Popularity of Reinforcement-Based and Belief-Based Learning Models: An Evolutionary Approach

Marcin Dziubiński\textsuperscript{b,1}, Jaideep Roy\textsuperscript{a,\*}

\textsuperscript{a}Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK,
Tel.: +44-121-4158362, Fax: +44-121-4147977

\textsuperscript{b}Faculty of Economics, University of Cambridge, Sidgewick Avenue, Cambridge CB3 9DD, UK,
Tel: +44-122-3335280

Abstract

In an evolutionary model, players from a given population meet randomly in pairs each period to play a coordination game. At each instant, the learning model used is determined via some replicator dynamics that respects payoff fitness. We allow for two such models: a belief-based best-response model that uses a costly predictor, and a costless reinforcement-based one. This generates dynamics over the choice of learning models and the consequent choices of endogenous variables. We report conditions under which the long run outcomes are efficient (or inefficient) and they support the exclusive use of either of the models (or their co-existence).

Keywords: Co-evolution, Best-response, Aspirations, Coordination games

JEL: D01, D03, D70

1. Introduction

Experiments demonstrate the power of learning models over static equilibrium notions in explaining actual play. These models fall under two broad classifications, namely, belief-based and reinforcement-based. In belief-based models, it is postulated that players use ‘predictors’ to form beliefs over the likely play of others, and choose actions based on expected payoffs. In contrast, reinforcement-based models do not adhere to such beliefs; rather, a player’s actions earn reinforcement or inertia based upon the payoffs they obtain, and over time the player adjusts his play so that actions with higher payoff experiences are used more frequently.

Mookherjee and Sopher [1994] were among the first to compare the performance of these two classes of models. They demonstrated through experiments that in relatively more complex games, reinforcement-based models explain actual play better than belief-based ones, a finding confirmed by Erev and Roth [1998] for aggregate behaviour. In contrast, Feltovich [2000] uses data from a large number of independent experiments,
including some of his own. His study suggests that belief-based models may not do so badly in general and that the relative abilities of these two classes of learning models depend on both the game used and the criterion of goodness of fit. Camerer and Ho (1999) demonstrate on the other hand that the best performing model is the one which combines both these features of learning, thereby opening up two interesting possibilities of either that individual players use a combined model or that players are heterogeneous in terms of the learning model they use so that aggregate (or statistically significant) behaviour is best described by a combined model. Such heterogeneity has also been reported in Cheung and Friedman (1997), albeit within the class of belief-based models. They show that majority of players use belief-based models and within them some use short memory while the rest use long memory (fictitious play) ones. Hopkins (2002) also compares these two learning models to find that the expected paths of play under both these models can be written as a perturbed form of a common evolutionary replicator dynamics, thus having similar asymptotic behaviour.

These (and other) studies ignite the following question: if different players use different learning models, how does the selection of learning models evolve? The influential work by Brock and Hommes (1997) (BH henceforth) initiated a related theoretical research and proposed what they called an adaptively rational equilibrium in an evolutionary framework. In their work, a player makes a rational decision between a sophisticated but costly (rational expectations) predictor, and a naive but costless one (naive expectations) and then according to the predictions (of the chosen predictor) of future play, plays a best response. They apply this meta model to a simple linear cobweb economy to study the dynamics of the evolution of the choice of these predictors and the consequent actions that determine equilibrium prices in the market. While the outputs (or forecasts) generated by these predictors are functions of past information, the performance (or fitness) measure of each predictor is the net realized payoffs in the most recent period that can be publicly observed. Based on these performance measures, players make rational decisions regarding which predictor to follow at each instant.

We adopt the general approach of BH to address the question we have emphasized above. However there are some differences between the two approaches. First, in BH, a model of learning is essentially a pair consisting of a predictor and a rule that suggests how to use it, where this rule is fixed to be a best response. Hence, in BH, while all learning models are belief-based, they differ in the degree of sophistication of the used belief (as in some sense similar to the models explored in Cheung and Friedman (1997)). In our case, there are two very different learning models: one that uses a costly predictor of other’s strategies and a best response rule to play accordingly; the other learning model is naive, and so unlike any of the learning models in BH, does not involve any use of predictors (and hence costless) but follows an experience-weighted attraction rule.

On the other hand, the rule by which players choose between the two models is similar to that in BH as in both the set ups, costless public information about the performance

---

2 See also Camerer et al. (2008) for popularity of reinforcement and belief learning models.
3 The terms predictor and learning strategy are used interchangeably in their article.
of the two models for the most recent history is used to determine this choice. The second point of departure is in the strategic environment. While BH and a series of related works concentrate on economic and financial markets that call for predicting the future market price, we study the problem of coordination failure that typically arises in environments with multiple Nash equilibria. This allows us to analyse an array of economic and social environments that demand the ability of players to coordinate their actions. In this respect, instead of analysing a specific economic model, we directly study an abstract $2 \times 2$ coordination game.

The list of works on how players solve the problem of coordination failure is large. See for example Van Huyck et al. (1990, 1991), Roth and Schoumaker (1983), Isaac et al. (1989), Cooper et al. (1990), Banks et al. (1988) and Crawford (1995). For an excellent summary of earlier work on this, see Ochs (1995). In a $2 \times 2$ coordination game where equilibria are Pareto ranked, the payoff dominant equilibrium may or may not be risk dominant. Several existing works in evolutionary game theory (see for example Kandori et al. (1993) and Young (1993)) suggest that in this class of games, and with best response type rules, the risk dominant equilibrium is more likely to be observed as such outcomes have larger basins of attraction. However, there is also a vast experimental literature (see for example Friedman (1996) and Bouchez and Friedman (2008)) that demonstrates that neither payoff-dominance nor risk dominance is either necessary or sufficient to predict behaviour, and hence lend a fair amount of support to the alternative theoretical hypothesis of Bergin and Lipman (1996) where the payoff dominant outcome becomes more likely to be observed in the long run. In the coordination game we study, the payoff dominated equilibrium is also risk dominated and our study shows that even then the payoff dominated equilibrium can be selected.

More directly related to our work are two papers, namely, Josephson (2009) and Juang (2002). In Josephson (2009), a model stochastic adaptation is studied with a population comprising of best repliers, better repliers, and imitators. Individuals select strategies according to the learning rule assigned to them a priori and play with respect to a sample from a finite history of past play. A set of sufficient conditions for convergence to minimal closed sets under better replies and selection of a Pareto dominant set of such outcomes are provided. Juang (2002) studies a $2 \times 2$ coordination game where the payoff dominant equilibrium is risk dominated and played amongst a finite population under random matching. As in our case, it allows for rule evolution by examining two learning rules: myopic best response (like ours) and a naive imitation rule. It shows that if players can change rules via experiments (that is, use the other existing rule with some probability, a clear departure from our case), then the Pareto efficient equilibrium strictly dominates the risk dominant one. While these works are closely related, in neither of them are players consciously adopting learning rules based on their relative performance, something that is central to our model.

More specifically, we study an evolutionary model where players from a large population are randomly matched in each period to play a strategic form stage game that is time invariant. There are two learning models available, a single-period memory belief-based best response model (the BR model in short), and a naive reinforcement-based model of
aspirations with a fixed social aspiration level (the AD model in short).\footnote{There are many studies of evolution of play where players are assumed to be AD (see Bendor et al. (2001b) for example and for an excellent survey, see Bendor et al. (2001a). See also Dixon (2000) and Palomino and Vega-Redondo (1999) for large population environments using AD type models.} The BR model we study is more akin with the notion of partial best-response play, rather than fictitious play\footnote{See books on evolutionary games by Weibull (1997), Vega-Redondo (1997), Samuelson (1997), Fudenberg and Levine (1998) and the remarkable lecture series of Young (2004) for several models and results on best response learning models.} Players switch each period from their current action to a best response to the aggregate statistic of play from the previous period, obtaining which requires some costly investment. This cost can also be attributed to the relative complexity of the model that requires computing best responses, a possibility that seems to be consistent with the findings of Mookherjee and Sopher (1994) for more complex games. Players consciously choose which model to follow and accordingly take their actions. The popularity of the two models is then established via a replicator that depends upon the relative performances of the two.\footnote{The literature on the evolution of preferences (see Samuelson (2001) for an introduction and the 2001 special issue in the Journal of Economic Theory, Volume 97, Issue 2) takes the stand that individual (subjective) preferences determine actions and these preferences may differ from the actual (objective) payoffs which on the other hand determine fitness of a given preference. The focus of these studies is on the evolution of these preferences while in our case, the focus is on the evolution of the learning models. There is a subtle but important distinction between this literature and our approach. While evolution of preferences addresses issues concerning the theory of why to play, evolution of learning models is more akin to a theory of how to play them.} We focus on $2 \times 2$ pure coordination games whose static Nash equilibria are strict-Pareto ranked (we call H the pure strategy that constitutes the Pareto efficient equilibrium, while the other pure strategy is called L).\footnote{See Section 4 for a discussion on other stage games.} We model this environment as a two dimensional dynamic system. The two dimensions we attend are the fraction of players in the population who use the AD model and the fraction of these AD players who play H.

With aggregate statistics as the common source of belief formation, in almost any situation it follows that all BR players will take the same pure strategy. Our analysis then shows the following. Consider the situation where all (if any) BR players play H. If H is also sufficiently popular amongst the AD players and the social aspiration is high, then, regardless of the cost parameter of the BR model, the system always converges to a unique stable rest point. In this rest point, the entire population uses the AD model and plays H. With low aspirations, on the other hand, there is a continuum of stable restpoints in each of which it is still true that the BR model is not observed. However, now the population, which uses exclusively the AD model, can only achieve partial coordination. Hence, co-existence of the two learning models and achieving Pareto efficiency are together impossible.

We then move to the situation where all (if any) BR players play L. If the social aspiration is high, then depending upon the cost of the BR model, the society converges to two possible long run outcomes. If the cost is sufficiently small, then there is a unique stable rest point where only the BR model survives. If the cost is larger, but not too high, the unique long run outcome is where both the BR and the AD models are used. However, when the cost is even higher, there are no restpoints; the system either gets stuck in an eternal flow of the BR players switching between the two actions H and L.
or they all play $H$ and we are back in the previous case where there is then a unique convergence point where the entire population uses the AD model and plays $H$. Lastly, if the social aspiration is low, then there is a continuum of stable restpoints. In each of these restpoints, the entire population uses the AD model with relatively little coordination on $H$.

The rest of the paper is structured as follows. We develop a formal model for our environment in Section 2. The results are stated in Section 3 with proofs moved to the Appendix. In Section 4, we discuss our results and their implications to other and more general environments with multiple equilibria.

2. The model

There is a continuum of players and time is continuous. At each instant $t \in [0, +\infty)$, players are randomly matched in pairs to play a coordination game presented in the Figure 1, where $0 < \delta < \sigma$. Without any loss of generality, we normalize the game assuming that $\delta + \sigma = 1$. The game has two pure strategy Nash equilibria, namely $(H, H)$ and $(L, L)$ and a unique mixed strategy Nash equilibrium where $H$ is played with probability $\delta$. These equilibria are Pareto ranked.

\[ \begin{array}{cc}
H & L \\
\sigma, \sigma & 0, 0 \\
L & 0, 0 \quad \delta, \delta \\
\end{array} \]

At any $t$, players can be classified into two distinct types, the set of Best Response (BR) players $R$, and the set of Aspiration driven (AD) players $A$. We use $a$ to denote the fraction of AD players in the population and $\mu$ to denote the fraction (with respect to $A$) of AD players playing $H$.

Simply put, the BR players are myopic and use the following belief-based model: they choose a best response to the most recent history of the frequency of actions in the society. We assume that obtaining reliable information about this frequency comes at a cost $\varrho$ such that $0 < \varrho < \delta$.

The AD players use the following reinforcement model. Denote the common social aspiration level for aspiring players by $\alpha$ which is fixed throughout. We restrict attention to the case where $0 < \alpha \leq \sigma$ (i.e. the social aspiration level does not exceed the highest possible payoff). With this payoff aspiration, an AD player takes an action from the set $\{H, L\}$, and receives an individual payoff of $\pi \in \{0, \delta, \sigma\}$. The social aspiration $\alpha$ and the realization of his current payoff $\pi$ gives rise to an individual dissatisfaction level $\chi = \alpha - \pi$. The probability rate at which an AD player with dissatisfaction level $\chi$ changes his current strategy is given by a continuously differentiable function $f(\chi)$ that is strictly increasing (on $[0, +\infty)$) and satisfying

\[ f(\chi) \begin{cases} 
0 & \text{if } \chi \leq 0, \\
> 0 & \text{if } \chi > 0 
\end{cases} \]  

11 Arguably this is the first explicit model of learning in games; see Cournot (1838) and Cheung and Friedman (1997). It is a special case of the ‘cautious’ fictitious play model with a ‘one-period’ memory.
This captures the notion that if an AD player is satisfied with his current payoff (that is $\chi \leq 0$), then he sticks to his current action in the next instant; otherwise, he 'experiments' with the other available action with a positive probability.

The fact that all BR players use a common information set about the society’s most recent play is helpful for the tractability of the analysis. Consider an initial condition with given values of $a$ and $\mu$. Since then, the history is empty, let $0 \leq p \leq 1$ be the mass of BR players who start playing $H$ while the remaining of them play $L$. Then, barring the non generic case that involves equalities, in the next instance, the following must be true: all BR players either play $H$ with probability 1 or $L$ with probability 1. By respecting genericity, we restrict attention to initial conditions where either all BR players play $H$ or they all play $L$. Given this, it is straightforward to verify that starting from any situation where all BR players play $H$, the action $H$ remains to be the best response as long as $\sigma - a(1 - \mu) > 0$. Similarly, starting from any situation, where all BR players play $L$, the action $L$ remains to be a best response as long as $\delta - a\mu > 0$.

