Games and Economic Behavior 72 (2011) 452–466

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

On the elimination of dominated strategies in stochastic models of evolution with large populations

Christoph Kuzmics *,[1]

J.L. Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2009, United States

ARTICLE INFO

Article history:
Received 14 October 2004
Available online 8 October 2010

JEL classification:
C62
C72
C73

Keywords:
Learning
Experimentation
$S^\infty W$-procedure
Weak dominance
Iterated strict dominance

ABSTRACT

A stochastic myopic best-reply dynamics is said to have property (W), for a given number of players n, if every pure weakly dominated strategy in every n-player game is eliminated in the long-run distribution of play induced by the dynamics. In this paper I give a necessary and sufficient condition that a dynamics has to satisfy in order for it to have property (W). The key determinant is found to be the sensitivity of the learning-rate to small payoff differences, inherent in the dynamics. If this sensitivity is higher than a certain cut-off, which depends on the number of players, then the dynamics satisfies property (W). If it is equal to or below that cut-off, then the dynamics does not satisfy property (W).

1. Introduction

1.1. What this paper does

In this paper I seek to answer the following question. Are there evolutionary dynamics which guarantee the elimination of weakly dominated strategies in all finite normal form games? I make this quest more specific by restricting attention to a class of stochastic myopic best-reply dynamics in the spirit of Kandori et al. (1993) and more specifically Samuelson (1994), inspired by Foster and Young (1990).

Stochastic myopic best-reply dynamics are defined as follows. For a given n-player game, for each player position, there is a finite population of individuals. Individuals, when given an opportunity to learn, play a best-reply to the empirical distribution of play of individuals in the other player positions. Dynamics differ in the learning rate, the probability with which individuals learn. I restrict attention to dynamics in which this learning-rate is a power-function of the payoff-difference between their current strategy and the best reply. The reciprocal of this power is the individuals' sensitivity to payoff-differences. Individuals also experiment or make mistakes.

I investigate the distribution of play, under such dynamics, in the invariant (long-run) distribution of the induced Markov chain, in the limit in which the experimentation-rate tends to zero, population sizes tend to infinity, while their product (the expected number of experiments in any given period) tends to infinity as well.

* Fax: +1 (847) 467 1220.
E-mail address: c-kuzmics@kellogg.northwestern.edu.
[1] I am very grateful to Wiola Dziuda, Florian Herold, Johannes Hörner, Marcin Peski, Hamid Sabourian, Larry Samuelson, Karl Schlag, two anonymous referees and one anonymous associate editor for helpful comments and suggestions.

0899-8256/$ – see front matter
doi:10.1016/j.geb.2010.10.002

A dynamics is then said to have property (W), for a given number of players n, if every weakly dominated strategy in every n-player game is eliminated in the limiting long-run distribution of play induced by the dynamics.

In the main Theorems 4 and 5 of this paper I provide a necessary and sufficient condition for a stochastic myopic best-reply dynamics to have property (W). This condition is in terms of the sensitivity of the learning-rate to payoff differences. If this sensitivity is greater than a certain threshold, which depends on the number of player positions, then all weakly dominated strategies are eliminated. If it is equal to or below that threshold then there is a game with the given number of player positions and a weakly dominated strategy which survives evolution.

1.2. Motivation and related literature

While it is possible to argue that a rational player will not use weakly dominated strategies[2] (see e.g. Dekel and Fudenberg, 1990; Brandenburger, 1992; Börgers, 1994; Gul, 1996; and Ben Porath, 1997), evolutionary models have, so far, mostly failed to support the elimination of weakly dominated strategies. The beginning was very promising, however. An evolutionary stable strategy (ESS), as defined by Maynard Smith and Price (1973), for symmetric 2-player games only, does not contain a weakly dominated strategy. In fact, van Damme (1991) shows that an ESS must be a proper equilibrium, as defined by Myerson (1978). Thus, ESS is a very strong refinement of Nash equilibrium, at the heart of which is not least the elimination of weakly dominated strategies. Thus, it was hoped that evolution could provide a strong tool which would help an applied researcher to pinpoint much more accurately the plausible equilibria of any given game. The concept of an ESS had only two shortcomings. First, it was only defined for symmetric 2-player games, and second, existence of an ESS is not guaranteed in all games. Also the concept of an ESS, being a static one, does not explicitly model the dynamic process of evolutionary selection. In essentially any model in which the dynamic process of selection was modeled directly, evolution failed to generally eliminate weakly dominated strategies. An example of this is given in Nachbar (1990) and a full discussion is provided in Samuelson (1993, 1994).

Samuelson (1993) shows that weakly dominated strategies can survive evolution in deterministic dynamic models such as the replicator dynamics of Taylor and Jonker (1978). Samuelson (1994) in his Theorem 3 shows that they can survive evolution in stochastic models such as that of Kandori et al. (1993), which is inspired by Foster and Young (1990). This is somewhat surprising, as evolutionary models directly, through low-probability mutations in stochastic models, or indirectly, by appealing to notions of stability to perturbations in deterministic models, allow for random mutation or experimentation, which should serve a similar purpose to trembles in strategic refinements. Yet, this is not so.

In deterministic models weakly dominated strategies can survive evolution when all opponents' strategies, against which the weakly dominated strategy performs poorly, vanish much faster than the weakly dominated strategy does (see e.g. Example 3.4 in Weibull, 1995).

In a stochastic finite-population model a la Kandori et al. (1993), weakly dominated strategies may feature in the support of the limiting invariant distribution of play because of the possibility of "evolutionary drift" (see Samuelson, 1994, Theorem 3). Suppose play is currently in a state in which the following is true. First, a given weakly dominated strategy is not played by anyone in the relevant player population. Second, this weakly dominated strategy is, however, a best-reply to the aggregate strategy profile of the opponents in the current state. Thus, the weakly dominated strategy is an alternative best reply in the given state, and if employed by one individual in the relevant population by mutation, there is no evolutionary pressure to remove it. In fact one could have a series of single mutations in this population toward more and more individuals playing the weakly dominated strategy. If nothing else changes, i.e. no other individual in any other population changes strategy, evolutionary pressure does not bear on individuals using the weakly dominated strategy, as it continues to be an alternative best reply in these circumstances.

Thus, Samuelson (1994) demonstrates that it is impossible for a stochastic best-reply model of evolution to eliminate weakly dominated strategies in all games if population sizes are fixed and finite. This implies, in turn, that a necessary condition for a stochastic evolutionary process to eliminate weakly dominated strategies in all games is that population sizes are large (i.e. are taken to infinity). Loosely speaking, large populations have the effect that "evolutionary drift" is less likely. Indeed, in his Theorem 5 Samuelson (1994) then shows that, for a particular 2-player game, a weakly dominated strategy, which is not eliminated under the finite population model, is eliminated in the limit in which population sizes tend to infinity.[3] While taking population sizes to infinity (without further qualification) is sufficient for evolution to eliminate the weakly dominated strategy in that particular game, it is, however, not generally sufficient for the elimination of all weakly

[2] The elimination of weakly dominated strategies is also at the heart of virtually every Nash equilibrium refinement based on strategic considerations. Selten's (1975) trembling hand perfect equilibrium and Myerson's (1978) proper equilibrium are examples for this. Kohlberg and Mertens (1986) even made it a requirement for a solution concept to be called strategically stable that it does not contain weakly dominated strategies.

