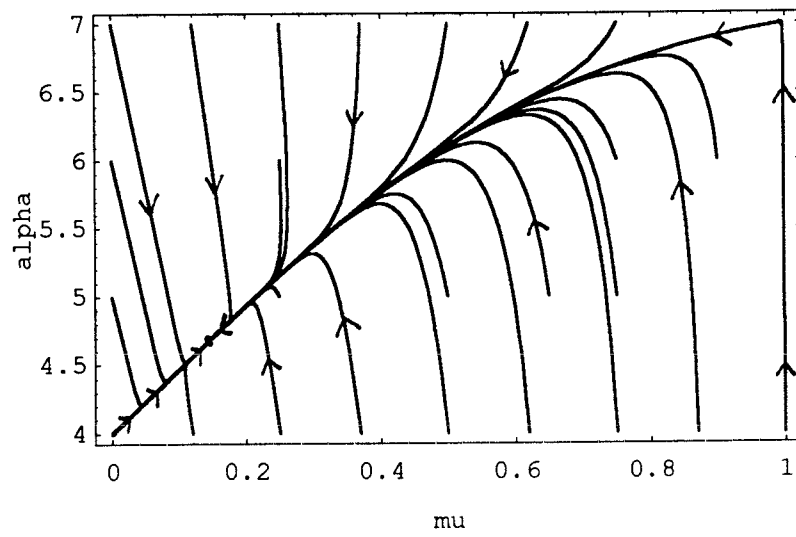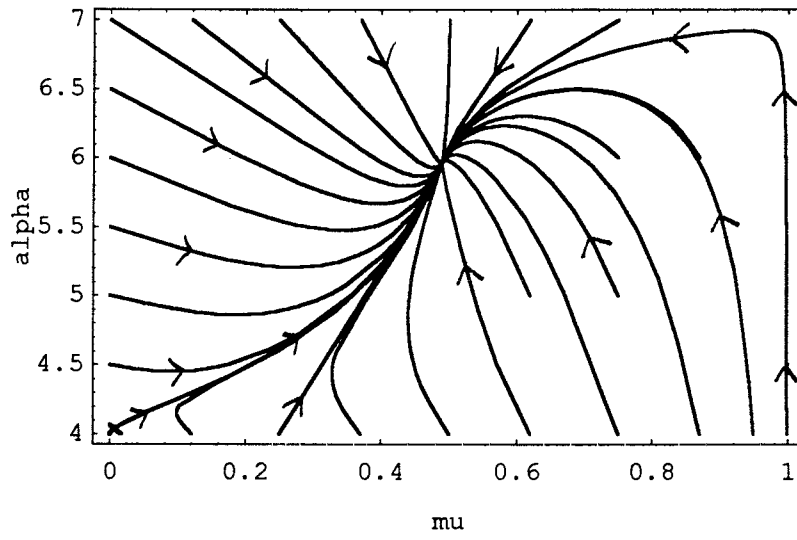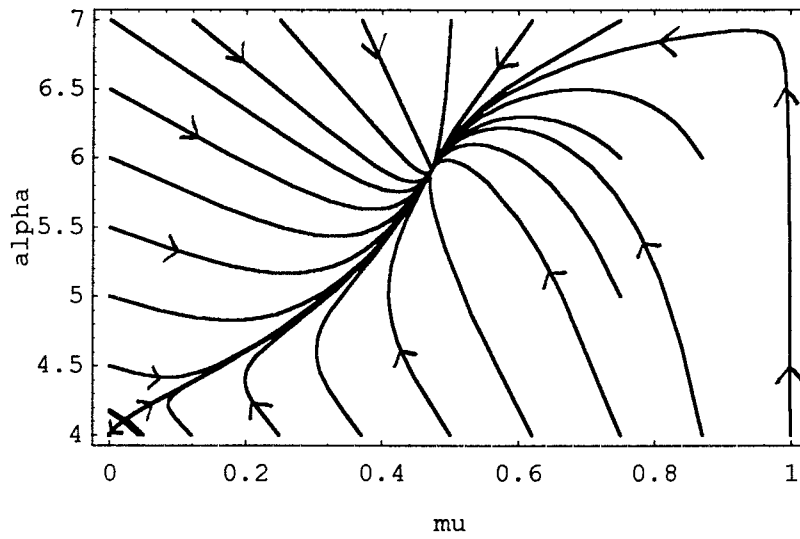Figure 3.a q=2 v=0.2