At each instant, each player can choose to change his learning model. This choice depends on the publicly observed difference between average payoffs of the two existing subpopulations, BR and AD. Given $a$, $\mu$ and the current common action $x \in \{H, L\}$ of the BR population, the average payoff of the AD population is denoted by $\bar{\pi}_x^a(\mu, a)$. Similarly denote the average payoff of the BR population at any instant by $\bar{\pi}_x^r(\mu, a)$.

Let
\[ \psi^x(a, \mu) = \bar{\pi}_x^a(\mu, a) - \bar{\pi}_x^r(\mu, a). \]

The probability rate at which a player changes his learning model is modelled by use of a strictly increasing (on $[0, +\infty)$) and continuously differentiable function $g$ satisfying
\[ g(z) \begin{cases} = 0 & \text{if } z \leq 0, \\ > 0 & \text{if } z > 0, \end{cases} \tag{2} \]

where $z$ reflects by how much the 'better-performing' rule is outperforming the other rule. Notice that continuous differentiability of $g$ implies that $g'(z) = 0$ for $z \leq 0$. These assumptions on $g$ are natural and reflect three things. Firstly, the higher is the difference between average payoffs of the two groups, the higher is the probability that a player with the inferior learning model will change his group. Secondly, if the difference between average payoffs is 0, then there is no incentive for players to change their models. Thirdly, the probability of this switch increases smoothly from the situation where the average payoffs of two groups are equal. Given $g(\cdot)$, the change in time of the fraction of AD

---

12It is unimportant if this frequency $p$ is an outcome of pure strategies or mixtures.
13If players could observe the average payoffs of both BR and AD populations individually, the population strategy distribution could be backed out without having to pay the explicit cost $\varrho$ of this information. However, we assume that players only observe the difference of the two and not each component of the difference. Thus they observe a single non-linear equation in two unknowns, namely, $\psi^x(a, \mu) = 0.6$, say. From that they cannot extrapolate $a$ and $\mu$ uniquely, required for them to determine their best response. In order to enable players to infer about the distribution of actions in the current population, we would require upon them to have finer social information that arguably can be costly as well.
14Analogous comments to the ones on $g$ hold for function $f$ as well. We would like to note that our results are independent of the exact nature of the functions $f(\cdot)$ and $g(\cdot)$, as long as they satisfy our general assumptions.
δ – The lower payoff in the coordination game  
σ – The higher payoff in the coordination game  
ϱ – The cost of BR model  
a – The fraction of AD players in the population  
µ – The fraction of AD players playing H  
α – Social aspiration level of AD players  
f(χ) – The probability rate of strategy change for AD players with given dissatisfaction level χ  
g(z) – The probability rate of learning model change, given the difference z of payoffs across the models

Table 1: A summary of the variables and parameters of the model.

players is given by

\[ \dot{a} = A^{\mu}(a, \mu) = \kappa_1 \left( -g(-\psi^x(a, \mu))a + g(\psi^x(a, \mu))(1 - a) \right), \quad (3) \]

where \( \kappa_1 > 0 \) is some arbitrary speed adjustment parameter. Notice that if \( \psi^x(a, \mu) > 0 \), then only the second component of the right hand side of Equation (3) is non-zero, and the fraction of AD players increases at the rate \( g(\psi^x(a, \mu))(1 - a) \). On the other hand, if \( \psi^x(a, \mu) < 0 \), then only the first component of the right hand side is non-zero, and the fraction of AD players increases at the rate \( -g(-\psi^x(a, \mu))a \).

Given the probability rate \( f \) defined above, we can now model the change in time of the fraction \( \mu \) of AD players playing H. There are three factors that affect \( \mu \):

(i). the switch of actions amongst current AD players who choose to remain AD,  
(ii). the mass of current AD players who chose to become BR and  
(iii). the mass of current BR players who choose to become AD.

First note that given \( g(\cdot) \), there can be no two-way flow in our model. Hence, at each instant, either no one changes his current learning model, or that either only some AD players become BR or some BR players become AD. In case of (i) above, the switch is entirely determined by the function \( f \). In case of (ii) the AD players who become BR are uniformly distributed over the set \{H, L\}. In case of (iii) the fraction \( \mu \) of the BR players who become AD start out by playing H while the remaining of them play L.

Thus we have the following expression for \( \mu \):

\[ \dot{\mu} = M^{\mu}(a, \mu) = \kappa_2 \begin{cases} f(\alpha - \delta)a(1 - \mu)^2 + f(\alpha)(1 - a)(1 - \mu) & \text{if } x = H, \\ f(\alpha - \delta)(1 - \mu)(1 - a\mu) - f(\alpha)\mu(1 - a) & \text{if } x = L, \end{cases} \quad (4) \]

where \( \kappa_2 > 0 \) is some arbitrary speed adjustment parameter.

We summarize all the variables and parameters of the model in Table 4.

In the rest of the paper, we ask the following question. Suppose the co-evolution of play and popularity of learning models is such that during the entire path, the BR players, if any, either play H or they play L. Where does the system converge to and what are the respective zones of attraction for these long run outcomes. In order to answer this question, we study the two dimensional dynamic system in the \( a - \mu \) plane characterized by the two equations of motion (3) and (4). This analysis is done in the following section.
3. The analysis

As discussed in the previous section, we attend to the situations where all BR players play either H or L. We study each of these two cases separately in the following subsections. In each of the cases, the only values that evolve are the fraction $a$ of AD players and the fraction $\mu$ of AD players playing H. The area of possible values of these parameters that will be considered is $\bar{Z} = [0, 1] \times [0, 1]$, with $(a, \mu) \in \bar{Z}$. In our analysis we will consider solutions starting from $(a, \mu) \in Z = (0, 1) \times [0, 1]$. Thus we assume that at the initial state both types of players exist (and so our model is not a model of mutation but replication).

3.1. The case of BR players playing H

We begin with the case that is relatively simple where we consider initial conditions such that all BR players play H. Intuitively, the properties of the rest points we identify in this section are not striking, though there are some interesting observations on the dynamics by which they are reached. For completeness, we will provide a full formal analysis of this case. Let

$$s^H(a, \mu) = \sigma (1 - a (1 - \mu)) - \delta a (1 - \mu) = \sigma - a (1 - \mu)$$

denote the difference between expected payoffs from playing H and L when the system is at the point $(a, \mu)$ and all BR players play H. As we explained in Section 2, if $s^H(a, \mu) > 0$, then all BR players continue playing H. The area $A_L = \{(a, \mu) \in \bar{Z} : s^H(a, \mu) \leq 0\}$ therefore constitutes the region in $\bar{Z}$ in which playing H by all BR players may not be sustained. On all the trajectories that never enter $A_L$, all BR players will keep playing H forever.

If all BR players play H, then the system of differential equations defined by (3) and (4) has the form:

$$\dot{a} = A^H(a, \mu) = \kappa_1 (g(\psi^H(a, \mu))(1 - a) - g(-\psi^H(a, \mu))a),$$
$$\dot{\mu} = M^H(a, \mu) = \kappa_2 (f(\alpha - \delta) a(1 - \mu)^2 + f(\alpha)(1 - \alpha)(1 - \mu)),$$

where

$$\psi^H(a, \mu) = a(1 - \mu)^2 - (1 - \mu)\sigma + \varrho.$$  

Further analysis will be divided into two cases: first, when the social aspiration level is high, that is $\delta < \alpha \leq \sigma$, and second, when it is low, that is $0 < \alpha \leq \delta$.

3.1.1. High aspiration level

Let $\delta < \alpha \leq \sigma$. We show that the system converges to a unique long run outcome in which all players adopt the AD model and play H. There are two main reasons behind this result: (i) when all BR players play H and the aspiration level is high, then the fraction of the AD players playing H is increasing and (ii) there is a positive cost of the BR model. Since all BR players play H, it follows that all AD players playing H who switch to L do so because they were matched with AD players playing H. The fraction of such players changing their strategy to L is equal to the fraction of AD players playing L who were matched with AD players playing H and then decided to change their strategy. Thus games between mismatched AD players do not affect the flow of usage
of different strategies within the AD population. The remaining AD players who change their strategy are those playing L and who were matched either with BR players or AD players playing L. In the second case this change is caused by the high level of social aspiration. Thus as long as all the BR players play H, the fraction $\mu$ of AD players playing H is increasing, converging to 1. Now, when $\mu$ is large enough ($> \delta$), then $s(a, \mu) > 0$ regardless of $a$. Hence starting from any state $(a, \mu)$ with such a level of $\mu$, all BR players keep playing H forever. If $\mu$ becomes sufficiently high ($> 1 - \varrho/\sigma$), the difference between the average payoffs of the AD and the BR populations become positive, as now the only significant difference between these payoffs is due to the cost of the BR model. As soon as this happens, the incentive to switch to the AD model increases significantly, driving the BR population to extinction. The phase plane diagram for the case of all BR players playing H and high aspiration level is presented in Figure 2 [in all such diagrams, restpoints are indicated by bold dots of lines when there is a continuum of them; please note that not all restpoints are stable].

Let $m \in \mathbb{R}$ and $\bar{Z}_m = \{(a, \mu) \in \bar{Z} : \mu > m\}$. We will show that if the initial conditions $(a_0, \mu_0)$ lie outside $A_L$, then any solution of the system of differential equations (5) and (6) converges to the unique (and asymptotically stable) rest point (1, 1) where all players are AD playing H. Moreover, for any initial conditions $(a_0, \mu_0) \in \bar{Z}_\delta$ and all BR players playing H, all BR players continue to play H forever. Hence any trajectory starting from within $\bar{Z}_{\delta}$ converges to the rest point (1, 1). This is stated formally below. The proof is moved to the Appendix.

**Proposition 1.** Consider any normalized coordination game such that $0 < \delta < \sigma < 1$. Suppose $a \in (\delta, \sigma]$ and $\varrho \in (0, \delta)$. Let $\bar{x} \in \bar{Z}$ and let $\phi_\varrho(t, \bar{x})$ be a solution of the system of differential equations (5) and (6). Then the following hold:

(i). If $\bar{x} \in \bar{Z}_\delta$, then $\phi_\varrho(t, \bar{x}) \in \bar{Z} \setminus A_L$ for all $t \geq 0$.

(ii). If $\phi_\varrho(t, \bar{x}) \in \bar{Z} \setminus A_L$, for all $t \geq 0$, then $\lim_{t \to +\infty} \phi_\varrho(t, \bar{x}) = (1, 1)$. 


3.1.2. Low aspiration level

Let \(0 < \alpha \leq \delta\). In this case \(f(\alpha - \delta) = 0\) and Equation (6) simplifies to

\[
\dot{\mu} = M^H(a, \mu) = \kappa_2 f(\alpha)(1 - a)(1 - \mu).
\] (8)

Unlike in the case of high aspiration level, there is no unique long run outcome when the aspiration level is low. The reason for this is as follows. Like in the case of high aspiration level, switch of strategies within the AD population resulting from miscoordinated game outcomes within this population do not affect the flow of usage of the strategies. The only games that could affect this are those in which AD players playing L are matched with BR players. Thus as long as BR population is not extinct, the fraction \(\mu\) of AD players playing H is increasing. However, if there are no BR players, that is \(a = 1\), then the value of \(\mu\) stabilizes, as the flow from H to L is the same as the flow from L to H within the AD population. Yet, when \(a = 1\), the flow from the AD population to the BR population is still possible, if the cost \(\rho\) of the BR model is not too high and the fraction of AD players playing H is neither too low nor too high. The phase plane diagram for the case of all BR players playing H and low aspiration level is presented in Figure 3.

![Phase plane diagram](image)

**Figure 3**: Phase plane diagram when all BR players play H and the aspiration level is low. There is a continuum of restpoints (bold line), where all players adopt the AD model, but the fractions of players playing H vary between the restpoints.

We will show that if the initial conditions \((a_0, \mu_0)\) lie outside \(A_L\), then any solution of the system of differential equations (5), (8) converges to the set of continuum stable rest points where all players are AD. Moreover, for any initial conditions \((a_0, \mu_0) \in Z_{>\delta}\) with all BR players playing H, all BR players continue to play H forever. Hence any trajectory starting from within \(Z_{>\delta}\) converges to a restpoint where all players are AD. This is stated formally below. The proof is moved to the Appendix.

**Proposition 2.** Consider any normalized coordination game such that \(0 < \delta < \sigma < 1\) and suppose \(\alpha \in (0, \delta]\). Let \(x \in Z\) and let \(\phi_\alpha(t, x)\) be a solution of the system of differential equations (5), (8). Then the following hold:
(i) If $\bar{x} \in \mathbb{Z}_{>\delta}$, then $\phi_v(t, \bar{x}) \in \mathbb{Z} \setminus A_L$ for all $t \geq 0$.

(ii) If $\phi_v(t, \bar{x}) \in \mathbb{Z} \setminus A_L$, for all $t \geq 0$, then:
   (a) If $\rho \in (0, \sigma^2/4)$, then $\lim_{t \to +\infty} \phi_v(t, \bar{x}) \in \{1\} \times ((\delta, \mu^H_1] \cup [\mu^H_2, 1])$, where

   \[
   \mu^H_1 = \delta + \frac{\sigma - \sqrt{\sigma^2 - 4\rho}}{2}, \quad \mu^H_2 = \delta + \frac{\sigma + \sqrt{\sigma^2 - 4\rho}}{2}.
   \]

   (b) If $\rho \in [\sigma^2/4, \delta)$, then $\lim_{t \to +\infty} \phi_v(t, \bar{x}) \in \{1\} \times (\delta, 1]$.