[3] In the special context of generic extensive form games, Nöldeke and Samuelson (1993) show that similar stochastic models of evolution do not necessarily lead to subgame perfect (i.e. undominated) equilibrium play if population sizes are finite. Hart (2002), in a slightly more biologically flavored stochastic model of evolution, shows that large populations (tending to infinity), while the product of mutation and population size is bounded from below, lead to the evolution of subgame perfect equilibrium play. Kuzmics (2004) shows that the same result can be obtained in the model of Nöldeke and Samuelson (1993) if population sizes are taken to infinity, while the product of experimentation probability and population sizes tend to infinity as well. Thus, both Hart (2002) and Kuzmics (2004) provide sufficient conditions for the evolutionary elimination of weakly dominated strategies in some contexts. They do not cover all normal form games and, as they do not discuss the player's sensitivity to payoff-differences when learning, do not provide necessary conditions for the elimination of weakly dominated strategies.

dominated strategies in all games. I show this by using another example, matching pennies with a weakly dominated strategy. In this game it is necessary to take the limit in which population sizes tend to infinity and the experimentation probability to zero in such a way that the expected number of experiments, which is the product of population size and the probability of experimentation, also tends to infinity. Given this pre-condition, I then move on to show that the crucial determinant for a stochastic best-reply model of evolution to eliminate all weakly dominated strategies in all games with a given number of player positions is that the individuals' sensitivity to payoff-differences (as verbally defined in the previous subsection and formally in Section 2), when learning to play a best reply, is above or below a certain threshold that depends on the number of player positions. If this sensitivity is above the given threshold then evolution eliminates all weakly dominated strategies in all games with the given number of player positions. If it is below that threshold then there is a game with the given number of player positions and a weakly dominated strategy in that game which is not eliminated by evolution.

The structure of this paper is as follows. Section 2 states the model. Section 3 discusses Samuelson's (1994) Theorem 5 as well as the conceptual difficulties involved in proving the main theorems of this paper by means of three examples. Some preliminary results are given in Section 4. Section 5 proves the elimination of strictly dominated strategies in Theorems 1 and 2 before Section 6 provides the main results on the evolutionary elimination of weakly dominated strategies, Theorems 3, 4, and 5, as well as an intuitive sketch of their proofs. Finally, Section 7 provides Theorem 6 showing that provided evolution eliminates all weakly dominated strategies it will then also eliminate all strategies which are not rationalizable (Bernheim, 1984 and Pearce, 1984) in the game obtained from the original game by removing all weakly dominated strategies. That is Theorem 6 provides some evolutionary support for the so-called $S^\infty W$-procedure of Dekel and Fudenberg (1990) in the sense that strategies which are eliminated under the $S^\infty W$-procedure are also eliminated by evolution under the assumptions stated in Theorem 6. Most proofs are relegated to Appendices A–F.

2. Model

For finite population sizes (m_i, see below) the following model is essentially the same as the stochastic[4] best-reply model of Samuelson (1994) and a special case of the evolutionary model of Kandori et al. (1993). The important difference to the model of Samuelson (1994) is that I will assume that any individual's learning rate depends on the difference between the payoff of the strategy currently used by the individual and the largest payoff this individual could obtain in the given situation. This is important when analyzing the limit in which population sizes tend to infinity.

The object of analysis in this paper is a (game/dynamics) pair $\langle \Gamma, \mathcal{D} \rangle$, where $\Gamma = (N, S, u)$ is a finite normal form game, and $\mathcal{D} = (f, \lambda, \kappa)$ is a stochastic best-reply dynamics on Γ. The normal form game $\Gamma = (N, S, u)$ is characterized by the set of n players $N = \{1, \dots, n\}$, the finite set of pure strategy profiles $S = \bigtimes_{i \in N} S_i$ (S_i is player i's set of pure strategies), and the vector of payoff functions $u = (u_1, \dots, u_n)$, where $u_i : S \to \mathbb{R}$. The dynamics $\mathcal{D} = (f, \lambda, \kappa)$ is characterized by the vector $f = (f_1, \dots, f_n)$ of best-reply learning functions (to be specified below), one for each player, the vector $\lambda = (\lambda_1, \dots, \lambda_n)$ of conditional experimentation probabilities (to be specified below), and the vector $\kappa = (\kappa_1, \dots, \kappa_n)$ of conditional learning probabilities (in case of 2 or more alternative best replies, again to be specified below).

In the remainder of this section I will specify the details of the dynamics $\mathcal{D} = (f, \lambda, \kappa)$. Let each player i be replaced by a population of individuals $M(i)$ with population size $m_i = |M(i)|$. Individuals are characterized by the pure strategy they are playing. A state is a characterization for each individual in each population. Let the state space be denoted by Ω.

Individuals in every period t play against every possible configuration of opponents. Between times t and $t+1$ each individual in each population first receives a draw from a Bernoulli random variable either to learn with probability σ or not to learn, and then, regardless of the outcome of the first Bernoulli draw, receives a second draw from an independent Bernoulli variable either to experiment with probability μ or not to experiment.

While μ is assumed to be a constant, the learning rate σ is assumed to be dependent on the payoffs obtainable by the various strategies. Suppose the current state is some $\omega \in \Omega$. Suppose a given agent in population $M(i)$ plays strategy $s \in S_i$ in this state ω. The probability that this agent will learn shall now depend on the payoff-difference between the payoff the agent could get when playing a best-reply (against state ω), and the payoff the agent receives currently. Slightly abusing notation, let the payoff function u_i be extended to the domain $S_i \times \Omega$ as follows. For a given state $\omega \in \Omega$ let $P^{i,s}(\omega)$ denote the proportion of individuals in player position i who play pure strategy $s \in S_i$. Let $P_i(\omega)$ be the vector of these proportions for player position $i \in N$ and let $P_{-i}(\omega) = \bigtimes_{j \in N,\, j \neq i} P_j(\omega)$. Then define $u_i(s, \omega) = u_i(s, P_{-i}(\omega))$. Let, furthermore, $u_i^*(\omega) = \max_{s' \in S_i} u_i(s', \omega)$ denote the maximal payoff an individual in player position i could have achieved given state ω. Then the probability that an agent (currently playing s) switches to a best-reply given state ω is given by $\sigma(s, \omega) = f_i(u_i^*(\omega) - u_i(s, \omega))$, where $f_i : \mathbb{R}_+ \to [0, 1]$.

If an agent learns, the agent chooses a best reply to the aggregate behavior of individuals at time t. If there are multiple best replies the agent chooses one according to the vector of conditional learning probabilities $\kappa = (\kappa_1, \dots, \kappa_n)$, with $\kappa_i = \{\kappa_{i,T_i}\}_{T_i \subset S_i}$ a family of probability distributions such that $\kappa_{i,T_i} : T_i \to R_+$ with $\kappa_{i,T_i}(s) > 0$ for all $s \in T_i$ (full support) and $\sum_{s \in T_i} \kappa_{i,T_i}(s) = 1$ for all $T_i \subset S_i$ and all $i \in N$. If at state ω the set of best-responses for players in position i is given by

[4] This type of stochastic evolutionary model originated in Foster and Young (1990).

$T_i \subset S_i$ then an individual, conditional on switching to a best reply, switches to strategy $s \in T_i$ with probability $\kappa_{i,T_i}(s)$. If she does not learn, the agent continues to play her old strategy.

If the agent receives an experimentation-draw she chooses an arbitrary strategy according to the vector of conditional experimentation probabilities $\lambda = (\lambda_1, \ldots, \lambda_n)$ with $\lambda_i(s) > 0$ for all $s \in S_i$ (full support) and $\sum_{s \in S_i} \lambda_i(s) = 1$ for all $i \in N$. That is, conditional on experimenting, an individual in player position i switches to strategy $s \in S_i$ with probability $\lambda_i(s)$. In the absence of an experimentation-draw the agent does not change her strategy. This completes the description of the stochastic dynamics $\mathcal{D} = (f, \lambda, \kappa)$.

Throughout the paper I restrict attention to a subclass of all such pairs $\langle \Gamma, \mathcal{D} \rangle$ of a game and a dynamics. I assume that f_i, a function from the non-negative part of the real line into the unit interval, satisfies $f_i(0) = 0$, $f_i(x) > 0$ for all $x > 0$, and that f_i is weakly increasing. Typical functions for f_i shall be a step function for which $f_i(x) = \sigma$ (constant) for all $x > 0$, a scaled identity function $f_i(x) = \alpha x$ for some α that guarantees $f_i(x) \in [0, 1]$ for all relevant x, or generally any power function $f_i(x) = \alpha x^\beta$, again with α such that $f_i(x) \in [0, 1]$ for all relevant x. Note that $f_i(0) = 0$ implies that when a learning agent already plays a best reply she is assumed to continue playing it. This assumption, however, is not important for the results of this paper.

For the learning-rate given by a power function $f_i(x) = \alpha x^\beta$, I call $\frac{1}{\beta}$ the **sensitivity** individuals display towards payoff-differences. It is e.g. 1 if the function is linear, 2 if the function is proportional to the square-root function, and infinite if the function is discontinuous at 0.