Figure 3.b q=2 v=0.5

Figure 3.c q=2 v=1

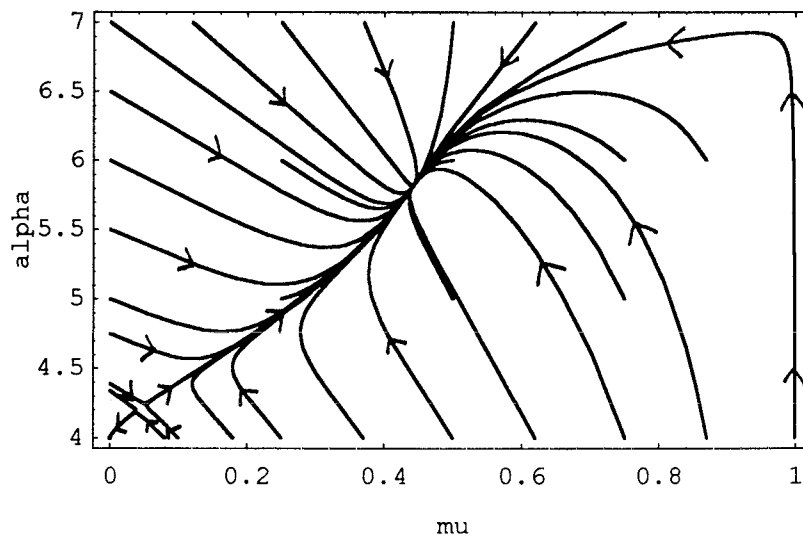## Figure 3.d q=2 v=2

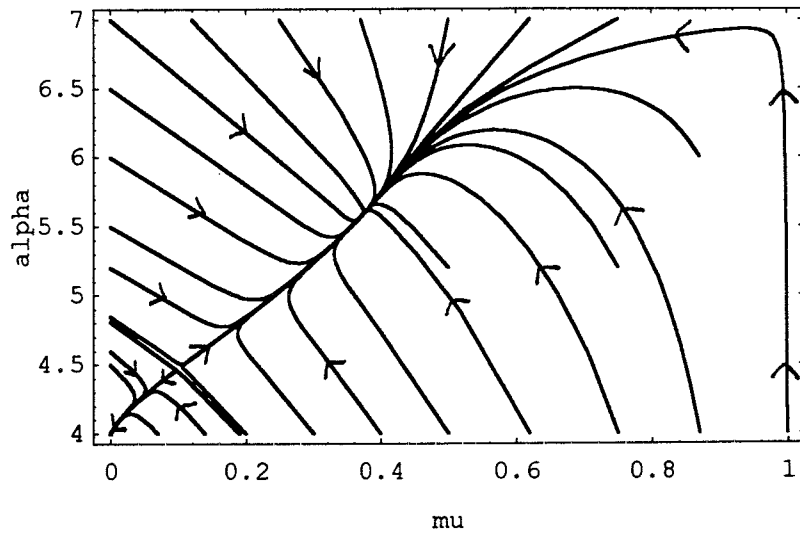## Figure 