3.2. The case of BR players playing $L$

We now study the more interesting case where we consider the trajectories of the system with initial conditions such that all BR players play $L$. Let

$$ s^L(a, \mu) = \delta(1 - a) - \sigma a \mu = \delta - a \mu $$

 denote the difference between expected payoffs from playing $L$ and $H$ when the system is at the point $(a, \mu)$ and all BR players play $L$. As we explained in Section 2, if $s^L(a, \mu) > 0$ then all BR players keep playing $L$. The area $A_H = \{(a, \mu) \in \mathbb{Z} : s^L(a, \mu) \leq 0\}$ constitutes the region in $\mathbb{Z}$ in which playing $L$ by the BR players may not be sustained.

On all the trajectories that never enter $A_H$, the BR players will keep playing $L$ forever.

If all BR players play $L$, then the system of differential equations defined by (3) and (4) has the form:

\[
\dot{a} = A^L(a, \mu) = \kappa_1(g(\psi^L(a, \mu))(1 - a) - g(-\psi^L(a, \mu))a), \tag{9}
\]

\[
\dot{\mu} = M^L(a, \mu) = \kappa_2(f(\alpha - \delta)(1 - \mu)(1 - a\mu) - f(\alpha)\mu(1 - a)), \tag{10}
\]

where

$$ \psi^L(a, \mu) = a\mu^2 - \mu\delta + \rho. \tag{11} $$

As in the case of all BR players playing $H$ we divide our analysis into two cases: first, when the social aspiration level is high, that is $\delta < \alpha \leq \sigma$, and second, when it is low, that is $0 < \alpha \leq \delta$.

3.2.1. High aspiration level

Let $\delta < \alpha \leq \sigma$. We show that two long run outcomes are possible. In the first one, all players adopt the BR model while in the second, both models coexist. Moreover, if the cost of the BR model is above some threshold value $\bar{\rho}$, then one cannot sustain the play of $L$ by the BR population.

To understand why these scenarios are possible notice first that unlike in the case of all BR players playing $H$, the fraction $\mu$ of AD players playing $H$ does not necessarily increase. Like in the previous case, the fraction of AD players playing $H$ who decide to change their strategy after being matched with AD players playing $L$ is exactly offset by the fraction of AD players playing $L$ who decide to change their strategy after being matched with AD players playing $H$. However this time the fraction of AD players playing $L$ who decide to change their strategy after being matched with AD players playing $L$ can be offset by the fraction of AD players playing $H$ who decide to change their strategy after being matched with BR players (all of whom play $L$). Hence a situation may arise where there is a constant (and of equal size) switch of AD players between the two pure
strategies. Now, if the cost of using the BR model is sufficiently small (at or below some threshold value $\hat{\varrho}$), and the mass of AD players is not too large, then there will be a positive flow from the AD to the BR population. Since this flow does not affect the distribution of players using different pure strategies within the AD population, it will lead to a state where the AD population goes extinct. If the cost of using the BR model is intermediate ($\in (\hat{\varrho}, \check{\varrho})$), then it is possible that in a situation where the flow between the two pure strategies within the AD population is constant, the expected payoffs from using the AD and the BR models are the same. This causes the flow between populations of different learning models to stop and the two models co-exist. If the cost of the BR model is too high ($> \check{\varrho}$), then in any situation where the flow between two pure strategies within the AD population is constant, the fraction of BR players decreases and when it reaches a sufficiently low level, the fraction of AD players playing L, who decide to change their strategy after being matched with other AD players playing L, cannot be offset. This leads to a constant flow to H within the AD population and to a constant flow from the BR to the AD population. The phase plane diagrams for the three situations discussed above are presented in Figures 4 – 6.

Let $\mathcal{B} = \{(a, \mu) \in \mathcal{Z} : \psi^L(a, \mu) \leq 0 \text{ and } M^L(a, \mu) < 0\}$. Thus $\mathcal{B}$ is a region in $\mathcal{Z}$ where the payoff of the BR players is at least as good as the payoff of the AD players when the mass of the AD players changing their strategy to L is larger than the mass of AD players changing their strategy to H. Let $x \in (0, 1]$ and

$$\mathcal{P}_x = \{(a, \mu) \in \mathcal{Z} : \psi^L(a, \mu) \leq 0 \text{ and } a < x\}.$$ 

Thus $\mathcal{P}_x$ is the region where the payoff of the BR players is at least as high as the payoff of the AD players given that the mass of AD players is below $x$. The results discussed above are stated formally below. The proof is moved to the Appendix.

**Proposition 3.** Consider any normalized coordination game such that $0 < \delta < \sigma < 1$. Suppose $a \in (\delta, \sigma]$ and $\varrho \in (0, \delta)$. Let $\bar{x} \in \mathcal{Z}$ and let $\phi_\varrho(t, \bar{x})$ be a solution of the system

\[
\frac{d}{dt} \psi^L(a, \mu) = -h(a, \alpha, \delta) \psi^L(a, \mu) \quad \text{and} \quad \frac{d}{dt} M^L(a, \mu) = -h(a, \alpha, \delta) M^L(a, \mu) - \varrho M^L(a, \mu)
\]
Figure 5: Phase plane diagram when all BR players play L, the aspiration level is high and the cost of the BR model is intermediate. There is only one stable restpoint (the bold dot closer to the $a$ axis), where BR and AD players coexist.

Figure 6: Phase plane diagram when all BR players play L, the aspiration level is high and the cost of the BR model is high. There are no stable restpoints in this case and the mass of players adopting AD model eventually increases, as does the mass of AD players playing H.

of differential equations [7], [10].

(i). If

$$\varphi \leq \hat{\varphi} = \frac{\delta}{1 + h(\alpha, \delta)}, \quad \text{where} \ h(\alpha, \delta) = \frac{f(\alpha)}{f(\alpha - \delta)},$$

then there exists $a \in (0, 1)$ such that $B \subseteq P_a$ and for any $\vec{x} \in P_a$, $\phi_v(t, \vec{x}) \in \hat{Z} \setminus \hat{A}_1$, for all $t \geq 0$. Moreover, for any $\vec{x} \in \bar{Z}$ such that for all $t \geq 0$, $\phi_v(t, \vec{x}) \in \bar{Z} \setminus \bar{A}_1$, $\lim_{t \to +\infty} \phi_v(t, \vec{x}) = (0, 1/(1 + h(\alpha, \delta))).$
(ii). If \( \varrho > \hat{\varrho} \) and

\[
\varrho < \hat{\varrho} = \left( h(\alpha, \delta) - \sqrt{(h(\alpha, \delta) - 1)(h(\alpha, \delta) + 1 - \delta)} \right)^2,
\]

then there is a point \((a^*, \mu^*)\) with \((a^*, \mu^*) \in \bar{Z} \setminus A_H\), \(a^* \in (0, 1)\) and \(\mu^* \in (0, 1)\), and a region \(X\) containing \((a^*, \mu^*)\) and such that for any \(x \in X\) and for all \(t \geq 0\), it holds that \(\phi_v(t, x) \in X\) and \(\lim_{t \to +\infty} \phi_v(t, x) = (a^*, \mu^*)\). This case is possible iff

\[
h(\alpha, \delta) \neq \frac{\sqrt{4\delta + 1}}{2\delta}.
\]

(iii). If \( \varrho \geq \hat{\varrho} \), then for any \(x \in Z\) there is \(t > 0\) such that \(\phi_v(t, x) \in A_H\). This case is possible iff \(h(\alpha, \delta) > \frac{1}{2\sqrt{\delta-\delta}}\).

Notice that if the cost of the BR model is small \((\varrho < \hat{\varrho})\), then whenever a situation is reached where the expected payoff of the BR population is higher than the expected payoff of the AD population and the mass of AD players is sufficiently small \((a < \hat{a})\), the system will converge to the restpoint where the AD model is extinct. If the cost of the BR model is intermediate \((\varrho \in (\hat{\varrho}, \hat{\varrho})\)), then we cannot, in general, characterize the basin of attraction of the restpoint where both populations co-exist. Moreover, we could not rule out other forms of stability (viz. stable orbits) where both populations co-exist, but their relative size and the fraction of AD players playing H oscillate. These possibilities will depend on the exact nature of \(\varrho\) and the speed adjustment factors \(\kappa_1\) and \(\kappa_2\). Nevertheless, our results guarantee that some form of stability where both populations coexist is indeed possible and it is asymptotically stable. Notice that for any function \(f\) and any value of the lower payoff \(\delta\) there is a range of aspiration levels \((\delta, \hat{\alpha}) \subseteq (\delta, \sigma)\) such that the mixed population restpoint exists. This is because \(h(\alpha, \delta) \to +\infty\) when \(\alpha \to \delta^+\) and so for sufficiently small \(\alpha\) it will be that \(h(\alpha, \delta) > (\sqrt{4\delta + 1} + 1)/(2\delta)\). It can be that for higher values of \(\alpha\) the inequality guaranteeing existence of this restpoint is also possible. It all depends on the character of the function \(f\). It should be noted as well that the condition \(\varrho > \hat{\varrho}\) that ensures non-existence of the restpoint may be empty under appropriate form of the function \(f\) and values of the aspiration level \(\alpha\) and payoff \(\delta\).

### 3.2.2. Low aspiration level

Finally let \(0 < \alpha \leq \delta\). In this case \(f(\alpha - \delta) = 0\) and \(\{10\}\) simplifies to

\[
\dot{\mu} = M_{\text{L}}(a, \mu) = -\kappa_2 f(\alpha)(1 - a)\mu.
\]

When the aspiration level is low, as long as there are some BR players, the fraction of AD players playing H falls. This is because, like in the case of high aspiration level, switch in strategies within the AD population resulting from mis-coordinated game outcomes within this population do not affect the flow of usage of the strategies. The only games that could affect that are those in which AD players playing H are matched with BR players (all of whom play L). Thus as long as the BR population is not extinct, the fraction \(\mu\) of AD players playing H is decreasing. However, if there are no BR players, that is \(a = 1\), then the value of \(\mu\) stabilizes, as the flow from H to L is the same as the flow from L to H within the AD population. Moreover when \(a = 1\), the flow from the AD to the BR population is still possible if the cost \(\varrho\) of the BR model is not too high and the fraction \(\mu\) of AD players playing H is not neither too low nor too high. The
phase plane diagram for the case of all BR players playing L and low aspiration level is presented in Figure 7.

Let $x \in \mathbb{R}$ and

$$Z_{<x} = \{(a, \mu) \in \mathbb{Z} : \mu < x\}.$$  

We will show that if the initial conditions $(a_0, \mu_0)$ lie outside $A_H$, then any solution of the system of differential equations (9), (12) converges to a set of continuum stable restpoints where all players are AD. Moreover, for any initial conditions $(a_0, \mu_0) \in Z_{<\delta}$ with all BR players playing L, all BR players continue to play L forever. Hence any trajectory starting from within $Z_{<\delta}$ converges to a restpoint where all players are AD. This is stated formally below. The proof is moved to the Appendix.

**Proposition 4.** Consider any normalized coordination game such that $0 < \delta < \sigma < 1$ and $\alpha \in (0, \delta]$. Let $\bar{x} \in \mathbb{Z}$ and let $\phi_{\nu}(t, \bar{x})$ be a solution of the system of differential equations (9), (12). Then the following hold:

(i). If $\bar{x} \in Z_{<\delta}$, then $\phi_{\nu}(t, \bar{x}) \in \bar{Z} \setminus A_H$ for all $t \geq 0$.

(ii). If $\phi_{\nu}(t, \bar{x}) \in \bar{Z} \setminus A_H$, for all $t \geq 0$, then:

(a) If $\varrho \in (0, \delta^2/4)$, then $\lim_{t \to +\infty} \phi_{\nu}(t, \bar{x}) \in \{1\} \times ([0, \mu_1^L] \cup [\mu_2^L, \delta])$, where

$$\mu_1^L = \frac{\delta - \sqrt{\delta^2 - 4\varrho}}{2}, \quad \mu_2^L = \frac{\delta + \sqrt{\delta^2 - 4\varrho}}{2}.$$  

(b) If $\varrho \in [\delta^2/4, \delta)$, then $\lim_{t \to +\infty} \phi_{\nu}(t, \bar{x}) \in \{1\} \times [0, \delta)$.

4. Discussion

4.1. Our findings

A summary of all the results obtained in this study can be found in Table 2.
Initially all BR players play L

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stable Restpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0 &lt; \varrho &lt; \delta^2/4)</td>
<td>({1} \times [0, \mu_1^L] \cup [\mu_2^L, \delta])</td>
</tr>
<tr>
<td>(\delta^2/4 &lt; \varrho &lt; \delta)</td>
<td>({1} \times {0, \delta})</td>
</tr>
</tbody>
</table>

\(0 < \alpha \leq \delta\)

\(\delta < \alpha \leq \sigma\)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stable Restpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0 &lt; \varrho \leq \hat{\varrho})</td>
<td>((0, 1/(1 + h(\alpha, \delta))))</td>
</tr>
<tr>
<td>(\hat{\varrho} &lt; \varrho &lt; \bar{\varrho})</td>
<td>((a^<em>, \mu^</em>)) with (0 &lt; a^<em>, \mu^</em> &lt; 1)</td>
</tr>
<tr>
<td>(\bar{\varrho} \leq \varrho &lt; \delta)</td>
<td>No stable restpoints nor stable orbits with all BR players playing L</td>
</tr>
</tbody>
</table>

Initially all BR players play H

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stable Restpoints</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0 &lt; \varrho &lt; \sigma^2/4)</td>
<td>({1} \times (\delta, \mu_H^1] \cup [\mu_H^2, 1])</td>
</tr>
<tr>
<td>(\sigma^2/4 &lt; \varrho &lt; \delta)</td>
<td>({1} \times (\delta, 1])</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Condition</th>
<th>Stable Restpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0 &lt; \varrho &lt; \delta)</td>
<td>(1, 1)</td>
</tr>
</tbody>
</table>

Table 2: A summary of model behaviour for different initial conditions and values of parameters.