A given game/dynamics pair, for a given experimentation probability μ and a given vector of population sizes $m = (m_1, \ldots, m_n)$, gives rise to a Markov chain on the state space Ω with transition probability matrix denoted by Q_μ^m. The transition probabilities also vary with different choices of f, λ, and κ. However, as I will study the limit of this process for any fixed f, λ, and κ, but taking μ to zero and m_i to infinity, I suppress f, λ, and κ in the notation.

The Markov chain induced by the above selection–mutation dynamics is aperiodic and irreducible. Hence, it has a unique stationary distribution, which shall be denoted by π_μ^m, and satisfies

$$\pi_\mu^m Q_\mu^m = \pi_\mu^m. \tag{1}$$

3. Motivating examples

A dynamics $\mathcal{D}$ is said to have property (W) for a given number of players n, if every pure weakly dominated strategy in every n-player game is eliminated in the long-run distribution of play induced by the dynamics. In this section I use a series of three examples of increasing complexity to demonstrate the conceptual difficulties in finding necessary and sufficient conditions for a dynamics to exhibit property (W). In some of the statements in the discussion I appeal to theorems from later sections.

The first example is a simple 2×2 game with a weakly dominated strategy that will be eliminated under any of the stochastic best reply models of this paper as long as the experimentation probability tends to zero, even if population sizes are finite. Thus, it demonstrates that while it may be difficult for evolution to eliminate weakly dominated strategies in all games, there are games in which the elimination of weakly dominated strategies is easily achieved by any reasonable model of evolution. The second example is Game G1 of Samuelson (1994), also called Game G1 in this paper. Samuelson's (1994) findings imply that, provided population sizes are finite, the single weakly dominated strategy in that game is not removed by any stochastic best-reply model of evolution. In Theorem 5 Samuelson (1994), furthermore, shows that for any stochastic best-reply model of evolution, if population sizes tend to infinity at the same time as experimentation probabilities tend to zero, the single weakly dominated strategy is eliminated. In the third example, Game G2, finally, I demonstrate that this result is not generally true in all games. Game G2 has a weakly dominated strategy that, under the conditions of Samuelson's (1994) Theorem 5 is not generally eliminated. It turns out, as discussed below, that, in order to obtain conditions under which property (W) holds, conditions need to be placed on the relative speed with which the experimentation probability tends to zero in relation to the speed with which population sizes tend to infinity. I demonstrate that it is essentially necessary to take this limit in such a way that the product of experimentation probability and population size, which corresponds to the expected number of experimentation draws in a population, needs to tend to infinity as well. This discussion, thus, leads me to conclude that in order to obtain necessary and sufficient conditions for a dynamics to satisfy property (W) I need to investigate exactly this limit.

3.1. A simple game

Consider, first, the following 2-player game, denoted Game G0. This game has a weakly dominated strategy, which is eliminated under any of the learning dynamics in this paper as long as the experimentation probability tends to zero, even if population sizes are fixed and finite.

	A	S
B	1, 1	1, 0
W	0, 1	1, 0

Note that S is strictly dominated for player 2. Thus, under any learning dynamics $\mathcal{D}$, by Theorem 2, we have that the expected proportion of S-players tends to 0, in any limit in which $\mu \to 0$. But then W is actually strictly worse than B for players in player position 1. Thus, by the same argument (see also Theorem 6), the expected proportion of W-players tends to 0 as well, as long as $\mu \to 0$. This is true whether or not μm_i tends to infinity or some positive number or even to 0.

3.2. Another simple example

Consider the following 2-player game (G1 in Samuelson, 1994), also denoted Game G1 here.

	A	S
B	$1, 1$	$1, 0$
W	$1, 1$	$0, 0$

Player 2's strategy S is strictly dominated, while player 1's strategy W is weakly dominated. Compared to the previous example, in this example it is harder for evolution to eliminate the weakly dominated strategy W, as this strategy in the limit will be essentially equally good in terms of payoff as the best strategy. In fact, Samuelson's (1994) results imply that, for this game and under any learning dynamics in this paper, weakly dominated strategy W survives evolution if population sizes are taken to be fixed and finite.

Now consider the following (extreme) best-reply dynamics. For $i = 1, 2$ let $f_i(x) = 0$ if $x = 0$ and $f_i(x) = \sigma > 0$ if $x > 0$. Samuelson (1994) in his Theorem 5 shows that the weakly dominated strategy W is eliminated by this dynamics under any limit in which the experimentation-rate (μ) tends to zero and population sizes (m_i) tend to infinity. Thus, this is true regardless of the limiting behavior of the product of the two (μm_i).

Thus, so far it seems that $\mu m_i \to \infty$ is not necessary to guarantee the elimination of weakly dominated strategies. Note, however, that this is due partly to the fact that even in the case when W is a best reply, individuals who do not play it, i.e. who play B, will not learn to play W, as they are already playing a best reply. If we changed the model to allow for arbitrary switches between alternative best replies even when individuals already play a best reply, this result would no longer hold. In that case we would need $\mu m_i \to \infty$ in order to prevent W to essentially ever be a best reply. The next example demonstrates that there are games in which $\mu m_i \to \infty$ is necessary even if we do not change the model.

3.3. Matching pennies with a weakly dominated strategy

Consider the following 2-player game G2 (matching pennies with a weakly dominated strategy).

	H_2	T_2
H_1	$1, -1$	$-1, 1$
T_1	$-1, 1$	$1, -1$
W	$-2, 1$	$1, -1$

Player 1's strategy W is weakly dominated. Note that player 2 is indifferent between player 1's strategies T_1 and W. Consider the following best reply dynamics with $\lambda_{H_2} = \lambda_{T_2} = \lambda_{H_1} = \frac{1}{2}$, $\lambda_{T_1} = \lambda_W = \frac{1}{4}$. Also $f_i(0) = 0$ and $f_i(x) = 1$ if $x > 0$ for $i = 1, 2$. Furthermore, when there are multiple best replies assume that individuals uniformly randomize between them when learning. For all $s \in S_i$ and all $i \in N$ let $P^{i,s}_{\mu,m} : \Omega \to \mathbb{R}_+$ denote the random variable that assigns to states $\omega \in \Omega$ the proportion of individuals in population $M(i)$ who play strategy s. For convenience assume that m_1, m_2 are odd numbers. Let the state space be partitioned into four sets. Let

$$NW = \left\{\omega \in \Omega \;\middle|\; P^{1,H_1}_{\mu,m}(\omega) > \frac{1}{2} \wedge P^{2,H_2}_{\mu,m}(\omega) > \frac{1}{2}\right\},$$

$$NE = \left\{\omega \in \Omega \;\middle|\; P^{1,H_1}_{\mu,m}(\omega) > \frac{1}{2} \wedge P^{2,H_2}_{\mu,m}(\omega) < \frac{1}{2}\right\},$$

$$SE = \left\{\omega \in \Omega \;\middle|\; P^{1,H_1}_{\mu,m}(\omega) < \frac{1}{2} \wedge P^{2,H_2}_{\mu,m}(\omega) < \frac{1}{2}\right\},$$

and

$$SW = \left\{\omega \in \Omega \;\middle|\; P^{1,H_1}_{\mu,m}(\omega) < \frac{1}{2} \wedge P^{2,H_2}_{\mu,m}(\omega) > \frac{1}{2}\right\}.$$

Given the symmetry in this game we must have $\pi_{\mu,m}(NW) = \pi_{\mu,m}(NE) = \pi_{\mu,m}(SE) = \pi_{\mu,m}(SW) = \frac{1}{4}$, for all μ and all (odd) m_1, m_2.

Now consider the following state, denoted ω^*, in which all individuals in population 1 play H_1 and all individuals in population 2 play T_2. Similarly, let $\tilde{\omega}$ be the state in which all individuals in population 1 play T_1 and all individuals in population 2 play T_2.

Lemma 1. *For the game G2 and dynamics such that* $f_i(0) = 0$ *and* $f_i(x) = 1$ *for* $x > 0$ *for both* $i = 1, 2$ *we have* $\pi_\mu^m(\omega^*) \geqslant (1 - \frac{1}{2}\mu)^{m_1}(1 - \frac{1}{2}\mu)^{m_2}\frac{1}{4}$ *and* $\pi_\mu^m(\tilde{\omega}) \geqslant (1 - \frac{1}{2}\mu)^{m_1}(1 - \frac{1}{2}\mu)^{m_2}\frac{1}{4}$.