3.e q=2 v=2.5

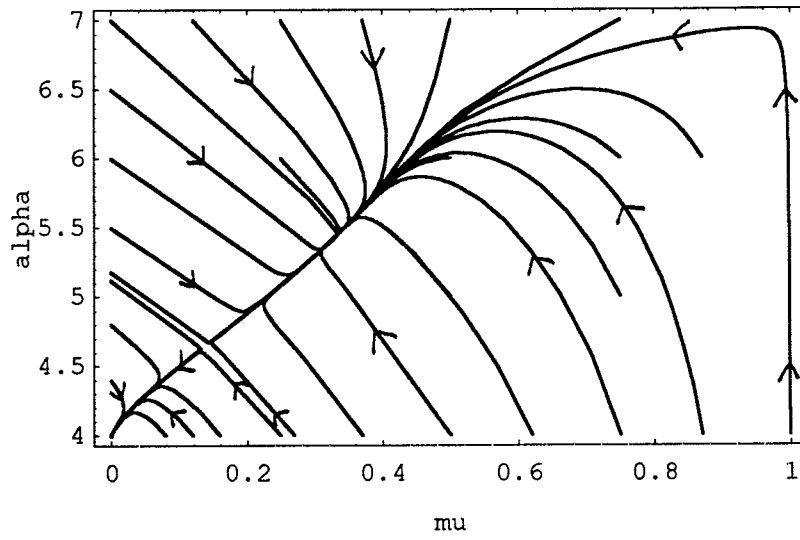## Figure 3.f q=2 v=3
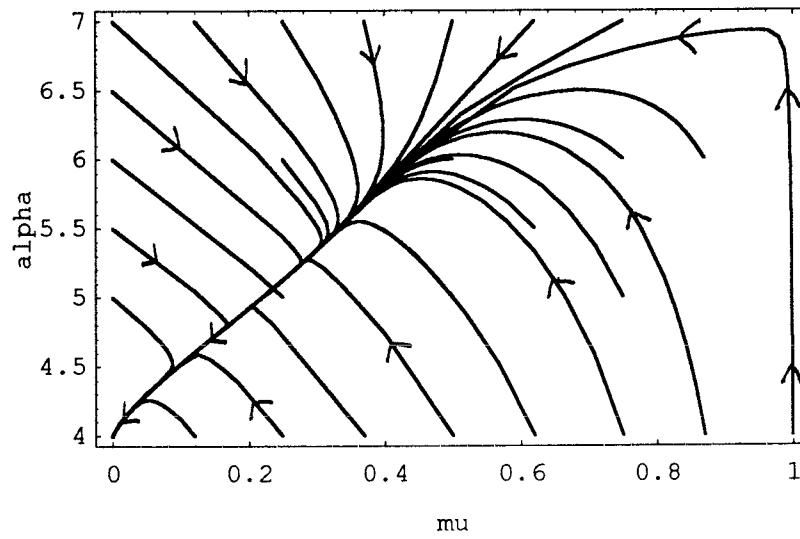