We have shown that when the social aspiration is high and when players initially following the BR model themselves play H, there is a basin of attraction \(\bar{Z}_{\geq \delta}\) (see Section 3.1.1) for a unique long run outcome where full coordination on the Pareto superior Nash equilibrium is achieved.\(^{15}\) While this convergence result is robust, it is mainly driven by the cost of the BR model. Nevertheless, the dynamics behind this convergence is interesting (see Figure 2): first the BR model gains popularity even to the extent that there is hardly any use of the AD model; but soon this leads to a sufficiently large fraction of the existing minority of AD players playing H, from when on, the popularity of the AD model starts growing and finally the point \((1, 1)\) is reached. It is important to note that the cost of the BR model may be arbitrarily small for the result to hold. This also supports the common convergence properties of the two models as in Hopkins (2002). The full extinction of the BR model can also be obtained when aspirations are low (see Figure 3), but now the long run population consisting of players using only the AD model may fail to achieve full coordination.

More interesting results appear in the case when initial conditions are such that the BR players play L. Starting from such a situation, if the social aspiration is high, then depending upon the cost of the BR model, the society converges to two possible long run outcomes. If the cost is sufficiently small (see Figure 4), then there is a unique stable rest point where only the BR model survives, implying full convergence to the Pareto inferior Nash equilibrium. It is important to note that this happens in a society where the AD model was built upon high aspirations (and if that model had survived, \(\hat{\varrho}\)We would like to note here that if in this case the initial conditions are outside \(\bar{Z}_{\geq \delta}\), then such a convergence is not guaranteed as the system may then enter and remain in the ‘switching region’ where the BR players ‘switch’ between H and L or transit for ever to the region \(A^L\).
then convergence to such an inefficient outcome could not be achieved). However, the relative ease of the BR model first destroys the popularity of the ambitious AD model with very high aspirations, and once that happens the system gets stuck in the inefficient equilibrium.

On the other hand, if the cost of the BR model is larger (see Figure 5), but not too high, there is a long run stable outcome where both learning models co-exist. Further, even though all players using the AD model in this restpoint aspire for high payoffs, a case when within the model convergence should select Pareto efficiency, the system may get stuck in a situation where a large majority of them play \( L \). This is due to the co-existence of BR players who keep playing \( L \) so that the AD population enters an eternal flux in which they keep switching their actions. What is interesting is that this balance is not fragile as the point is indeed asymptotically stable (see Figure A.8 in the appendix and Lemma 5 therein that constructs a basin of attraction for such a mixed-population restpoint).

Although we have not undertaken a full formal analysis of it, we note that not all trajectories starting from initial conditions with the BR players playing \( L \) would end up at this restpoint when the AD model comes with high aspirations as on some paths it may eventually become impossible to sustain the action \( L \) as the best response for the BR population so that they would then switch to the action \( H \). This becomes more likely when the cost of the BR rule is very high (see Figure 6). In this case, we have shown that there are no restpoints in the area where the BR players still play \( L \) and the system may then transit into the region \( \mathcal{Z}_{>\delta} \) to then converge to the restpoint where the entire population uses the AD rule and plays \( H \). It may also be that the system transits to a zone where the BR players play \( H \) but remains outside the region \( \mathcal{Z}_{>\delta} \), leading to more nuanced long run behaviour where the BR players keep switching between \( H \) and \( L \). With low social aspiration on the other hand (see Figure 7), there is a continuum of stable restpoints. In each of these restpoints, the entire population uses the AD model with relatively little (less than the fraction \( \delta \)) coordination on the Pareto superior outcome.

4.2. On Cycles

While we find the stable co-existence of learning rules an important result, one may ask if there can also be cyclical co-existence of the two, i.e., cycles where players’ “moods” switch from BR to AD and reversely (though existence of such mood switches is not sufficient for stable orbits to exist). Interestingly, such stable orbits could be possible if the assumption about a smooth change between AD and BR rules was changed from \( g'(0) = 0 \) to \( g'(0) > 0 \). In such a case the stable point with both rules coexisting would become an unstable source, but there would be a stable orbit around it to which all the trajectories leaving from the source would converge. This stable orbit would involve continuous change of fractions of players using the two heuristics.

Another possibility for cycles is near the region where BR players are indifferent between playing \( H \) and \( L \). The system might then enter a cycle were BR players oscillate between different fractions playing \( H \), which in turn could result in the fraction of AD players oscillating. We omitted the analysis of behaviour of the system in such cases, for two reasons: firstly, the system would become three dimensional (opening up the possibility of chaotic dynamics as well), with the fraction of BR players playing \( H \) being the third parameter of the system and, secondly, the system would depend too much on
the assumptions of the rate of the switch, thereby making such results of little theoretical interest.

4.3. Conjectures and Questions on more general models

What do our results on a $2 \times 2$ coordination game tell us about comparisons between dumb/cheap (like our AD model) and smart/costly (like our BR model) learning rules in general? In this subsection, we provide some conjectures and raise questions on this issue. We have analysed a specific strategic environment defined by a pure coordination game. In that game, the Nash equilibrium $(H, H)$ is both payoff and risk dominant. It would be interesting to also study other related games of coordination, like the Stag Hunt, where the payoff dominant equilibrium is risk dominated so that long run outcomes give clearer perspectives between the two opposing forces of risk and return. As mentioned in Section 1, in a stochastic environment, Juang (2002) shows that competition between a myopic BR rule and an imitation rule in a $2 \times 2$ coordination game with two Nash equilibria leads to play of the payoff dominant outcome and there is no co-existence of the two rules. Our results show, there are long run outcomes where play converges to the equilibrium which is both payoff and risk dominated. From this one can conclude that if the payoff dominated equilibrium was risk dominant, such a convergence may still be guaranteed, though in such a case, either the BR rule would be extinct (see Figure 7) or the AD rule would go extinct (see Figure 4). We conjecture that on the other hand, for a sufficiently small degree of risk dominance of the payoff dominated equilibrium, our results of convergence to the outcome $(H, H)$ (which is payoff dominant) and where the BR goes extinct should remain valid. Although a rigorous analysis of these issues is beyond the scope of the current paper, this remains an interesting avenue for future research.

Next consider coordination games where each player has $n > 2$ strategies, all off-diagonal payoffs are 0 and the Nash equilibria on the diagonal are successively strict Pareto ranked. We conjecture that in such games, if for example the social aspiration is very high and if the BR population starts playing the Pareto dominant equilibrium action, convergence to that outcome is likely. Along such paths of play, the popularity of the BR rule grows initially until most existing AD players play the action constituting the efficient equilibrium from whereon, the BR rule starts getting discarded and eventually eliminated. In other cases where the BR population starts coordinating on equilibria that are not Pareto efficient, it is expected to survive in the long run. In summary, it turns out that unless the BR players coordinate on the Pareto superior equilibrium at the very start of play and that the social aspiration is high, one cannot ensure convergence to full Pareto efficiency.

For arbitrary finite games with multiple Nash equilibria, it seems similar results should continue to hold so that the broader message from our study is this: if the BR players start coordinating on the Pareto efficient equilibrium and if the social aspiration is high enough, the BR population is likely to die out in the long run; for other cases, it is expected to survive either alone or along with the AD rule, and in such situations, convergence to full efficiency is not achievable. The general message therefore for arbitrary strategic environments with multiple Nash equilibria is as follows. One would expect to observe two of the most central phenomena we have shown: (a) convergence to the point where the entire population uses the dumb/cheap rule and successfully coordinates on the Pareto dominant equilibrium provided players using the smart/costly rule initially start using
this Pareto dominant action and (b) the co-existence of dumb/cheap and smart/costly rules where full population wide coordination is unachievable even if the cost of such smart rules is small provided players using the smart/costly rule initially starts using actions which are not Pareto dominant.

Next in order are strictly competitive games, like for example the Matching Pennies or other games with strictly dominant strategies for both players leading to inefficient Nash equilibria, like for example the Prisoners' Dilemma. Strictly competitive games are strategically very different and further separate analysis is required to understand the co-evolution of smart/costly and dumb/cheap rules and such analysis may throw out interesting convergence results (in terms of whether play converges around mixed strategy equilibria) and the possibility of the existence of long run outcomes where both models can co-exist. If it comes to the Prisoners' Dilemma, it would be hard to justify the fact that a BR type rule needs any associated costs since in such games, a BR type player would not care to know about what is the distribution of possible play from his current opponent. One may suspect that the BR rule would gain more popularity over time. However each time such players are paired with each other, they would do strictly worse than a matched pair of AD players if they still exist in sufficiently large proportions. This may imply that there exists a balanced proportion of the two types of models in the population where AD players typically cooperate while BR players keep playing the strictly dominant strategy of defection. But to know whether such mixed population environments can be asymptotically stable, one would have to undertake a full analysis that is beyond the scope of this paper. An application of our environment to a cobweb model a la BH also remains to be done. It would be interesting to know under what conditions our environment can generate similar complex dynamics as shown in BH and other related works.

4.4. Some variations on the two learning rules

We have addressed the co-evolution of a one-period memory BR rule and the reinforcement based AD rule. While the choice of homogeneous BR players simplifies the analysis, it may add realism if two types of BR rules are considered instead: e.g. one period beliefs as in the current model versus longer-memory beliefs of the fictitious-play type. One would conjecture that different BR players may coordinate on different strategies \( H \) and \( L \) in a mirror image of the heterogeneity of AD play. Moreover the fraction of BR players using \( H \) could then be endogenised. However, this would most certainly complicate the analysis as we would then need to study a higher dimensional dynamic system, and it certainly is an interesting and important problem for future investigation.

We would like to suggest another extension that leads to a three dimensional system that we feel is important to address. This is an environment where the common aspiration level of the players is allowed to evolve as a function of the average payoff of the concerned sub-population. In particular, is it still possible to have stable points or orbits where AD rule does not go extinct? Also, while we do not intend to extend the conclusions drawn from the analysis in this paper to broader classes of belief-based and reinforcement-based models, it seems reasonable to speculate that an analysis of this nature will also be successful in identifying conditions for long run popularity of different models when applied to stage games with similar strategic properties.

One may ask if the qualitative results we obtained remain robust to adding a little bit of heterogeneity in the level of aspiration or the cost of the BR rule where a more
costly BR rule uses longer history to form more accurate beliefs? While a full analysis
is beyond the scope of this paper, since our results are structurally stable (that is, they
hold for open sets of the values of the parameters such as $\alpha$ and $\varrho$), we conjecture that
our results will remain qualitatively intact as long as these heterogeneities are not too
large.

Acknowledgements

We are grateful to three anonymous referees of this journal and to Cars H. Hommes
for many questions and detailed suggestions that have greatly improved the paper in
content and exposition. We also thank Sandro Brusco, Kalyan Chatterjee, Huw Dixon,
Nick Feltovich, Martin Jensen, Ehud Kalai, Rajeeva Karandikar, Larry Samuelson, Rajiv
Sarin and Jörgen Weibull for various comments and suggestions during several stages of
this research. Roy thanks the London School of Economics for hospitality during the
final stages of this work.

Appendix A. Proofs

Linearisation at the restpoints of the dynamic systems studied in the proofs below
results in the Jacobian with one zero eigenvalue. For this reason we cannot use the
standard techniques to reason about stability of these restpoints and we proceed by
elementary methods. In particular, when showing asymptotic stability we construct
regions which are positively invariant (i.e. any trajectory that starts within them stays
inside them as $t \to +\infty$, see for example Hirsch and Smale (1974)).

Proof of Proposition 1. We start with a lemma characterizing the restpoints of the
dynamic system defined by Equations (5) and (6).

Lemma 1. The only restpoint $(a, \mu) \in \bar{Z}$ of the dynamic system defined by Equations (5)
and (6) is the point $(1, 1)$, where the entire population consists of AD players only and
all of them play H. Moreover, this restpoint is asymptotically stable.

Proof. It is easily seen that for any $a \in (0, 1]$, $\mu \in [0, 1)$ implies $M^H(a, \mu) > 0$ and if
$\mu = 1$, then $M^H(a, \mu) = 0$. On the other hand $\psi^H(a, 1) = \varrho > 0$, so $A^H(a, 1) > 0$ for all
$a \in (0, 1)$ and $A^H(1, 1) = 0$. Thus $(1, 1)$ is the only rest point in $\bar{Z}$. What remains to be
shown is its asymptotic stability.

Let $v$ be the vector field defined by (5) and (6). Consider any region $R_\theta = [1 - \theta, 1] \times
[1 - \theta, 1]$ with $\theta \in (0, \varrho/\sigma)$. We can see that on its boundary $A^H(a, \mu) \geq 0$, $M^H(a, \mu) \geq 0$
and at least one of them is positive there (in other words the vector field $v$ points inwards
on the boundary of $R_\theta$). This is because, as we observed already, $M^H(a, \mu) > 0$ for all
$(a, \mu) \in (0, 1) \times (0, 1)$ and $M^H(a, 1) = 0$ for all positive $a \leq 1$. Also, $A^H(a, \mu) > 0$ for all
$\mu \in (1 - \varrho/\sigma, 1)$, as in this case

$$\psi^H(a, \mu) = a(1 - \mu)^2(\delta + \sigma) - (1 - \mu)\sigma + \varrho > a(1 - \mu)^2(\delta + \sigma) > 0.$$ 

Thus $R_\theta$ is positively invariant. Since this is true for any $\theta \in (0, \varrho/\sigma)$ so $(1, 1)$ is
asymptotically stable. \qed
What remains to be shown to complete the proof of Proposition 1 is that all trajectories starting from within the region $Z_{>\delta}$ end up in the restpoint $(1,1)$ and never enter the area $A_L$. Let $v$ be the vector field defined by Equations (5) and (8). Notice that $v$ never points outwards on the boundary of $Z_{>\delta}$ and it points inwards on the boundary where $\mu = \delta$ and $a \in (0,1)$. Notice also that $Z_{>\delta}$ and $A_L$ are disjoint. This implies that any trajectory starting from $Z_{>\delta}$ will stay in $Z_{>\delta}$ and, in particular, will never enter $A_L$.