Proof. Note that for all $\omega \in NW$ we have $Q_\mu^m(\omega \to \omega^*) = (1 - \frac{1}{2}\mu)^{m_1}(1 - \frac{1}{2}\mu)^{m_2}$, which converges to some positive γ in the limit, when $\mu \to 0$, while $\mu m_i \to \delta_i \in (0, \infty)$. Then

$$\begin{aligned}
\pi_\mu^m(\omega^*) &= \sum_{\omega \in \Omega} \pi_\mu^m(\omega) Q_\mu^m(\omega \to \omega^*) \\
&\geqslant \sum_{\omega \in NW} \pi_\mu^m(\omega) Q_\mu^m(\omega \to \omega^*) \\
&\geqslant \left(1 - \frac{1}{2}\mu\right)^{m_1}\left(1 - \frac{1}{2}\mu\right)^{m_2} \sum_{\omega \in NW} \pi_\mu^m(\omega) \\
&\geqslant \left(1 - \frac{1}{2}\mu\right)^{m_1}\left(1 - \frac{1}{2}\mu\right)^{m_2} \frac{1}{4}.
\end{aligned}$$

The proof that $\pi_\mu^m(\tilde{\omega}) \geqslant (1 - \frac{1}{2}\mu)^{m_1}(1 - \frac{1}{2}\mu)^{m_2}\frac{1}{4}$ is completely analogous. □

Proposition 1. *For the game G2 and dynamics such that* $f_i(0) = 0$ *and* $f_i(x) = 1$ *for* $x > 0$ *for both* $i = 1, 2$ *we have* $\lim_{\mu \to 0,\, \mu m_i \to \delta_i} \mathbb{E}[P_{\mu,m}^{1,W}] > 0$ *if* $\delta_i \in (0, \infty)$.

Proof. By the law of iterated expectation we have

$$\begin{aligned}
\mathbb{E}\big[(P_{\mu,m}^{1,W})_{t+1}\big] &= \mathbb{E}\big[\mathbb{E}\big[(P_{\mu,m}^{1,W})_t + dP_{\mu,m}^{1,W} \big| \omega_t\big]\big] \\
&\geqslant \pi_\mu^m(\tilde{\omega})\mathbb{E}\big[(P_{\mu,m}^{1,W})_t + dP_{\mu,m}^{1,W} \big| \omega_t = \tilde{\omega}\big] + \pi_\mu^m(\omega^*)\mathbb{E}\big[(P_{\mu,m}^{1,W})_t + dP_{\mu,m}^{1,W} \big| \omega_t = \omega^*\big] \\
&\geqslant \pi_\mu^m(\tilde{\omega})\mathbb{E}\big[dP_{\mu,m}^{1,W} \big| \omega_t = \tilde{\omega}\big] + \pi_\mu^m(\omega^*)\mathbb{E}\big[dP_{\mu,m}^{1,W} \big| \omega_t = \omega^*\big] \\
&\geqslant \left(1 - \frac{1}{2}\mu\right)^{m_1}\left(1 - \frac{1}{2}\mu\right)^{m_2} \frac{1}{2}\left(\frac{1}{2}\big(1 - \mu(1 - \lambda_W)\big)\right),
\end{aligned}$$

which tends to $\frac{1}{4}\gamma$ for some $\gamma > 0$ as $\mu \to 0$ and $\mu m_i \to \delta_i$, and where the last inequality follows from Lemma 1 and the observation that $\mathbb{E}[dP_{\mu,m}^{1,W} | \omega_t = \omega^*] = \mathbb{E}[dP_{\mu,m}^{1,W} | \omega_t = \tilde{\omega}] = \frac{1}{2}(1 - \mu(1 - \lambda_W))$ by an analogous argument as in the proof of Theorem 1. □

Thus, Game G2 has a weakly dominated strategies, which is eliminated under the given dynamics in the limit in which $\mu m_i \to \infty$ (by Theorem 4), but not in the limit in which $\mu m_i \to \delta < \infty$ (by Proposition 1). The dynamics is as in Samuelson (1994). Thus, this example demonstrates that Samuelson's (1994) Theorem 5 does not extend to all games. That is, it is not enough to look at the limit in which $\mu \to 0$ while $m_i \to \infty$ without specifying the limiting behavior of the product μm_i. In fact, this implies that taking $\mu \to 0$ and $m_i \to \infty$ together is not sufficient to yield a well-defined limit in general games, i.e. the limit inferior and limit superior are not equal. In the main part of the paper I, thus, focus attention on the limiting case in which the experimentation probability μ tends to zero, while the expected number of experimentation draws μm_i in any given period tends to infinity. This provides a well-defined limit in all cases I consider.

4. Preliminary results

This section provides some preliminary results needed for the analysis, and, before that, some additional notation. Let $i \in N$ and $s \in S_i$. Let $\Lambda_k^{i,s}$ denote the set of states in which the proportion of individuals at population $M(i)$ playing strategy s is $\frac{k}{m_i}$. Let $\Phi_\tau^{i,s} = \bigcup_{k \leqslant \tau m_i} \Lambda_k^{i,s}$ denote the set of states in which not more than a proportion of τ individuals play s at player population $M(i)$. Let $P_{\mu,m}^{i,s} : \Omega \to \mathbb{R}$ denote a random variable (given probability space (Ω, π_μ^m)) such that $P_{\mu,m}^{i,s}(\omega)$ denotes the proportion of s-players in population $M(i)$ given state ω. Thus $P_{\mu,m}^{i,s}(\omega) = \frac{k}{m_i}$ if $\omega \in \Lambda_k^{i,s}$.

Note that $\pi_\mu^m(P_{\mu,m}^{i,s} \leqslant \epsilon) = \pi_\mu^m(\Phi_\epsilon^{i,s})$. Throughout this section the conditional mutation-probability vector, λ is arbitrary. Hence, the results hold for any such λ. Then λ_s shall denote the probability λ puts on pure strategy $s \in S_i$.

Let $\rho_\mu^m = (\mu, \frac{1}{m_1\mu}, \ldots, \frac{1}{m_n\mu})$ and let $\rho_\mu^m \to 0$ mean that each component of ρ_μ^m tends to zero.

Lemma 2. *Let* $\langle \Gamma, \mathcal{D} \rangle$ *be an arbitrary game/dynamics pair. Let* $s \in S_i$ *be an arbitrary pure strategy of player* i *in the game. Then*

1. $\pi_\mu^m(\Lambda_0^{i,s}) \leqslant (1 - \lambda_s\mu)^{m_i}$;

2. *this bound is tight* (*achieved by some* $\langle \Gamma, \mathcal{D}\rangle$);
3. *and* $\lim_{\rho_\mu^m \to 0} \pi_\mu^m(\Lambda_0^{i,s}) = 0$.

Proof. Points (1) and (2) follow from Lemma 3 in Appendix A. From (1) it follows that $\lim_{\rho_\mu^m \to 0} \pi_\mu^m(\Lambda_0^{i,s}) \leqslant \lim_{\rho_\mu^m \to 0}(1 - \lambda_s \mu)^{m_i}$, which is equal to zero. This proves point (3). □

Note that the limit I am considering here is the only limit in which I can guarantee for any finite normal form game and for any choice of learning functions $f = (f_1, \ldots, f_n)$, that every pure strategy is played by at least one person in the game. To see this, suppose that $\mu \to 0$ and $\mu m_i \to \delta < \infty$. But then from point (2) of Lemma 2 it follows that for some game, dynamics, and pure strategy s it is true that $\lim(1 - \lambda_s \mu)^{m_i} = \lim(1 - \lambda_s \frac{\delta}{m_i})^{m_i} = e^{-\lambda_s \delta} > 0$. Hence, under this limit, one cannot guarantee that a strictly dominated strategy s is always played by at least 1 person. The following corollary is immediate from Lemma 2.