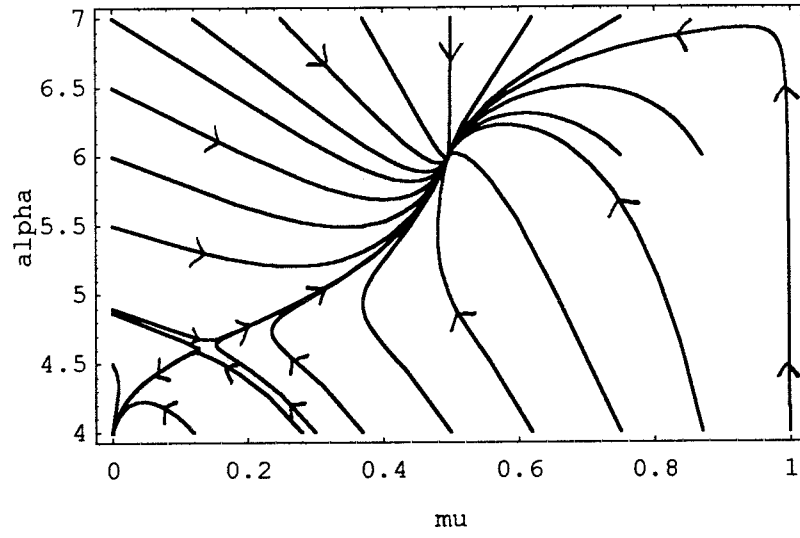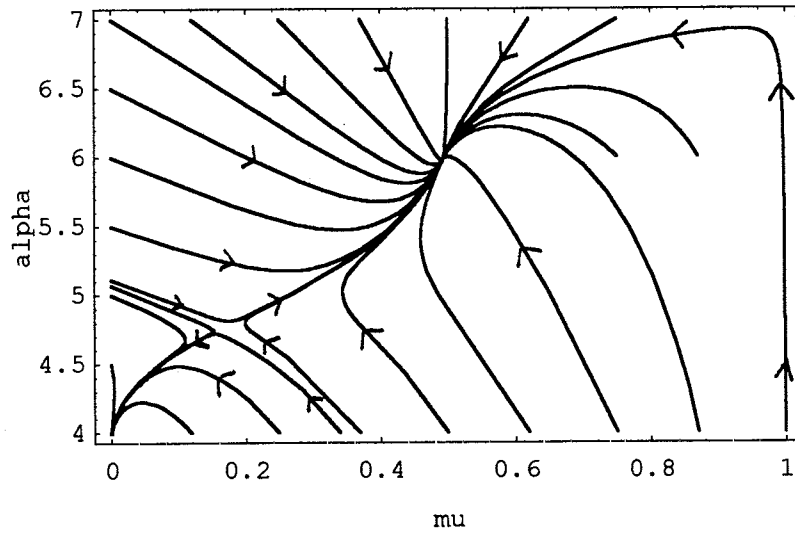
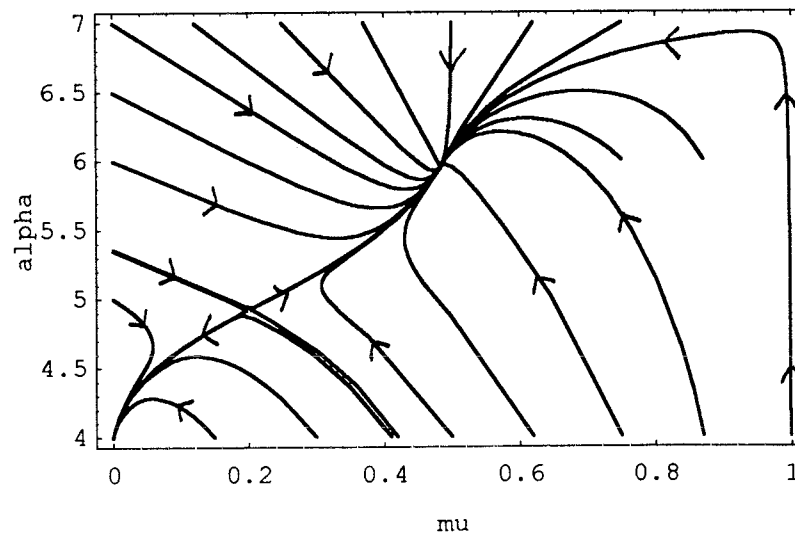Figure 4.a q=5 v=0.5