Furthermore, by the Poincaré-Bendixson Theorem (see for example Hirsch and Smale (1974) or Hubbard and West (1991)) limit sets of solutions of two dimensional differential equations either include a rest point or are closed orbits. Since $M^H(a,\mu) > 0$ for all $a \in (0,1)$ and $\mu > 0$, as well as $A^H(a,1) > 0$ for $a \in (0,1)$, so no close orbit within $Z_{>\delta}$ is possible and so all the trajectories must reach the asymptotically stable restpoint $(1,1)$. Hence $\lim_{t \to +\infty} \phi_v(t,x) = (1,1)$. This completes the proof of Proposition 1.

**Proof of Proposition 2**

Like in the proof of Proposition 1 we start with a lemma characterizing the restpoints of the dynamic system defined by Equations (5) and (8).

**Lemma 2.** If $\varrho \in (0,\sigma^2/4)$, then the set of restpoints in $Z \setminus A_L$ of the dynamic system defined by Equations (5) and (8) is $\{(1,\tilde{\mu}) : \tilde{\mu} \in (\delta,\mu_1] \cup [\mu_2,1]\}$ where

$$
\mu_1 = \delta + \frac{\sigma - \sqrt{\sigma^2 - 4\varrho}}{2}, \quad \mu_2 = \delta + \frac{\sigma + \sqrt{\sigma^2 - 4\varrho}}{2}.
$$

If $\varrho \in [\sigma^2/4,\delta)$, then the set of restpoints in $Z \setminus A_L$ of the dynamic system defined by Equations (5) and (8) is $\{(1,\bar{\mu}) : \bar{\mu} \in (\delta,1]\}$.

**Proof.** It is easily seen that $M^H(a,\mu) = 0$ for $(a,\mu) \in Z$ iff $a = 1$ or $\mu = 1$. Since $A^H(a,1) > 0$ for $a \in (0,1)$, as $\psi^H(a,1) = \varrho > 0$ for $a \in (0,1)$, so any restpoint $(a,\mu) \in Z$ of the system defined by Equations (5) and (8) must satisfy $a = 1$. Since $A^H(1,\mu) = 0$ holds if $\psi^H(1,\mu) \leq 0$, so the set of restpoints of the system defined by Equations (5) and (8) lying in $Z \setminus A_L$ is the set of all points $(1,\bar{\mu})$ such that $\bar{\mu} \in (\delta,1]$ and satisfies the inequality

$$
(1 - \bar{\mu})^2 - (1 - \bar{\mu})\sigma + \varrho \leq 0. \quad (A.1)
$$

Solving it we find that either $\bar{\mu} \in (\delta,\mu_1] \cup [\mu_2,1]$ with

$$
\mu_1 = \delta + \frac{\sigma - \sqrt{\sigma^2 - 4\varrho}}{2}, \quad \mu_2 = \delta + \frac{\sigma + \sqrt{\sigma^2 - 4\varrho}}{2},
$$

if $\varrho \in [\sigma^2/4,\delta)$, or $\bar{\mu} \in (\delta,1]$, if $\varrho \in [\sigma^2/4,\delta)$.

Let $R$ be the set of restpoints of the system defined by Equations (5) and (8) lying in $Z \setminus A_L$, as identified in Lemma 2. What remains to be shown to complete the proof of Proposition 2 is that all trajectories starting from within the region $Z_{>\delta}$ end up in a restpoint $(1,\bar{\mu}) \in R$ and never enter the area $A_L$. Let $v$ be the vector field defined by Equations (5) and (8). Like in proof of Proposition 1, $v$ never points outwards on the boundary of $Z_{>\delta}$ and it points inwards on the boundary where $\mu = \delta$ and $a \in (0,1)$. Thus, by the same arguments as used there, any trajectory starting from $Z_{>\delta}$ will stay in $Z_{>\delta}$ and will never enter $A_L$. Furthermore, since $M^H(a,\mu) > 0$ for all $a \in (0,1)$ and $\mu > 0$, as well as $A^H(a,1) > 0$ for $a \in (0,1)$, so no close orbit within $Z_{>\delta}$ is possible and so, by the Poincaré-Bendixson Theorem, all the trajectories must reach a restpoint in the set $R$. Hence $\lim_{t \to +\infty} \phi_v(t,x) \in R$. This completes the proof of Proposition 2. \qed

21
**Proof of Proposition 3** We start with a lemma identifying the restpoints of the dynamic system defined by Equations (11) and (10) and conditions of their existence. This is followed by a corollary that identifies and characterizes conditions of existence of the restpoints of the dynamic system defined by Equations (9) and (10). Throughout the proof we will use $\nu$ to refer to the vector field defined by Equations (11) and (10).

**Lemma 3.** Let $0 < \delta < \sigma$, $\delta + \sigma = 1$, $\varrho \in (0, \delta)$ and $\alpha \in (\delta, \sigma]$. Then system of equations
\[
\begin{cases}
\psi^{\text{L}}(a, \mu) = 0, \\
M^{\text{L}}(a, \mu) = 0.
\end{cases}
\tag{A.2}
\]
either has two, one or zero solutions in $\hat{Z} \setminus A_{\text{H}}$, depending on the value of $\varrho$. Let
$$
\hat{\varrho} = \frac{\delta}{1 + h(\alpha, \delta)},
\tilde{\varrho} = \left(h(\alpha, \delta) - \sqrt{(h(\alpha, \delta) - 1)(h(\alpha, \delta) + 1 - \delta)}\right)^2,
$$
where $h(\alpha, \delta) = \frac{f(\alpha)}{f(\alpha - \delta)}$. The system of Equations (A.2) has two solutions in $\hat{Z} \setminus A_{\text{H}}$ if and only if $\varrho \in (\hat{\varrho}, \tilde{\varrho})$. The system has one solution in $\hat{Z} \setminus A_{\text{H}}$ if and only if $\varrho \in (0, \hat{\varrho}] \cup \{\tilde{\varrho}\}$. The system has no solutions if and only if $\varrho > \tilde{\varrho}$.

**Proof.** Throughout the proof we will write $h$ instead of $h(\alpha, \delta)$, for short. Notice that if $\alpha \in (\delta, \sigma]$, then $h > 1$ and the system of Equations \((A.2)\) is equivalent to the following:
\[
\begin{cases}
\mu a^2 - \mu \delta + \varrho = 0, \\
(1 - \mu)(1 - a\mu) - h\mu(1 - a) = 0.
\end{cases}
\tag{A.3}
\]
The line $\mu = 0$ is an asymptote of both curves defined by the above equations, and so the following analysis will be conducted for $\mu \neq 0$. By solving these two equations for $a$ we get
\[
a = \frac{\delta \mu - \varrho}{\mu}, \quad \text{or} \quad a = \frac{\mu(h + 1) - 1}{\mu(h + 1 - h)}.
\tag{A.4}
\]
Subtracting the second from the first equation in \((A.4)\) and multiplying both sides by $\mu(\mu + h - 1)$ (for $\mu \neq 0$) we get the following quadratic equation:
\[
\mu^2(h + 1 - \delta) - \mu(\delta(h - 1) + 1 - \varrho) + \varrho(h - 1) = 0.
\tag{A.5}
\]
Notice that the function $a^{\text{L}}(\mu) = \frac{\mu(h + 1) - 1}{\mu(\mu + h - 1)}$, implicitly defined by the equation $M^{\text{L}}(a, \mu) = 0$, is strictly increasing in $\mu$ on the interval $(0, 1)$ (which can be easily seen by looking at $a^{\text{L}}(\mu)$). Thus solutions of Equation \((A.4)\) uniquely identify the solutions of \((A.2)\).

Equation \((A.3)\) has zero, one or two solutions in real values. Its discriminant is $\Delta = (\delta(h - 1) + 1 - \varrho)^2 - 4\varrho(h - 1)(h + 1 - \delta)$. Solving the inequality $\Delta \geq 0$ for $\varrho$ we get $\varrho \leq \varrho_1$ or $\varrho \geq \varrho_2$, where
\[
\varrho_1 = \left(h - \sqrt{(h - 1)(h + 1 - \delta)}\right)^2, \quad \varrho_2 = \left(h + \sqrt{(h - 1)(h + 1 - \delta)}\right)^2.
\]
Notice that $\varrho_1 \geq 0$ and, since $h > 1$, so $\varrho_2 > 1$. Moreover, since $(h - 1)(h + 1 - \delta) = h^2 - 1 - (h - 1)\delta$, so $\varrho_1 > 0$ for $\delta \in (0, 1)$ and $h > 1$. Since we restrict the attention to $\varrho \in (0, \delta)$, so \((A.2)\) has two solutions if $\varrho \in (0, \varrho_1)$, one solution if $\varrho = \varrho_1$ and no solutions if $\varrho \in (\varrho_1, \delta)$. 

22
Observation 1. Notice that if we restrict our attention to $\varrho \in (0, \delta)$, then the case of no solutions is possible only if $\varrho_1 < \delta$. This is possible if and only if $\alpha$, $h$ and $\delta$ satisfying our assumptions are such that $h > 1/(2\sqrt{\delta} - \delta)$. To see this, inserting the formula for $\varrho_1$ to this inequality we get

$$\left(h - \sqrt{(h - 1)(h + 1 - \delta)}\right)^2 < \delta.$$ \hfill (A.6)

Expanding the left hand side, subtracting the right hand side from both sides and adding $2h\sqrt{(h - 1)(h + 1 - \delta)}$ to both sides we get:

$$2h^2 - h\delta - 1 < 2h\sqrt{(h - 1)(h + 1 - \delta)}.$$ \hfill (A.7)

Both sides of this inequality are positive for $h > 1$ and $\delta \in (0, 1)$. Squaring them and subtracting right hand side from both sides we get:

$$(\delta h + 1)^2 - 4h^2\delta < 0.$$ \hfill (A.8)

For $h > 1$ and $\delta \in (0, 1)$ this is equivalent to

$$h > \frac{1}{2\sqrt{\delta} - \delta}.$$ \hfill (A.9)

The expressions for solutions of Equation (A.5) are:

$$\mu_1 = \frac{\delta(h - 1) + 1 - \varrho - \sqrt{\Delta}}{2(h + 1 - \delta)}, \quad \mu_2 = \frac{\delta(h - 1) + 1 - \varrho + \sqrt{\Delta}}{2(h + 1 - \delta)}.$$ \hfill (A.10)

The function $a^M(\mu)$ intersects the line $a = 0$ at $\mu = 1/(h + 1)$ and it intersects $s^L(a, \mu)$ at the unique point within $\bar{Z}$ with $\bar{\mu} = (\delta(h - 1) + 1)/(h + 1 - \delta)$ and whenever $\mu \in [0, \bar{\mu})$, then $M^L(a, \mu) = 0$ lies outside $A_H$. Since $\Delta < (\delta(h - 1) + 1 - \varrho)^2$, so $0 < \mu_1 < \mu_2$. Moreover,

$$\mu_2 < \frac{\delta(h - 1) + 1 - \varrho}{h + 1 - \delta} < \frac{\delta(h - 1) + 1}{h + 1 - \delta} = \bar{\mu}.$$ \hfill (A.11)

Thus both $\mu_1$ and $\mu_2$ lie outside $A_H$.

Observation 2. The analysis above implies that

$$\frac{\delta \mu - \varrho}{\mu^2} > \frac{\mu(h + 1) - 1}{\mu(\mu + h - 1)}, \text{ for } \mu \in (0, \mu_1) \cup (\mu_2, 1) \text{ and}$$ \hfill (A.12)

and

$$\frac{\delta \mu - \varrho}{\mu^2} < \frac{\mu(h + 1) - 1}{\mu(\mu + h - 1)}, \text{ for } \mu \in (\mu_1, \mu_2).$$ \hfill (A.13)

Now we will find conditions under which $\mu_1 > 1/(h + 1)$. We will show first that $\mu_1$ is strictly increasing with $\varrho$ increasing on the interval $[0, \varrho_1]$. The part of $\mu_1$ which depends on $\varrho$ is: $-\varrho - \sqrt{\Delta}$. Let

$$\tau = \delta(h - 1) + 1 + 2(h - 1)(h + 1 - \delta) \text{ and } v = \delta(h - 1) + 1,$$

then

$$-\varrho - \sqrt{\Delta} = -\varrho - \sqrt{\varrho^2 - 2\varrho \tau + \tau^2}.$$ \hfill (A.14)
Differentiating the right hand side with respect to \( \varrho \) we get 
\[
\frac{(\tau - v) - \sqrt{\varrho^2 - 2\varrho v + v^2}}{\sqrt{\varrho^2 - 2\varrho v + v^2}}
\] which is positive for \( h > 1 \) and \( \delta \in (0, 1) \), as \( \tau > v \) in this case and, consequently, \((\tau - v)^2 > \varrho^2 - 2\varrho v + v^2\). This shows that \( \mu_1 \) is strictly increasing with \( \varrho \) increasing on the interval \([0, g_1]\). Since \( \mu_1 = 1/(h + 1) \) for \( \varrho = \delta/(h + 1) \), so if \( \varrho \in [\delta/(h + 1), g_1] \), then \( (A.2) \) has two solutions in \( \mathcal{Z} \setminus \mathcal{A}_H \).