Corollary 1. *Let* $\langle \Gamma, \mathcal{D}\rangle$ *be an arbitrary game/dynamics pair. Denote by* Ψ *the set of states, in which there is a population such that at least one strategy is not played by any individual at this population, i.e.*

$$\Psi = \bigcup_{i=1}^{n} \bigcup_{x \in S_i} \Lambda_0^{i,x}. \tag{2}$$

Then

$$\lim_{\rho_\mu^m \to 0} \pi_\mu^m(\Psi) = 0. \tag{3}$$

5. Strictly dominated strategies

So far we know that, in the limit considered here, every strictly dominated strategy will be played by at least one person. In this section I am interested in the expected number and proportion of people who play any given strictly dominated strategy. Recall that $P_{\mu,m}^{i,s}(\omega)$ denotes the proportion of s-players in population $M(i)$ given state ω.

Let $s \in S_i$ be a strictly dominated strategy. Then the difference between the payoff derived from using strategy s and the maximal obtainable payoff in a given state ω must be positive. That is, $u_i^*(\omega) - u_i(s, \omega) > 0$. In fact we must have that $\min_{\omega \in \Omega}(u_i^*(\omega) - u_i(s, \omega)) = a > 0$. But then under the assumptions about f_i we must have that there is a $\tilde{\sigma} > 0$ such that $\sigma(s, \omega) = f_i(u_i^*(\omega) - u_i(s, \omega)) \geqslant \tilde{\sigma}$ for all $\omega \in \Omega$. On the other hand, we, of course, have that $\sigma(s, \omega) \leqslant 1$ for all $\omega \in \Omega$. In the following the expectation $\mathbb{E}$ is always understood to be the expectation given the invariant distribution π_μ^m. Most proofs are in Appendices A–F.

Theorem 1. *Let* $s \in S_i$ *be a strictly dominated strategy.*[5] *Then*

$$\mu \lambda_s \leqslant \mathbb{E}\big[P_{\mu,m}^{i,s}\big] \leqslant \frac{\mu \lambda_s}{\tilde{\sigma}(1 - \mu) + \mu}.$$

Note that the expectation in Theorem 1 does not depend on the population size. Hence, in any limit in which μ tends to zero, regardless of the limiting behavior of population sizes m_i, we must have that the expected proportion of s-players tends to zero. In the case of fixed population sizes this implies that not only the expected proportion, but also the expected number of s-players tends to zero. In fact this also implies that in this limit (with fixed m_i) the event that no individual plays s has probability 1.

Theorem 1 has the following corollary, which I will also call a theorem, which is somewhat of an analogue to Proposition 5.6 in Weibull (1995), due to Samuelson and Zhang (1992), which proves the same in the context of deterministic payoff-monotonic dynamics.

Theorem 2. *Let* $s \in S_i$ *be a strictly dominated strategy. Then*

$$\lim_{\mu \to 0} \mathbb{E}\big[P_{\mu,m}^{i,s}\big] = 0.$$

[5] In fact this theorem extends to any pure strategy which is never a best-reply. In 2-player games a strategy is strictly dominated if and only if it is a never best-reply. In more than 2 player games every strictly dominated strategy is obviously never a best reply, while there may be a strategy which is never a best reply yet not strictly dominated (see e.g. Ritzberger, 2002, Example 5.7).

Proof. Immediate from Theorem 1. □

Theorem 1, thus, implies, for the limit I consider in this paper, where μ tends to zero while μm_i tends to infinity, that the expected number of s-players tends to infinity, while the expected proportion tends to zero.

6. Weakly dominated strategies

Let $w \in S_i$ be a weakly dominated strategy which is not strictly dominated. Let w be in fact weakly dominated by some mixed strategy $x \in \Delta(S_i)$. We then have that $u_i^*(\omega) - u_i(w,\omega) \geqslant u_i(x,\omega) - u_i(w,\omega) \geqslant 0$, where $u_i^*(\omega)$ is, as previously defined, the maximal payoff a player in player position i could achieve, by playing a best-reply, given state ω. Let $S_{-i} = \bigtimes_{j\neq i} S_j$. Now, by definition, for any $x \in \Delta(S_i)$,

$$u_i(x,\omega) = \sum_{s_{-i}\in S_{-i}} u_i(x,s_{-i}) P_{\mu,m}^{-i,s_{-i}}(\omega),$$

where $P_{\mu,m}^{-i,s_{-i}}(\omega) = \prod_{j\neq i} P_{\mu,m}^{j,s_j}(\omega)$, where s_j is player j's part of the strategy combination s_{-i}. Given that we have that

$$u_i(x,\omega) - u_i(w,\omega) = \sum_{s_{-i}\in S_{-i}} \big(u_i(x,s_{-i}) - u_i(w,s_{-i})\big) P_{\mu,m}^{-i,s_{-i}}(\omega),$$

and, given that all elements in the sum are non-negative,

$$u_i^*(\omega) - u_i(w,\omega) \geqslant \big(u_i(x,s_{-i}) - u_i(w,s_{-i})\big) P_{\mu,m}^{-i,s_{-i}}(\omega) \tag{4}$$

for any $s_{-i} \in S_{-i}$.

By definition of a weakly dominated strategy we know that there must be at least one strategy combination s_{-i} such that $u_i(x,s_{-i}) > u_i(w,s_{-i})$. The prevalence of these strategy combinations will then be the determinant as to whether this weakly dominated strategy will or will not survive evolution as modeled in this paper. For the given weakly dominated strategy $w \in S_i$ let $A_{-i}(w) \subset S_{-i}$ be the set of all these strategy combinations against which x does strictly better than w, i.e. $A_{-i} = \{s_{-i} \in S_{-i} \mid u_i(x,s_{-i}) > u_i(w,s_{-i})\}$. Let $P_{\mu,m}^{-i,A_{-i}}(\omega) = \sum_{s_{-i}\in A_{-i}} P_{\mu,m}^{-i,s_{-i}}(\omega)$. The following theorem is somewhat of an analogue to Proposition 5.8 in Weibull (1995), which proves the same in the context of 2-player games and deterministic payoff-linear dynamics.

Theorem 3. *Let $\langle \Gamma, \mathcal{D}\rangle$ be an arbitrary game/dynamics pair. Let $w \in S_i$ be weakly dominated. Then* $\lim_{\rho_\mu^m \to 0} \mathbb{E}[P_{\mu,m}^{i,w} P_{\mu,m}^{-i,A_{-i}}] = 0$.

Theorem 4. *Let $\Gamma = (N,S,u)$ be an n-player game, i.e. $|N| = n$. Let the learning function f_i for player i be $f_i(x) = \alpha x^\beta$ for some $\alpha > 0$. Let $w \in S_i$ be a weakly dominated strategy. If the sensitivity to payoff-differences satisfies $\frac{1}{\beta} > n-1$ then* $\lim_{\rho_\mu^m \to 0} \mathbb{E}[P_{\mu,m}^{i,w}] = 0$.

Theorem 4 provides sufficient conditions on the learning function f_i under which any weakly dominated strategy in any finite n-player normal form game is eliminated in the course of evolution. In fact this condition is also necessary in the following sense.

Theorem 5. *Let $f_i(x) = x^\beta$ such that the sensitivity to payoff-differences satisfies $\frac{1}{\beta} \leqslant n-1$. Then there is a finite n-player normal form game, a set of learning functions $\{f_j\}_{j\neq i}$, and a weakly dominated strategy w for player i such that* $\lim_{\rho_\mu^m \to 0} \mathbb{E}[P_{\mu,m}^{i,w}] > 0$.

The proofs of Theorems 4 and 5 are given in Appendices E and F through a series of lemmas leading to the two proofs. While making sure that every detail works out is somewhat tedious, the intuition behind these results is relatively straightforward, and shall be given here.

Consider an n-player game in which player i has a pure weakly dominated strategy w, weakly dominated by another strategy x, which is always best. Suppose that there is only one pure strategy combination $s_{-i} = \bigtimes_{j\neq i} s_j$ of the opponents against which any player i's pure strategy w is actually strictly worse than x. Suppose, furthermore, that all opponents of player i find their pure strategy s_j strictly dominated. Suppose, finally, that all individuals in opponent-populations use an extreme learning function $f_j(x) = 1$ if $x > 0$ and $f(0) = 0$. This creates the environment that makes the survival of weakly dominated strategy w most likely.[6]

[6] I actually do not prove this statement, although I conjecture it is true. I find it sufficient to prove that this is the environment that makes the survival of any strictly dominated strategy s_j of the opponents least likely. This is made precise in Lemma 3 and then has consequences for the survival/elimination of the weakly dominated strategy w as eventually used in Lemma 7.