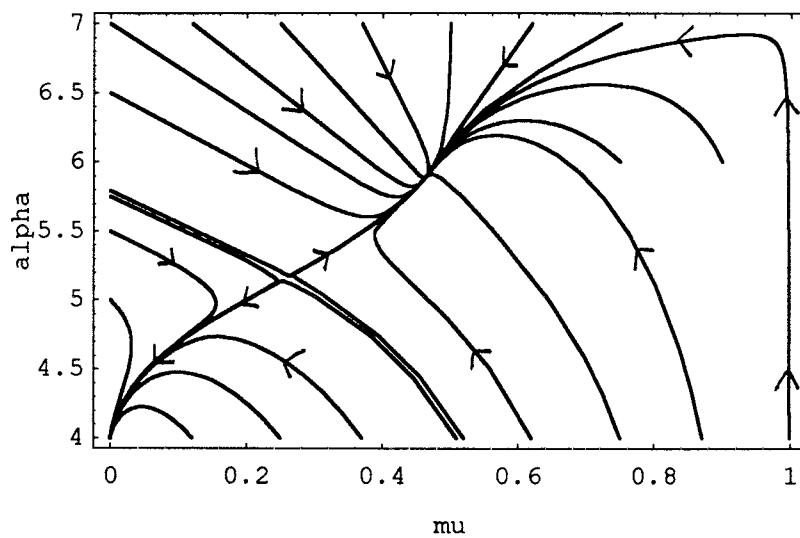Figure 4.b q=5 v=1

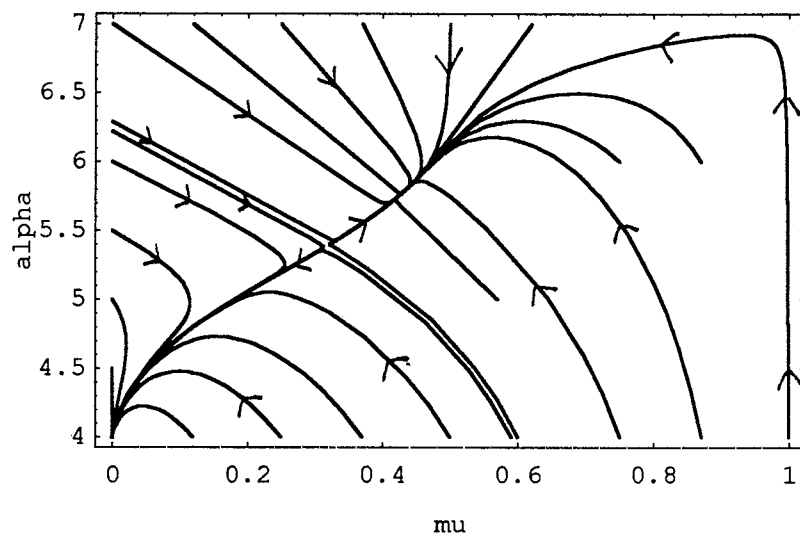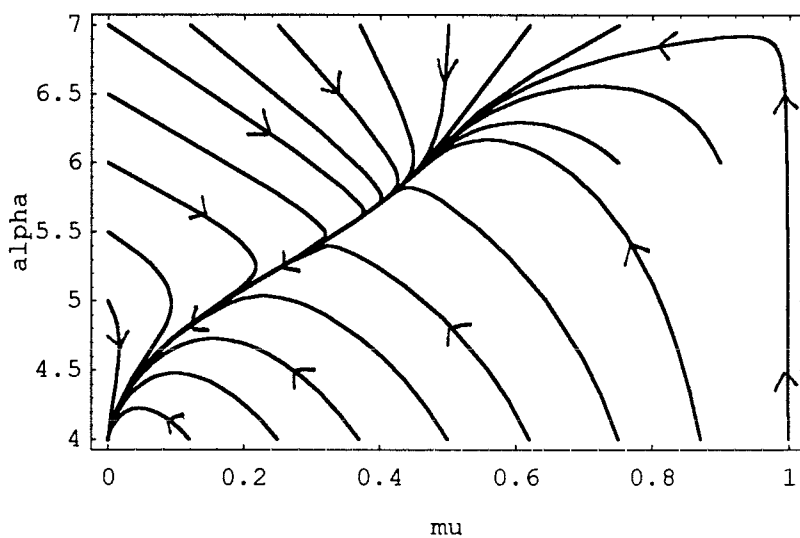Figure 4.c q=5 v=2

Figure 4.d q=5 v=4

Figure 4.e q=5 v=6

Figure 4.f q=5 v=8

References

[1] Arnold, V.I. (1973): *Ordinary Differential Equations*, Cambridge: The MIT Press.