To show that if \( \varrho = g_1 \), then the solution of \( (A.2) \) is in \( \mathcal{Z} \setminus \mathcal{A}_H \), we will show the following fact.

**Fact 1.** For all \( h > 1 \) and \( \delta \in (0, 1) \), \( g_1 \geq \delta/(h + 1) \) and the equality holds only if \( h = (\sqrt{4\delta + 1} + 1)/(2\delta) \).

**Proof.** Inserting the formula for \( g_1 \) to the inequality we get
\[
\left( h - \sqrt{(h - 1)(h + 1 - \delta)\right)^2 \geq \frac{\delta}{h + 1}.
\] (A.14)

Expanding the right hand side, subtracting the left hand side from both sides, adding \( 2h/(h - 1)(h + 1 - \delta) \) to both sides and, finally, multiplying both sides by \( h + 1 \) we get:
\[
h^2(h - 1 - \delta) + (h + 1)^2(h - 1) \geq 2h(h + 1)\sqrt{(h - 1)(h + 1 - \delta)}.
\] (A.15)

Both sides of this inequality are positive for \( h > 1 \) and \( \delta \in (0, 1) \). Squaring them and subtracting right hand side from both sides we get:
\[
(\delta h^2 - (h + 1))^2 \geq 0.
\] (A.16)

This shows that the inequality \( g_1 \geq \delta/(h + 1) \) is satisfied for all \( h > 1 \) and \( \delta \in (0, 1) \). Moreover it can be easily checked that in the case of \( h > 1 \) equality holds only if \( h = (\sqrt{4\delta + 1} + 1)/(2\delta) \), \( \square \).

**Observation 3.** Fact 2 implies that the case of \( \alpha, h \) and \( \delta \) satisfying our assumptions, we have \( g_1 > \delta/(h + 1) \) if and only if \( h \neq (\sqrt{4\delta + 1} + 1)/(2\delta) \).

If \( \varrho = g_1 \), then, by Fact 1, \( g \geq \delta/(h + 1) \). Since in this case \( \mu_1 = \mu_2 \) and since, by what was shown above, \((a^M(\mu_1), \mu_1) \in \mathcal{Z} \setminus \mathcal{A}_H \) for \( g \geq \delta/(h + 1) \), so the solution of \( (A.2) \) is in \( \mathcal{Z} \setminus \mathcal{A}_H \) in the case of \( g = g_1 \).

Lastly, if \( \varrho \in (0, \delta/(h + 1)) \), then Equation \( (A.5) \) has two solutions, but one of them, \( \mu_1 < 1/(h + 1) \) and so \((a^M(\mu_1), \mu_1) \notin \mathcal{Z} \setminus \mathcal{A}_H \). We will show that in this case \( \mu_2 \geq 1/(h + 1) \), and so the second solution of \( (A.2) \) is in \( \mathcal{Z} \setminus \mathcal{A}_H \). We will show first that \( \mu_2 \) is strictly decreasing with \( \varrho \) increasing on the interval \([0, g_1]\). Similarly to the case of \( \mu_1 \), we will look at the part of \( \mu_2 \) which depends on \( g \). Differentiating it with respect to \( g \) we get 
\[
\frac{\sqrt{\varrho^2 - 2\varrho v + v^2 - (\tau - v)^2}}{\sqrt{\varrho^2 - 2\varrho v + v^2}}
\] which is negative for \( h > 1 \) and \( \delta \in (0, 1) \), as \( \tau > v \) in this case (c.f. analogous arguments for \( \mu_1 \) above). Thus \( \mu_2 \) is strictly decreasing with \( \varrho \) increasing on the interval \([0, g_1]\). Since \( \mu_2 = 1/(h + 1) \) for \( \varrho = \delta/(h + 1) \), so if \( \varrho \in (0, \delta/(h + 1)) \), then \( \mu_2 > 1/(h + 1) \) and \( (A.2) \) has one solution in \( \mathcal{Z} \setminus \mathcal{A}_H \), \( \square \).
Then the following cases are possible, for different values of $\varrho$:

(i). If $\varrho \leq \hat{\varrho}$ and $\varrho < \check{\varrho}$, then \( A(a, \mu) = 0 \) has two distinct solutions in $\bar{Z} \setminus \mathcal{A}_H$, $x_a = (a_u, \mu_u)$ and $x_u = (a_u, \mu_0)$, with $1 > a_u > a_s = 0$ and $1 > \mu_u > \mu_a$.

(ii). If $\varrho = \hat{\varrho} = \check{\varrho}$, then \( A(a, \mu) = 0 \) has one solution in $\bar{Z} \setminus \mathcal{A}_H$, $x_a = (a_u, \mu_u)$, with $a_u = 0$.

This case is possible iff $h = \frac{\sqrt{3} + 1}{25}$. \hfill \Box

(iii). If $\varrho > \hat{\varrho}$ and $\varrho < \check{\varrho}$, then \( A(a, \mu) = 0 \) has two distinct solutions in $\bar{Z} \setminus \mathcal{A}_H$, $x_a = (a_u, \mu_u)$ and $x_u = (a_u, \mu_0)$, with $1 > a_u > a_s > 0$ and $1 > \mu_u > \mu_a > 0$. This case is possible iff $h \neq \frac{\sqrt{3} + 1}{25}$.

(iv). If $\varrho > \hat{\varrho}$ and $\varrho = \check{\varrho}$, then \( A(a, \mu) = 0 \) has one solution in $\bar{Z} \setminus \mathcal{A}_H$, $x_a = (a_u, \mu_u)$, with $a_u > 0$.

(v). If $\varrho > \check{\varrho}$, then \( A(a, \mu) = 0 \) has no solutions in $\bar{Z} \setminus \mathcal{A}_H$. This case is possible iff $h > \frac{1}{2\sqrt{3} - 3}$.

Proof. It is easy to check that if $A(a, \mu) = 0$ then either $\varphi^L(a, \mu) = 0$ or $a = 0$ or $a = 1$. Since $M^L(1, \mu) = 1$ for $\mu \in [0, 1]$ only if $\mu = 1$ and $(1, 1) \in \mathcal{A}_H$, so the only cases we need to consider are the first two. Notice that $A^L(0, \mu) = 0$ only if $\varphi^L(0, \mu) \leq 0$, that is $\mu \in [\varrho/\delta, 1]$. On the other hand $M^L(0, \mu) = 0$ only if $\mu = 1/(h + 1)$. Also, as we have shown in proof of Lemma 3 the function $a^L(\mu)$ defined implicitly by the equation $M^L(a, \mu) = 0$ is strictly increasing on the interval $(0, 1]$. Now it is easy to see that the cases listed above are immediate consequence of Lemma 3 (in particular of Observations 3 and 4).

Having identified possible restpoints of the dynamic system defined by Equations (11) and (10) and conditions of their existence, we will proceed with their characterization. Firstly, we will show that the restpoint $x_a$ is unstable in all the cases listed in Corollary 1.

Lemma 4. The restpoint $x_a$ is unstable.

Proof. Consider the functions

\[
\begin{align*}
a^L(\mu) &= \frac{\delta \mu - \varrho}{\mu}, \\
a^M(\mu) &= \frac{\mu(h + 1) - 1}{\mu(\mu + h - 1)}
\end{align*}
\]

(A.18)

defined implicitly by the Equations (A.2). As was shown in Lemma 3 these functions intersect in at most two points (for the considered values of the parameters $\alpha$, $\delta$ and $\varrho$) and $x_a$ is one of these solutions with larger $\mu$. Notice that, by Observation 2 $a^M(\mu) > a^L(\mu)$ for $\mu \in (\mu_a, 1)$. Thus $\hat{a} > 0$ for $\mu \in (\mu_a, 1)$ and $\bar{a} > \frac{\sqrt{\mu^2 - \varrho^2}}{\mu^2}$. Let

\[ R = \{(a, \mu) \in \bar{Z} \setminus \mathcal{A}_H : \mu > \mu_a \text{ and } a \geq a^M(\mu)\} \]

be the region bounded by the line $\mu = \mu_a$, the isocline $M^L(a, \mu) = 0$ and the boundaries of the region $\mathcal{A}_H$. Let $\mathbf{v}$ be the vector field defined by (9) and (10). Since $a^M(\mu)$ is strictly increasing with $\mu$ increasing on the interval $(0, 1]$, $\mu = 0$ and $\bar{a} > 0$ for all points $(a^M(\mu), \mu)$ with $\mu \in (\mu_a, 1)$, so the vector field $\mathbf{v}$ points inwards on this boundary of $R$. \hfill \Box
for $\mu \in (\mu_a, 1)$. Similarly with the boundary $(a, \mu_a)$ for $a > a_u$, as both $\dot{\mu} > 0$ and $\dot{a} > 0$ there. Also $\dot{\mu} > 0$ within $\mathcal{R}$.

Now, consider any neighbourhood $\mathcal{N}$ of $x_u$. It must be that $\mathcal{X} = \mathcal{N} \cap \mathcal{R} \neq \emptyset$. Since $\dot{\mu} > 0$ within $\mathcal{X}$, so there can be no restpoints nor stable orbits there and so, by the Poincaré-Bendixson Theorem, any trajectory starting within $\mathcal{X}$ must eventually leave it. Hence $x_u$ is unstable. \qed

Secondly we will show that the restpoint $x_a$ is asymptotically stable.

**Lemma 5.** The restpoint $x_a$ is asymptotically stable.

**Proof.** We divide the proof into three cases: (a) $\varrho < \dot{\varrho}$, (b) $\varrho > \dot{\varrho}$ and (c) $\varrho = \dot{\varrho}$. In all the cases we will construct a positive invariant region $\mathcal{R}_\varrho$ (with some sufficiently small $\theta > 0$) containing $x_a$. The construction depends on the value of $\varrho$ and is different in each case. Case (a) is relatively straightforward, case (b) is the most involved and case (c) relies on the construction from case (b).

**Case (a):** $\varrho < \dot{\varrho}$. Consider a region $\mathcal{R}_\varrho = [0, \theta] \times [1/(h+1) - \theta, 1/(h+1) + \theta]$ defined for $\theta > 0$. Notice that for any $\theta > 0$, $x_a \in \mathcal{R}_\varrho$. Also, there exists $\varrho > 0$ such that for any $\theta \in (0, \varrho)$ and any $(a, \mu) \in \mathcal{R}_\varrho$ it holds that $\psi^\varrho(a, \mu) < 0$ and, consequently, $\dot{\varrho} \leq 0$ with equality for $a = 0$ only. This is because $a^\varrho(\mu)$ (c.f. Equation (A.18)) is continuous, strictly increasing for $\mu \in [0, 2\varrho/\delta)$ and then strictly decreasing for $\mu \in (2\varrho/\delta, 1]$ (which can be easily seen by analysing $a^\varrho(\mu)$) until it reaches $\delta - \varrho$ at $\mu = 1$. Thus we could take $\varrho = \min(\varrho, a^\varrho(1/(h+1) - \tau))$ with sufficiently small $\tau$ where $\tau \in (0, 1/(h+1) - \varrho/\delta)$. Moreover, since $a^M(1/(h+1)) = (h+1)^3/h^2 > 1$ (where $a^M(\mu)$ is defined by Equation (A.18)) and is continuous in the neighbourhood of $1/(h+1)$, so there exists $\varrho \in (0, \theta)$ such that $a^M(\mu) > 1$ for all $\mu \in (0, \theta)$. Thus $\mu > 0$ for all $(a, 1/(h+1) - \theta)$ and $\mu < 0$ for all $(a, 1/(h+1) + \theta)$, with $a \in (0, \theta)$ and $\theta < \theta$. This shows that $\varrho$ points inwards on the boundary of $\mathcal{R}_\varrho$ for all $\theta \in (0, \theta)$ and so $\mathcal{R}_\varrho$ is positively invariant for any such $\theta$. Hence $x_a$ is asymptotically stable.

**Case (b):** $\varrho > \dot{\varrho}$. Let

$$K_1 = \frac{M^L_1(x_a) \cdot M^L_1(x_a)}{M^L_1(x_a) - M^L_1(x_a)}, \quad K_2 = \frac{M^L_2(x_a) \cdot M^L_2(x_a)}{M^L_2(x_a) - M^L_2(x_a)},$$

and let $K = 1 + \max(K_1, K_2)$. Notice that $K_1$ and $K_2$ are well defined for $(a, \mu) \in \hat{Z}$, as for such values of $(a, \mu)$, $M^L(a, \mu) \geq 0$, $M^L_1(a, \mu) \leq 0$ and they are never equal. This implies also that $K_1, K_2, K > 0$ for $(a, \mu) \in \hat{Z}$. Let

$$\eta(x) = \max_{(a, \mu) \in C(Kx)} g(|\psi^L(a, \mu)|),$$

and $C(x) = [a_a - x, a_a + x] \times [\mu_a - x, \mu_a + x]$.

Consider the following curves, defined for $\theta > 0$

$$\varphi^L_\theta(a, \mu) = M^L(a, \mu)^2 - 2\kappa_1\kappa_2 f(\alpha)\eta(\theta)(a - \mu_a + \mu - \mu_a + \theta) = 0,$$

$$\varphi^L_\theta(a, \mu) = M^L(a, \mu)^2 + 2\kappa_1\kappa_2 f(\alpha)\eta(\theta)(a - \mu_a + \mu - \mu_a - \theta) = 0.$$

The following two facts are important for construction of the region $\mathcal{R}_\varrho$. Their proofs are given at the end of the Appendix.
Fact 2. There exists $\bar{\theta} > 0$ such that for all $\theta \in (0, \bar{\theta})$ there is an intersection point $x_1^0 = (a_1^0, \mu_1^0)$ of the curves $\varphi_1^0(a, \mu) = 0$ and $M^1(a, \mu) = 0$ and an intersection point $x_2^0 = (a_2^0, \mu_2^0)$ of the curves $\varphi_2^0(a, \mu) = 0$ and $M^1(a, \mu) = 0$. Moreover, these intersection points have the following properties:

(i). $\psi^1(x_1^0) > 0$ and $-\kappa < a_1^0 - a_1 < 0$, $-\kappa < \mu_1^0 - \mu_1 < 0$, and $\psi^1(x_2^0) < 0$ and $0 < a_2^0 - a_2 < \kappa$, $0 < \mu_2^0 - \mu_2 < \kappa$.