Then by Theorem 1 the expected proportion of individuals playing s_j is of the order of μ. The expected frequency with which a player i encounters opponent profile s_{-i} is then of the order of μ^{n-1}. Given, the above assumptions, the frequency of opponent profile s_{-i} is also the payoff-difference for players i between using x and w. Thus a player-i individual, currently playing w will learn to use something else with probability given by f_i evaluated at the frequency of s_{-i}. In expectation this frequency is of the order of μ^{n-1}. Applying $f_i(x) = \alpha x^\beta$, the expected learning rate away from w is of the order of $\mu^{\beta(n-1)}$. This is so, because the variance of this frequency is an order of magnitude smaller than the expectation. This is demonstrated in Lemma 4 and is only true under the limiting condition $\mu m_i \to \infty$. This is important in the proof, as otherwise, by Jensen's inequality the expected learning rate could be well below the learning rate evaluated at the expected payoff-difference. The exact consequences of Lemma 4 which are used in the eventual proof of Theorem 4, and are derived from Chebyshev's inequality, are given in Lemmas 5 and 6.

Given all this, we thus have an evolutionary force away from w given roughly by this probability of the order of $\mu^{\beta(n-1)}$. Of course, through mutations we also have an evolutionary force towards w, which is roughly given by a probability of the order of μ. Thus, finally, if $\beta(n-1) < 1$ the evolutionary force away from w is stronger than the one towards. This leads to the elimination of the weakly dominated strategy. This is made precise in the proof of Theorem 4. On the other hand, if $\beta(n-1) \geqslant 1$ the two forces are at best equivalent or the force towards w might even be stronger. Then, the weakly dominated strategy w cannot be eliminated by evolution. This is made precise in the proof of Theorem 5.

Note that Theorems 4 and 5, for the special case of 2-player games, imply the following. If the learning-rate depends on the payoff differences in a linear fashion evolution does not necessarily eliminate weakly dominated strategies. If this learning function, however, is a power function with any power less than 1, evolution does eliminate all weakly dominated strategies. The learning rate, thus, does not need to be discontinuous, but needs to have infinite slope at a payoff-difference of 0.

7. $S^\infty W$-procedure

I now turn to a brief discussion about which other strategies will have to be eliminated by evolution as modeled in this paper, supposing evolution eliminates all weakly dominated strategies. In the previous section I investigated under what circumstances, for a given strategy $w \in S_i$, does $\mathbb{E}[P^{i,w}_{\mu,m}]$ tend to zero as ρ^m_μ tends to 0. It is immediate that whenever $\lim_{\rho^m_\mu \to 0} \mathbb{E}[P^{i,w}_{\mu,m}] = 0$ then it must be true that, for any $\epsilon \in (0,1)$ we have $\pi^m_\mu(P^{i,x}_{\mu,m} \leqslant \epsilon) = \pi^m_\mu(\Phi^{i,x}_\epsilon)$ tends to 1 in the limit. This means that with probability 1 the proportion of individuals playing this strategy w is below any $\epsilon > 0$. Given this, however, it must be true that strategies which are strictly dominated once all weakly dominated strategies are thus eliminated, must also be eliminated in the course of evolution.

Let Γ^1 denote the game which remains when all such weakly dominated strategies are eliminated. That is, Γ^1 is derived from Γ by reducing each player's pure strategy set by all weakly dominated strategies, while the payoff function is the same (with restricted domain). Let S^1_i denote the restricted strategy set for player i. If indeed all weakly dominated strategies are eliminated, then strategies which are strictly dominated in Γ^1 must also disappear in the limit I consider. In fact, this argument can be iterated any finite number of times. A strategy which survives the iterated deletion of never best replies is called rationalizable (Bernheim, 1984 and Pearce, 1984). Let a strategy which is rationalizable in the game obtained from the original by deletion of all weakly dominated strategies be termed **strongly rationalizable**. We then have the following

Theorem 6. *For $i \in N$, let $s \in S_i$ be a strategy which is not strongly rationalizable. Whenever $\lim_{\rho^m_\mu \to 0} \mathbb{E}(P^{j,w}_{\mu,m}) = 0$ for every weakly dominated strategy $w \in S_j$ for every player j, then*

$$\lim_{\rho^m_\mu \to 0} \mathbb{E}\big(P^{i,s}_{\mu,m}\big) = 0. \tag{5}$$

Proof. Note that in Γ^1 there is a payoff-wedge between strategy s and the strategy by which it is strictly dominated. But as all the strategies which are available only on Γ but not Γ_1 are played by a vanishing fraction in the limit, this payoff-wedge is present with probability 1. But then a straightforward adaptation of the proof of Theorem 1 yields the result. This argument can be iterated any finite number of times. □

While epistemic conditions for the use of what has been termed the $S^\infty W$-procedure, which stands for the deletion of first all weakly dominated strategies and then iteratively all strictly dominated strategies, have been identified by Dekel and Fudenberg (1990), Brandenburger (1992), Börgers (1994), Gul (1996), and Ben Porath (1997), the above theorem provides an evolutionary justification for its use in the sense that every strategy which is eliminated by the $S^\infty W$-procedure will also be eliminated by evolution under the stated assumptions. The plausibility of this justification depends only on the plausibility of the degree of sensitivity in payoff differences required to eliminate all weakly dominated strategies.

Appendix A. Lemma needed to prove Lemma 2

The following result states that the distribution of the proportion with which any given strategy is played for any given game/dynamics pair under the invariant distribution first order stochastically dominates the distribution of the proportion of a strictly dominated strategy under the most extreme dynamics.

Lemma 3. *Let $\langle \Gamma, \mathcal{D}\rangle$ with $\Gamma = (N, S, u)$ and $\mathcal{D} = (f, \lambda, \kappa)$ be an arbitrary game/dynamics pair. Let $\langle \hat{\Gamma}, \hat{\mathcal{D}}\rangle$ with $\hat{\Gamma} = (N, S, \hat{u})$ and $\hat{\mathcal{D}} = (\hat{f}, \lambda, \kappa)$ be another game/dynamics pair with the properties that $\hat{u}'_j = \hat{u}_j$ for all $j \neq i$, $\hat{u}_i(s', \cdot) = u_i(s', \cdot)$ for all $s' \in S_i$ with $s' \neq s$, and $\hat{u}_i(s, \cdot) = u_i(s, \cdot) - A$, where A is large enough to make s strictly dominated in $\hat{\Gamma}$. Let furthermore $\hat{f}_j = f_j$ for all $j \neq i$ and $\hat{f}_i(0) = 0$ and $\hat{f}_i(x) = 1$ if $x > 0$. Let π^m_μ and $\hat{\pi}^m_\mu$ denote the invariant distributions induced by the two processes respectively. Then $\pi^m_\mu(P^{i,s}_{\mu,m} \geqslant \frac{j}{m_i}) \geqslant \hat{\pi}^m_\mu(P^{i,s}_{\mu,m} \geqslant \frac{j}{m_i})$ for all $j \in \{0, \ldots, m_i\}$.*

Proof. Let $A^j = \{\omega \in \Omega \mid P^{i,s}_{\mu,m}(\omega) \geqslant \frac{j}{m_i}\}$. Then $\hat{Q}(\omega, A^j) = \sum_{k \geqslant j} (\lambda_s\mu)^k (1-\lambda_s\mu)^{m_i-k}\binom{m_i}{k}$, which is independent of ω. Thus we can denote this expression by $\hat{Q}(A^j)$. Furthermore we clearly have that $Q(\omega, A^j) \geqslant \hat{Q}(A^j)$ for all ω. Then, $\pi^m_\mu(P^{i,s}_{\mu,m} \geqslant \frac{j}{m_i}) = \sum_{\omega\in\Omega} \pi^m_\mu(\omega) Q(\omega, A^j) \geqslant \sum_{\omega\in\Omega} \pi^m_\mu(\omega)\hat{Q}(A^j) = \hat{Q}(A^j) = \hat{\pi}^m_\mu(P^{i,s}_{\mu,m} \geqslant \frac{j}{m_i})$. □

Appendix B. Proof of Theorem 1

Let $\{\Omega \times \Omega, \Pr\}$[7] denote a probability space, where Pr is such that[8] $\Pr(\omega, \omega') = \pi^m_\mu(\omega)(Q^m_\mu)_{\omega,\omega'}$ for all $(\omega, \omega') \in \Omega \times \Omega$.