[2] Axelrod, R. (1984): *The Evolution of Cooperation*, New York: Basic Publ.

[3] Bendor, J., D. Mookherjee and D. Ray (1992): "Aspirations, adaptative learning and cooperation in repeated games," Discussion Paper, Planning Unit, Indian Statistical Institute, New Delhi. Revised version: June, 1995.

[4] Binmore, K. and L. Samuelson (1992): "Evolutionary stability in games played by finite automata," *Journal of Economic Theory* **57**, 278-305.

[5] Binmore, K. and L. Samuelson (1993): "Muddling through: noisy equilibrium selection," mimeo, University College, London.

[6] Börges, T. and R. Sarin (1994): "Learning through reinforcement and the Replicator Dynamics," mimeo, University College, London, and Texas A&M University.

[7] Bush, R. and F. Mosteller (1955): *Stochastic Models of Learning*, New York: Wiley and Sons.

[8] Corchón, L. and A. Mas-Colell (1996): "A note on the stability of best reply and gradient systems with applications to imperfectly competitive models," *Economic Letters* **51**, 59-65, .

[9] Ellison, G. (1994): "Cooperation in the Prisoner's Dilemma with anonymous random matching", *Review of Economic Studies* **61**, 567-88.

[10] Friedman, J. (1977): *Oligopoly and the Theory of Games*, Amsterdam: North Holland Publishing Co.

[11] Fudenberg, D. & E. Maskin (1990): "Evolution and cooperation in noisy repeated games," *American Economic Review* **80**, 274-90.

[12] Gilboa, I. and D. Schmeidler (1995): "Case-Based Decision Theory," *Quarterly Journal of Economics* **110**, 605-40.

[13] Hirsch, M. and S. Smale (1974): *Differential Equations, Dynamical Systems, and Linear Algebra*, New York: Academic Press.

[14] Hofbauer, J. and K. Sigmund (1988): *The Theory of Evolution and Dynamical Systems*, Cambridge: Cambridge University Press.

[15] Kandori, M. (1992): "Social norms and community enforcement", *Review of Economic Studies* **59**, 63-80.

[16] Karandikar, R., D. Mokherjee, D. Ray, and F. Vega-Redondo (1995): "Evolving aspirations and cooperation," mimeo, Indian Statistical Institute, Boston University, and University of Alicante.

[17] Keenan, D. and T. Rader (1985): "Market dynamics and the Law of Demand," *Econometrica* **53**, 465-71.

[18] Kreps, D., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma", *Journal of Economic Theory* **27**, 245-52.

[19] Kim, Y. (1995): "Satisficing, cooperation and coordination," mimeo, Dept. of Economics, Queen Mary and Westfield College, University of London.

[20] Mokherjee, D. and B. Sopher (1994): "Learning behavior in experimental matching pennies games," *Games and Economic Behavior* **7**, 62-91.

[21] Neyman, A. (1985): "Bounded complexity justifies cooperation in the finitely repeated Prisoner's dilemma," *Economic Letters* **19**, 227-29.

[22] Pazgal, A. (1995): "Satisficing leads to cooperation in mutual interest games," mimeo, Kellog School of Managament, Northwestern University.

[23] Roth, A. and I. Erev (1993): "Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term," *Games and Economic Behavior*, forthcoming.

[24] Rubinstein, A. (1986): "Finite automata play the repeated Prisoner's Dilemma," *Journal of Economic Theory* **39**, 83-96.

[25] Selten, R. (1991): "Evolution, learning, and economic behavior," *Games and Economic Behavior* **3**, 3-24.

[26] Selten, R. and R. Stoecker (1986): "End behavior in sequences of finite Prisoner's Dilemma supergames," *Journal of Economic Behavior and Organization* **7**, 47-70.

[27] Vega-Redondo, F. (1996): *Evolution, Games, and Economic Behaviour*, Oxford: Oxford University Press.