Fact 3. There exists $\bar{\theta} > 0$ satisfying Fact 2 and such that for all $\theta \in (0, \bar{\theta})$

(i). there is an intersection point $x_2^0 = (a_2^0, \mu_2^0)$ of the curves $\varphi_2^0(a, \mu) = 0$ and $\psi^1(a, \mu) = 0$ s.t. $M^1(x_2^0) > 0$, $a_2 < a_2^0 < a_2^0$ and $\mu_2 < \mu_2^0 < \mu_2^0$,

(ii). if $M^1(a, \mu) > 0$, $\psi^1(a, \mu) > 0$ and $\varphi_2^0(a, \mu) = 0$, then $a < a_2 + \kappa$ and $\mu < \mu_2 - \kappa$,

(iii). there is an intersection point $x_3^0 = (a_3^0, \mu_3^0)$ of the curves $\varphi_3^0(a, \mu) = 0$ and $\psi^1(a, \mu) = 0$ s.t. $M^1(x_3^0) < 0$, $a_3 < a_3^0 < a_3$ and $\mu_3^0 < \mu_3 < \mu_3$,

(iv). if $M^1(a, \mu) < 0$, $\psi^1(a, \mu) < 0$ and $\varphi_2^0(a, \mu) = 0$, then $a > a_2 - \kappa$ and $\mu < \mu_2 + \kappa$.

Let

$$\ell^1_1(a, \mu) = (\mu_1 - \mu_2^0)(a - a_2^0) - (a_3 - a_2^0)(\mu - \mu_2^0) = 0,$$

$$\ell^2_2(a, \mu) = (\mu^0_2 - \mu_2^0)(a - a_2^0) - (a_2^0 - a_2^0)(\mu - \mu_2^0) = 0,$$

be the lines going through points $x_2^0$, $x_3^0$, and $x_4^0$, respectively. Let $R_0 = R_0^1 \cup R_0^2 \cup R_0^3 \cup R_0^4$, where

$$R_0^1 = \{(a, \mu) : \varphi_1^0(a, \mu) \leq 0 \text{ and } A^1(a, \mu) \leq 0 \text{ and } M^1(a, \mu) \leq 0\},$$

$$R_0^2 = \{(a, \mu) : \varphi_2^0(a, \mu) \leq 0 \text{ and } A^1(a, \mu) \geq 0 \text{ and } M^1(a, \mu) \geq 0\},$$

$$R_0^3 = \{(a, \mu) : \ell^1_1(a, \mu) \leq 0 \text{ and } A^1(a, \mu) \geq 0 \text{ and } M^1(a, \mu) \leq 0\},$$

$$R_0^4 = \{(a, \mu) : \ell^2_2(a, \mu) \leq 0 \text{ and } A^1(a, \mu) \leq 0 \text{ and } M^1(a, \mu) \geq 0\},$$

(see Figure A.8).

Facts 2 and 3 guarantee that there exists $\bar{\theta}$ such that for any $\theta \in (0, \bar{\theta})$, $x_\theta \in R_\theta$ and $R_\theta$ is positively invariant, as implied by the following lemma.

Lemma 6. Let $\theta$ be such that Facts 2 and 3 hold. Then for any $\theta \in (0, \bar{\theta})$ the vector field $v$ points inwards on the boundary of $R_\theta$.

Proof. To check that the vector field $v$ points inwards we will compare the direction of the vector field and the slope of the tangent to the boundary along all four parts of the boundary. That is we have to check whether $D(a, \mu) = F_0(a, \mu)A^1(a, \mu) + F_1(a, \mu)M^1(a, \mu) < 0$ is true along the entire boundary $F(a, \mu) = 0$ of $R_\theta$.

If $M^1(a, \mu) \geq 0$ and $A^1(a, \mu) \geq 0$, then the corresponding part of the boundary is the curve $\varphi_3^0(a, \mu) = 0$ and we have

$$D(a, \mu) = 2A^1M^1M^1_{\alpha} - \theta^2A + 2(M^1)^2M^1_{\mu} - 2M^{1}\kappa_1\kappa_2 f(\alpha)(\eta(\kappa, \theta))$$

$$= 2M^1(A^1M^1_{\alpha} - \kappa_1\kappa_2 f(\alpha)(\eta(\theta))) - \theta^2A + 2(M^1)^2M^1_{\mu} < 0,$$
\[ M_1^L(a, \mu) = \kappa_2 \mu (f(\alpha) - f(\alpha - \delta)(1 - \mu)) < \kappa_2 f(\alpha), \]
\[ M_1^R(a, \mu) = -\kappa_2 (f(\alpha - \delta)(a(1 - \mu) + 1 - a\mu) + f(\alpha)(1 - a)) < 0, \]
\[ A^L(a, \mu) = \kappa_1 (1 - a) g(\psi^L(a, \mu)) < \eta(\theta), \]

for all \( \theta \in (0, \bar{\theta}) \). Similarly it can be shown that \( D(a, \mu) < 0 \) for \( M^L(a, \mu) \leq 0 \) and \( A^L(a, \mu) \leq 0 \), when the corresponding part of the boundary is the curve \( \varphi^L_1(a, \mu) = 0 \).

If \( M^L(a, \mu) \geq 0 \) and \( A^L(a, \mu) \leq 0 \), then the corresponding part of the boundary is the line \( \ell^1_\theta(a, \mu) = 0 \) and we have
\[ D(a, \mu) = (\mu^3_3 - \mu^3_2) A^L - (a^3_3 - a^3_2) M^L < 0, \]
as, by Fact 3, \( \mu^3_3 > \mu^3_2 \) and \( a^3_3 > a^3_2 \). Similarly it can be shown that \( D(a, \mu) < 0 \) holds for \( M^L(a, \mu) \leq 0 \) and \( A^L(a, \mu) \geq 0 \), when the corresponding part of the boundary is the curve \( \ell^2_\theta(a, \mu) = 0 \).

Let \( \bar{\theta} > 0 \) be such that Facts 2 and 3 hold. Since for any \( \theta \in (0, \bar{\theta}) \), \( x_s \in R_\theta \) and \( R_\theta \) is positively invariant, so \( x_s \) is asymptotically stable.

**Case (c):** \( \rho = \rho \). Construction of the region \( R_\theta \) that would be positive invariant for all sufficiently small \( \theta > 0 \) combines the constructions used in the two previous cases and we will omit the details here, just giving the idea of this construction. In the case of \( a \) and \( \mu \) such that \( M^L(a, \mu) \geq 0 \) the region is the intersection of \( R_\theta \) used to show Case (ii) and \( \{(a, \mu) : M^L(a, \mu) \geq 0\} \), while in the case of \( a \) and \( \mu \) such that \( M^L(a, \mu) \leq 0 \) it is the intersection of \([0, \theta'] \times [\mu_s, \mu_s + \theta']\) and \( \{(a, \mu) : M^L(a, \mu) \leq 0\} \), where \( \theta' > 0 \) is such that the two regions ‘connect well’ on the line \( M^L(a, \mu) = 0 \). Showing that \( R_\theta \) is positively invariant for sufficiently small \( \theta > 0 \) combines arguments used for Cases (a) and (b) (and is easier than Case (b)).

Having characterized the restpoints of the system and conditions of their existence we are ready to prove Proposition 3. We will address each of the points (i) – (iii) separately.
Point (i) If \( \varrho \leq \hat{\varrho} \), then, by Corollary 1, the dynamic system under consideration has two distinct restpoints in \( \bar{Z} \setminus A_H \) and, by Lemmas 4 and 5 only one of them, \( x_a = (0, 1/(1 + h)) \), is stable. Notice that the vector field \( \bar{v} \) never points outwards on the boundary of the set \( P_{a} \) (containing \( B \)) and it does not point inward at \( x_a \) and on the curve \( A^1(a, \mu) = 0 \) only. Thus for any \( \bar{x} \in P_a \) and all \( t > 0 \), \( \phi_v(t, \bar{x}) \in P_{a} \) and \( \lim_{t \to +\infty} \phi_v(t, \bar{x}) = (0, 1/(1 + h(\alpha, \delta))) \). To show that all the trajectories that stay in \( \bar{Z} \setminus A_H \) converge to \( x_a \) we need to show that there is no stable orbit within \( \bar{Z} \setminus A_H \).

The open set enclosed by a stable orbit must contain a restpoint (c.f. (Hirsch and Smale 1974, Theorem 2, Chapter 11)). Thus any such orbit would have to enclose an open set containing either \( x_a \) or \( x_u \). The first case is impossible, as \( a = 0 \). The second case is impossible as well, as any such orbit would have to enter the region \( P_{a} \) and, as we have argued above, it would converge to \( x_a \). Thus there is no stable orbit within \( \bar{Z} \setminus A_H \) and, by the Poincaré-Bendixson Theorem, for any \( \bar{x} \in \bar{Z} \) such that for all \( t \geq 0 \), \( \phi_v(t, \bar{x}) \in \bar{Z} \setminus A_H \), \( \lim_{t \to +\infty} \phi_v(t, \bar{x}) = (0, 1/(1 + h(\alpha, \delta))) \).

Point (ii) If \( \varrho \leq \hat{\varrho} \), then, by Corollary 1, the dynamic system under consideration has two distinct restpoints in \( \bar{Z} \setminus A_H \) and, by Lemmas 4 and 5 only one of them, \( x_a = (0, 1) \) and \( x_u = (0, 1) \). Since, by Lemma 5, \( x_a \) is asymptotically stable, so there exists a set \( X \subseteq \bar{Z} \setminus A_H \) such that for any \( \bar{x} \in X \) and for all \( t \geq 0 \), \( \phi_v(t, \bar{x}) \in X \) and \( \lim_{t \to +\infty} \phi_v(t, \bar{x}) = x_a \). By Corollary 1, this case is possible if \( h(\alpha, \delta) \neq \frac{\sqrt{4\delta + 1}}{2\delta - 1} \).

Point (iii) If \( \varrho \leq \hat{\varrho} \), then, by Corollary 1, the dynamic system under consideration has no restpoints in \( \bar{Z} \setminus A_H \). Thus there cannot be any closed orbit in \( \bar{Z} \setminus A_H \) neither, as the open set enclosed by any such orbit must contain a restpoint. Hence, by the Poincaré-Bendixson Theorem, for any \( \bar{x} \in \bar{Z} \) there is \( t \geq 0 \) such that \( \phi_v(t, \bar{x}) \in A_H \). By Corollary 1, this case is possible if \( h(\alpha, \delta) > \frac{1}{2\sqrt{\delta - 1}} \).

Proof of Proposition 2. The proof is analogous to proof of Proposition 2. We start with a lemma characterizing the restpoints of the dynamic system defined by Equations 11 and 12.

Lemma 7. If \( \varrho \in (0, \delta^2/4) \), then the set of restpoints in \( \bar{Z} \setminus A_H \) of the dynamic system defined by Equations 11 and 12 is \( \{(1, \mu) : \mu \in [0, \mu_1] \cup [\mu_2, \delta]\} \) where

\[
\mu_1 = \frac{\delta - \sqrt{\delta^2 - 4\varrho}}{2}, \quad \mu_2 = \frac{\delta + \sqrt{\delta^2 - 4\varrho}}{2}.
\]

If \( \varrho \in [\delta^2/4, \delta) \), then the set of restpoints in \( \bar{Z} \setminus A_H \) of the dynamic system defined by Equations 11 and 12 is \( \{(1, \mu) : \mu \in [0, \delta]\} \).

Proof. It is easily shown that \( M^L(a, \mu) = 0 \) for \( (a, \mu) \in \bar{Z} \) iff \( a = 1 \) or \( \mu = 0 \). Since \( A^L(a, 0) > 0 \) for \( a \in [0, 1] \), as \( \psi^L(a, 0) = \varrho > 0 \) for \( a \in [0, 1] \), so any restpoint \( (a, \mu) \in \bar{Z} \) of the system defined by Equations 11 and 12 must satisfy \( a = 1 \). Since \( A^L(1, \mu) = 0 \) holds iff \( \psi^L(1, \mu) \leq 0 \), so the set of restpoints of the system defined by Equations 11 and 12 lying in \( \bar{Z} \setminus A_H \) is the set of all points \( (1, \mu) \) such that \( \mu \in [0, \delta] \) and satisfies the inequality \( \mu^2 - \mu \varrho + \varrho \leq 0 \). Solving it we find that either \( \mu \in [0, \mu_1] \cup [\mu_2, \delta] \) with

\[
\mu_1 = \frac{\delta - \sqrt{\delta^2 - 4\varrho}}{2}, \quad \mu_2 = \frac{\delta + \sqrt{\delta^2 - 4\varrho}}{2}.
\]

if \( \varrho \in (0, \delta^2/4) \), or \( \mu \in [0, \delta] \), if \( \varrho \in [\delta^2/4, \delta) \).
The rest of the argumentation is analogous to that used in proof of Proposition 2 and is based on the observation that the vector field defined by \( \hat{\epsilon} \) and 12 never points outwards on the boundary of \( \mathcal{Z}_{<\delta} \) and it points inwards on the boundary where \( \mu = \delta \) and \( a \in (0,1) \).

In the remaining part of the Appendix we prove Facts 2 and 3 that were used in proof of Lemma 5 to show that construction of region \( \mathcal{R}_\theta \) is feasible.