Let $(P^{i,s}_{\mu,m})_t$ denote the proportion of s-players in population $M(i)$ at time t. Let $dP^{i,s}_{\mu,m}$ denote the change in proportion of s-players in population $M(i)$ between times t and $t+1$. That is,

$$\left(P^{i,s}_{\mu,m}\right)_{t+1} = \left(P^{i,s}_{\mu,m}\right)_t + dP^{i,s}_{\mu,m}. \tag{6}$$

If $(P^{i,s}_{\mu,m})_t$ is distributed according to the invariant distribution π^m_μ then so is $(P^{i,s}_{\mu,m})_{t+1}$ and, hence, the expected value $\mathbb{E}[dP^{i,s}_{\mu,m}] = 0$. Also all these three random variables are measurable given the above stated probability space.

By the law of iterated expectations the last expectation can be written as $\mathbb{E}[dP^{i,s}_{\mu,m}] = \mathbb{E}[\mathbb{E}\, dP^{i,s}_{\mu,m} | (P^{i,s}_{\mu,m})_t]$, and hence

$$0 = \mathbb{E}\left[dP^{i,s}_{\mu,m}\right] = \sum_{k=0}^{m_i} \pi^m_\mu\left(\Lambda^{i,s}_k\right)\mathbb{E}\left(dP^{i,s}_{\mu,m} \middle| \left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}\right). \tag{7}$$

Conditional on $(P^{i,s}_{\mu,m})_t = \frac{k}{m_i}$, the change $dP^{i,s}_{\mu,m}$ can be viewed as the difference of two random variables $\frac{Y}{m_i}$ and $\frac{X}{m_i}$, again both measurable given our specification of the probability space above, where $X(\omega, \omega')$ is the number of individuals at $M(i)$ who, in the transition from ω to ω', switch strategy from something other than s to s, and $Y(\omega, \omega')$ is the number of individuals at $M(i)$ who, in the transition from ω to ω', switch strategy from s to anything other than s. Conditional on $(P^{i,s}_{\mu,m})_t = \frac{k}{m_i}$, X and Y are binomially distributed, i.e. $X \sim \mathrm{Bin}(m_i - k, \mu\lambda_s)$ and $Y \sim \mathrm{Bin}(k, \sigma(s,\omega)(1-\mu) + \mu(1-\lambda_s))$. Given that s is a strictly dominated strategy we know that $\tilde{\sigma} \leqslant \sigma(s,\omega) \leqslant 1$. Given all this, the term

$$\mathbb{E}\left(dP^{i,s}_{\mu,m} \middle| \left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}\right)$$

is the difference between the expectation of these two binomial variables, divided by m_i, and bounded from above by

$$\mathbb{E}\left(dP^{i,s}_{\mu,m} \middle| \left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}\right) \leqslant \mu\lambda_s - \frac{k}{m_i}\left(\tilde{\sigma}(1-\mu) + \mu\right)$$

and from below by

$$\mathbb{E}\left(dP^{i,s}_{\mu,m} \middle| \left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}\right) \geqslant \mu\lambda_s - \frac{k}{m_i}.$$

Plugging the lower bound back into Eq. (7) we obtain

$$0 \geqslant \left[\tilde{\sigma}(1-\mu) + \mu\right] \sum_{k=0}^{m_i} \frac{k}{m_i}\pi^m_\mu\left(\Lambda^{i,s}_k\right) - \mu\lambda_s, \tag{8}$$

[7] As the state space is finite I omit the sigma-algebra, which can be taken as the set of all subsets of $\Omega \times \Omega$, in the description of the probability space.

[8] Given the axioms of a probability measure this is sufficient to uniquely define Pr.

which by the assumptions of the lemma and by the fact that $\sum_{k=0}^{m_i} \frac{k}{m_i}\pi_\mu^m(\Lambda_k^{i,s}) = \mathbb{E}(P_{\mu,m}^{i,s})$ yields $\mathbb{E}(P_{\mu,m}^{i,s}) \leqslant \frac{\mu\lambda_s}{\tilde{\sigma}(1-\mu)+\mu}$. Doing the same with the upper bound yields $\mathbb{E}(P_{\mu,m}^{i,s}) \geqslant \mu\lambda_s$. □

Appendix C. Proof of Theorem 3

I shall prove the statement here under the assumption that $\lim_{\rho_\mu^m\to 0}\mathbb{E}P_{\mu,m}^{-i,A_{-i}} > 0$. Note that if $\lim_{\rho_\mu^m\to 0}\mathbb{E}P_{\mu,m}^{-i,A_{-i}} = 0$ then $\lim_{\rho_\mu^m\to 0}\mathbb{E}[P_{\mu,m}^{i,w}P_{\mu,m}^{-i,A_{-i}}] \leqslant \lim_{\rho_\mu^m\to 0}\mathbb{E}[P_{\mu,m}^{-i,A_{-i}}] = 0$ directly.

Reconsider Eq. (6), now for strategy w, $(P_{\mu,m}^{i,w})_{t+1} = (P_{\mu,m}^{i,w})_t + dP_{\mu,m}^{i,w}$. Let $B^w \subset \Omega$ denote the set of states in which w is a best reply for individuals at population $M(i)$. The expectation $\mathbb{E}[dP_{\mu,m}^{i,w}]$, which as in the proof of Theorem 1 must be zero, using the law of iterated expectations, can be written as

$$\mathbb{E}\big[dP_{\mu,m}^{i,w}\big] = \pi_\mu^m\big(B^w\big)\mathbb{E}\big[dP_{\mu,m}^{i,w}\big|B^w\big] + \big(1-\pi_\mu^m\big(B^w\big)\big)\mathbb{E}\big[dP_{\mu,m}^{i,w}\big|B^{w,c}\big], \tag{9}$$

where $B^{w,c}$ is the complement of B^w in Ω. Much like in the proof of Theorem 1 the expectation $\mathbb{E}[dP_{\mu,m}^{i,w}|B^{w,c}]$ can be rewritten with the recurrent use of the law of iterated expectations as

$$\mathbb{E}\big[\mathbb{E}\big[dP_{\mu,m}^{i,w}\big|B^{w,c}\wedge\big(P_{\mu,m}^{i,w}\big)_t\wedge\big(P_{\mu,m}^{-i,A_{-i}}\big)_t\big]\big].$$

Given $B^{w,c}$ w is not a best reply and we can, again as in Theorem 1, write this conditional expectation as the expectation of the difference between two random variables $\frac{Y}{m_i}$ and $\frac{X}{m_i}$, with the same interpretation as in Theorem 1. Given $P_{\mu,m}^{i,w} = \frac{k}{m_i}$ we still have $X \sim \mathrm{Bin}(m_i - k, \mu\lambda_w)$ as well as $Y \sim \mathrm{Bin}(k, \sigma(w,\omega)(1-\mu)+\mu(1-\lambda_w))$. Of course, $\sigma(w,\omega) = f_i(u_i^*(\omega) - u_i(w,\omega))$ by the model assumptions. Given the definition of $A_{-i}(w)$ we have that $\min_{s_{-i}\in S_{-i}} u_i(x,s_{-i}) - u_i(w,s_{-i}) = a > 0$. Using inequality (4), and the fact that f_i is weakly increasing, we obtain that $\sigma(w,\omega) \geqslant f_i(aP_{\mu,m}^{-i,A_{-i}}(\omega))$. Putting all this together we obtain

$$\mathbb{E}\big[\mathbb{E}\big[dP_{\mu,m}^{i,w}\big|B^{w,c}\wedge\big(P_{\mu,m}^{i,w}\big)_t\wedge\big(P_{\mu,m}^{-i,A_{-i}}\big)_t\big]\big] \geqslant \mathbb{E}\big[P_{\mu,m}^{i,w}\big(f_i\big(aP_{\mu,m}^{-i,A_{-i}}\big)(1-\mu)+\mu\big) - \mu\lambda_w\big]. \tag{10}$$