**Proof of Fact 2** From \( M^L(a,\mu) = 0 \) and \( \varphi_0^L(a,\mu) = 0 \) we have that

\[
a_1^\theta - a_s + \mu_1^\theta - \mu_s = -\theta. \tag{A.20}
\]

Since the function \( a^M(\mu) \) defined implicitly by the equation \( M^L(a,\mu) = 0 \) (c.f. Equation (A.18)) is strictly increasing for \( \mu \in [0,1) \) so, for sufficiently small \( \theta > 0 \) there is \( \mu_\theta \in [0,\mu_s) \) such that Equation (A.20) is satisfied for \( x_1^\theta = (a^M(\mu_\theta),\mu_\theta) \). Now we will show that if \( \theta > 0 \) is sufficiently small, then Point (ii) is satisfied for \( x_1^\theta \).

Expanding \( M^L(a,\mu) \) in a Taylor polynomial around \( x_s \) we get

\[
M^L(a,\mu) = M^L_s(x_s)(a - a_s) + M^L_\mu(x_s)(\mu - \mu_s) + o(a - a_s) + o(\mu - \mu_s).
\]

Since \( x_1^\theta \to x_s \) as \( \theta \to 0 \), so using the Taylor expansion of \( M^L(a,\mu) \) around \( x_s \) for \( \theta \) close to 0 we get

\[
(M^L_s(x_s)(a_1^\theta - a_s) + M^L_\mu(x_s)(\mu_1^\theta - \mu_s) + o(a_1^\theta - a_s) + o(\mu_1^\theta - \mu_s) =
(M^L_s(x_s) - M^L_\mu(x_s))(a_1^\theta - a_s) + M^L_\mu(x_s)(a_1^\theta - a_s + \mu_1^\theta - \mu_s) + o(a_1^\theta - a_s + \mu_1^\theta - \mu_s)) =
(M^L_s(x_s) - M^L_\mu(x_s))(a_1^\theta - a_s) - M^L_\mu(x_s)\theta + o(\theta) = 0.
\]

From this (and by a similar method for \( \mu_1^\theta - \mu_s \) we get

\[
a_1^\theta - a_s = \frac{M^L_s(x_s)}{M^L_s(x_s) - M^L_\mu(x_s)}\theta + o(\theta) > -K\theta,
\]

\[
\mu_1^\theta - \mu_s = -\frac{M^L_\mu(x_s)}{M^L_s(x_s) - M^L_\mu(x_s)}\theta + o(\theta) > -K\theta,
\]

for \( \theta \) small enough (where \( K \) is defined by Equation (A.19)). Since \( \mu_1^\theta < \mu_s \) so, by Observation 2 we have \( \psi^L(x_1^\theta) > 0 \).

Similarly for \( M^L(a,\mu) = 0 \) and \( \varphi^2_\mu(a,\mu) = 0 \) we can find the intersection point \( x_3^\theta \) which satisfies

\[
a_3^\theta - a_s = -\frac{M^L_s(x_s)}{M^L_s(x_s) - M^L_\mu(x_s)}\theta + o(\theta) < K\theta,
\]

\[
\mu_3^\theta - \mu_s = \frac{M^L_\mu(x_s)}{M^L_s(x_s) - M^L_\mu(x_s)}\theta + o(\theta) < K\theta.
\]

for \( \theta \) small enough. Since \( \mu_3^\theta > \mu_s \) so, by Observation 2 we have \( \psi^L(x_3^\theta) > 0 \).

Before we prove Fact 3, we will show a fact characterising partial derivatives of \( \psi^L(a,\mu) \) in the neighbourhood of \( x_s \) and the function \( a^\psi(\mu) \), implicitly defined by the equation \( \psi^L(a,\mu) = 0 \) (see Equation (A.18)), in the neighbourhood of \( \mu_s \).
**Fact 4.** In the neighbourhood of \( \mathbf{x}_s \) it holds that \( \psi^1_\alpha(a, \mu) > 0 \) and \( \psi^1(a, \mu) < 0 \). In particular \( a^\psi(\mu) \), implicitly defined by the equation \( \psi^1(a, \mu) = 0 \), is strictly increasing in the neighbourhood of \( \mu_s \).

**Proof.** Partial derivatives of \( \psi^1(a, \mu) \) are given by the following formulas:

\[
\psi^1_\alpha(a, \mu) = \mu^2, \quad \psi^1_\mu(a, \mu) = 2a\mu - \delta.
\]

Since \( \mu_s > 0 \), so \( \psi^1_\alpha(a, \mu) > 0 \) in the neighbourhood of \( \mathbf{x}_s \). To show that in the neighbourhood of \( \mathbf{x}_s \), \( \psi^1_\mu(a, \mu) < 0 \), we will show that if \( (a, \mu) \) is sufficiently close to \( \mathbf{x}_s \), then \( 2a\mu - \delta < 0 \). Consider a hyperbola \( a = \delta/(2\mu) \). The intersection of this hyperbola and the function \( a^M(a, \mu) \), implicitly defined by the equation \( M^1(a, \mu) = 0 \) (see Equation (A.18)) is at \( a = \frac{\delta}{2(\mu + 1) + \delta/2} \) It can be easily checked that \( \frac{\delta}{2(\mu + 1) + \delta/2} > \delta/(2\mu) > \mu_s \) (c.f. Equation (A.10)). Since \( a^M(a, \mu) \) is strictly increasing for \( \mu \in [0, 1) \), so for \( (a, \mu) \) is sufficiently close to \( \mathbf{x}_s \) it holds that \( 2a\mu - \delta < 0 \).

The fact that \( a^\psi(\mu) \) is strictly increasing in the neighbourhood of \( \mu_s \) follows from the fact that \( a^\psi(\mu) = -\psi^1_\alpha(a^M(\mu), \mu)/\psi^1_\mu(a^M(\mu), \mu) > 0 \) for \( \mu \) in the neighbourhood of \( \mu_s \), as in this case \( (a^M(\mu), \mu) \) is in the neighbourhood of \( \mathbf{x}_s \).

**Remark 3** Throughout the entire proof we will assume that any considered \( \theta \) is small enough to satisfy Fact 2.

For Point 5 we will show that the point \( \mathbf{x}^2_2 \) having the desired properties can be obtained for sufficiently small \( \theta \). Let \( a \in (a_s, a_s + K\theta), \mu \in (\mu_s, \mu_s + K\theta) \) and suppose that \( M^1(a, \mu) > 0 \). Since in this case \( \psi^1_\mu(a, \mu) \leq 0 \) so

\[
M^1(a, \mu) < \sqrt{2\kappa_1\kappa_2f(\alpha)(2K + 1)}\eta(\theta)\theta. \tag{A.21}
\]

Let \( d(\theta) = \sqrt{(2\kappa_1\kappa_2f(\alpha)(2K + 1)}\eta(\theta)\theta \). We will need the following fact.

**Fact 5.** If \( \theta \) is close to 0, then \( \eta(\theta) = o(\theta) > 0 \).

**Proof.** Since, by Fact 4, in the neighbourhood of \( \mathbf{x}_s \) we have \( \psi^1_\alpha(a, \mu) > 0 \) and \( \psi^1_\mu(a, \mu) < 0 \) and \( g \) is strictly increasing for non negative arguments, so for sufficiently small \( \theta \) we have

\[
\max_{(a, \mu) \in C(K\theta)} g(\psi^1(a, \mu)) = g(\max\{-\psi^1(a_s - K\theta, \mu_s + K\theta), \psi^1(a_s - K\theta, \mu_s - K\theta)\}).
\]

One can easily check that

\[
\max\{-\psi^1(a_s - K\theta, \mu_s + K\theta), \psi^1(a_s - K\theta, \mu_s - K\theta)\} = L\theta + o(\theta),
\]

where \( L > 0 \) is some constant. On the other hand using Taylor expansion of \( g \) around 0 we get \( g(x) = g(0) + g'(0)x + o(x) = o(x) \). Thus if \( \theta \) is close to 0, then \( g(\max\{-\psi^1(0, \mu_s + \theta), \psi^1(1, \mu_s - \theta)\}) = o(\theta) > 0, \) as \( g(x) > 0 \) for \( x > 0 \). Hence \( \eta(\theta) = o(\theta) > 0 \).

By Fact 5 for sufficiently small \( \theta \) it holds that \( \eta(\theta) = o(\theta) \). Hence we have \( d(\theta) = o(\theta) \).

Therefore,

\[
\lim_{\theta \to 0} \frac{\sqrt{2\kappa_1\kappa_2f(\alpha)(2K + 1)}\eta(\theta)\theta}{\theta} = \lim_{\theta \to 0} \sqrt{\frac{2\kappa_1\kappa_2f(\alpha)(2K + 1)}{\theta}\eta(\theta)} = 0.
\]
Consider the curve $M^L(a, \mu) = d(\theta)$. Notice that for sufficiently small $\theta > 0$ there is an intersection point $x_0^\theta = (a_0^\theta, \mu_0^\theta)$ of that curve and the curve $\psi^L(a, \mu) = 0$. This is because, by Observation 2, $M^L(a, \mu) > 0$ for sufficiently small $\mu > \mu_s$ and $(a, \mu)$ on the curve $\psi^L(a, \mu) = 0$. It also holds that $x_0^\theta \to x_s$ when $\theta \to 0$. Using Taylor expansion of $M^L(a, \mu)$ around $x_s$, for $\theta$ close to 0 and from the fact that

$$M_\mu^L(x_s)(a_0^\theta - a_s) + M_\mu^L(x_s)(\mu_0^\theta - \mu_s) + o(a_0^\theta - a_s) + o(\mu_0^\theta - \mu_s) = d(\theta),$$

and $M_\mu^L(x_s) \neq 0$ (as $M_\mu^L(a, \mu) < 0$ for all $(a, \mu) \in \mathcal{Z} \setminus \{(1,1)\}$) we get

$$M_\mu^L(x_s)(a_0^\theta - a_s) + \mu_0^\theta - \mu_s = \frac{d(\theta)}{M_\mu^L(x_s)} + o(d(\theta)). \quad (A.22)$$

Moreover, using Taylor expansion of $\psi^L(a, \mu)$ around $x_s$ we get

$$\psi^L_\mu(x_s)(a_0^\theta - a_s) + \psi^L_\mu(x_s)(\mu_0^\theta - \mu_s) + o(a_0^\theta - a_s) + o(\mu_0^\theta - \mu_s) =$$

$$\left(\psi^L_\mu(x_s) - \frac{\psi^L_\mu(x_s) M_\mu^L(x_s)}{M_\mu^L(x_s)}(a_0^\theta - a_s) + \psi^L_\mu(x_s) \frac{M_\mu^L(x_s)}{M_\mu^L(x_s)}(a_0^\theta - a_s) + \frac{M_\mu^L(x_s)}{M_\mu^L(x_s)}(a_0^\theta - a_s) + o(\mu_0^\theta - \mu_s)\right)$$

$$= \left(\psi^L_\mu(x_s) - \frac{\psi^L_\mu(x_s) M_\mu^L(x_s)}{M_\mu^L(x_s)} \left(\frac{M_\mu^L(x_s)}{M_\mu^L(x_s)}(a_0^\theta - a_s) + \frac{M_\mu^L(x_s)}{M_\mu^L(x_s)}(a_0^\theta - a_s) + o(\mu_0^\theta - \mu_s)\right) \right)$$

Thus (and by similar method for $\mu_0^\theta$) we have

$$a_0^\theta - a_s = - \frac{\psi^L_\mu(x_s)}{\psi^L_\mu(x_s) M_\mu^L(x_s) - \psi^L_\mu(x_s) M_\mu^L(x_s)} d(\theta) + o(d(\theta)) < K\theta,$$

$$\mu_0^\theta - \mu_s = \frac{\psi^L_\mu(x_s)}{\psi^L_\mu(x_s) M_\mu^L(x_s) - \psi^L_\mu(x_s) M_\mu^L(x_s)} d(\theta) + o(d(\theta)) < K\theta,$$

for $\theta$ sufficiently small, as $d(\theta) = o(\theta)$.

Let $\theta$ be sufficiently small, so that $a_0^\theta - a_s < K\theta$ and $\mu_0^\theta - \mu_s < K\theta$. Consider a function $c(\mu) = \varphi_\theta^L(a^\delta(\mu), \mu)$, where $a^\delta(\mu)$ is the function implicitly defined by the equation $\psi^L(a, \mu) = 0$ (see Equation (A.18)). Since $c(\mu_s) < 0$, $c(\mu_0^\theta) > 0$ and $c$ is continuous, so there exists $\mu_0^\theta \in (\mu_s, \mu_0^\theta)$ such that $c(\mu_0^\theta) = 0$. From this it follows that $a_0^\theta = a_\psi(\mu_0^\theta) \in (a_s, a_0^\theta)$, because, by Fact 4, $a^\delta(\mu)$ is strictly increasing in the neighbourhood of $\mu_0^\theta$. Thus for $\theta$ sufficiently small the intersection point $x_0^\theta = (a_0^\theta, \mu_0^\theta)$ exists with $a_0^\theta - a_s < K\theta$ and $\mu_0^\theta - \mu_s < K\theta$.

For Point (iii) suppose that $M^L(a, \mu) > 0$ and $\psi^L(a, \mu) > 0$. From the analysis above we know that, for $(a, \mu) \in C(K\theta)$, $\varphi_\theta^L(a, \mu) = 0$ implies $a < a_0^\theta$. Hence $a - a_s < K\theta$. Moreover, from $\varphi_\theta^L(a, \mu) = 0$ we get $\mu - \mu_s \geq -(a_0^\theta - a_s) - \theta \geq -K\theta$.

Points (iii) and (iv) can be shown by analogous arguments.

References