By the fact that w is weakly dominated we have that $B^w \subset \Psi$, and, hence, by Corollary 1 we have that $\lim_{\rho_\mu^m\to 0}\pi_\mu^m(B^w) = 0$. Hence, from Eq. (9) we have that $\lim_{\rho_\mu^m\to 0}\mathbb{E}[dP_{\mu,m}^{i,w}|B^{w,c}] = 0$. But then, by inequality (10), we have that

$$\lim_{\rho_\mu^m\to 0}\mathbb{E}\big[P_{\mu,m}^{i,w}\big(f_i\big(aP_{\mu,m}^{-i,A_{-i}}\big)(1-\mu)+\mu\big) - \mu\lambda_w\big] \leqslant 0,$$

which, given the assumption that $\lim_{\rho_\mu^m\to 0}\mathbb{E}P_{\mu,m}^{-i,A_{-i}} > 0$ and, hence, that $\lim_{\rho_\mu^m\to 0}\mathbb{E}f_i(aP_{\mu,m}^{-i,A_{-i}}) > 0$ implies that

$$\lim_{\rho_\mu^m\to 0}\mathbb{E}\big[P_{\mu,m}^{i,w}f_i\big(aP_{\mu,m}^{-i,A_{-i}}\big)\big] \leqslant 0. \tag{11}$$

In fact, given both random variables $P_{\mu,m}^{i,w}$ and $f_i(aP_{\mu,m}^{-i,A_{-i}})$ are strictly non-negative, we must have that

$$\lim_{\rho_\mu^m\to 0}\mathbb{E}\big[P_{\mu,m}^{i,w}f_i\big(aP_{\mu,m}^{-i,A_{-i}}\big)\big] = 0.$$

Given the assumption that any $f_i(x) > 0$ for all $x > 0$ and that $\mathbb{E}[P_{\mu,m}^{-i,A_{-i}}] > 0$ this implies that $\lim_{\rho_\mu^m\to 0}\mathbb{E}[P_{\mu,m}^{i,w}P_{\mu,m}^{-i,A_{-i}}] = 0$. □

Appendix D. Lemmas needed to prove Theorem 4

The following result about the variance of $P_{\mu,m}^{i,s}$ when $s \in S_i$ is a strictly dominated strategy is used in the proof of Lemma 5 below, which in turn is used in Lemma 6, which in turn is used in the proof of Lemma 7, which finally is used to prove Theorem 4.

Lemma 4. *Let $s \in S_i$ be strictly dominated and let f_i be such that $f_i(0) = 0$ and $f_i(x) = 1$ for all $x \neq 0$. Then $V(P_{\mu,m}^{i,s}) = \frac{\mu\lambda_s(1-\mu\lambda_s)}{2m_i}$.*

Proof. From Eq. (6) we obtain

$$V\big[\big(P_{\mu,m}^{i,s}\big)_{t+1}\big] = V\big[\big(P_{\mu,m}^{i,s}\big)_t\big] + 2\,\mathrm{Cov}\big[\big(P_{\mu,m}^{i,s}\big)_t, dP_{\mu,m}^{i,s}\big] + V\big[dP_{\mu,m}^{i,s}\big].$$

As we assume that at time t behavior is governed by the stationary invariant distribution, we then have that

$$2\,\mathrm{Cov}\big[\big(P_{\mu,m}^{i,s}\big)_t, dP_{\mu,m}^{i,s}\big] + V\big[dP_{\mu,m}^{i,s}\big] = 0. \tag{12}$$

By definition

$$\mathrm{Cov}\big[\big(P^{i,s}_{\mu,m}\big)_t, dP^{i,s}_{\mu,m}\big] = \mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)_t\, dP^{i,s}_{\mu,m}\big] - \mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)_t\big]\mathbb{E}\big[dP^{i,s}_{\mu,m}\big].$$

Given the assumption that time t behavior is governed by the stationary invariant distribution we have that $\mathbb{E}[dP^{i,s}_{\mu,m}] = 0$. Hence,

$$\mathrm{Cov}\big[\big(P^{i,s}_{\mu,m}\big)_t, dP^{i,s}_{\mu,m}\big] = \mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)_t\, dP^{i,s}_{\mu,m}\big].$$

By the law of iterated expectation we have

$$\mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)_t\, dP^{i,s}_{\mu,m}\big] = \mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)_t \mathbb{E}\big[dP^{i,s}_{\mu,m}\big|\big(P^{i,s}_{\mu,m}\big)_t\big]\big].$$

Recall the argument given in the proof of Theorem 1 that $dP^{i,s}_{\mu,m}$ conditional on $(P^{i,s}_{\mu,m})_t = \frac{k}{m_i}$ can be written as the difference between two random variables $\frac{Y}{m_i}$ and $\frac{X}{m_i}$ (given there). Under the additional assumption about f_i this yields the result that $\mathbb{E}[dP^{i,s}_{\mu,m}|(P^{i,s}_{\mu,m})_t = \frac{k}{m_i}] = \mu\lambda_s - \frac{k}{m_i}$ and, hence $\mathbb{E}[(P^{i,s}_{\mu,m})_t \mathbb{E}[dP^{i,s}_{\mu,m}|(P^{i,s}_{\mu,m})_t]] = \mu\lambda_s \mathbb{E}[(P^{i,s}_{\mu,m})_t] - \mathbb{E}[(P^{i,s}_{\mu,m})^2_t]$. By Theorem 1 and the given assumption about f_i we have that $\mathbb{E}[(P^{i,s}_{\mu,m})_t] = \mu\lambda_s$, and, thus, finally

$$\mathrm{Cov}\big[\big(P^{i,s}_{\mu,m}\big)_t, dP^{i,s}_{\mu,m}\big] = \mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)_t\big]^2 - \mathbb{E}\big[\big(P^{i,s}_{\mu,m}\big)^2_t\big] = -V\big[\big(P^{i,s}_{\mu,m}\big)_t\big].$$

Turning to the second term in Eq. (12) note that

$$V\big[dP^{i,s}_{\mu,m}\big] = \mathbb{E}\big[V\big[dP^{i,s}_{\mu,m}\big|\big(P^{i,s}_{\mu,m}\big)_t\big]\big],$$

again, by the law of iterated expectation. Recall again that $dP^{i,s}_{\mu,m}$ conditional on $(P^{i,s}_{\mu,m})_t = \frac{k}{m_i}$ can be written as the difference between two random variables $\frac{Y}{m_i}$ and $\frac{X}{m_i}$ as given in the proof of Theorem 1. These are independent of each other, conditional on $(P^{i,s}_{\mu,m})_t = \frac{k}{m_i}$, and hence, the variance of their difference is the sum of their variances. Given the fact that Y is a binomial random variable, the variance of $\frac{Y}{m_i}$ is given by $\frac{k}{m_i^2}(1-\mu\lambda_s)\mu\lambda_s$. Similarly, the variance of $\frac{X}{m_i}$ is given by $\frac{m_i-k}{m_i^2}(1-\mu\lambda_s)\mu\lambda_s$. The sum of the two variances is then given by $\frac{1}{m_i}(1-\mu\lambda_s)\mu\lambda_s$ regardless of the value of k. This then finally yields that

$$V\big[dP^{i,s}_{\mu,m}\big] = \frac{1}{m_i}(1-\mu\lambda_s)\mu\lambda_s. \tag{13}$$

Using both intermediate results in Eq. (12) we obtain the desired result. □

The next lemma is used in the proof of Lemma 6, which in turn is used in the proof of Lemma 7, which in turn is used in the proof of Theorem 4.

Lemma 5. *Let $s \in S_i$ be strictly dominated and let f_i be such that $f_i(0) = 0$ and $f_i(x) = 1$ for all $x \neq 0$. Then it is true that*

$$\pi^m_\mu\left(P^{i,s}_{\mu,m} \leqslant \frac{\mu\lambda_s}{2}\right) \leqslant \frac{2}{\mu\lambda_s m_i}.$$

Proof. The proof uses Chebyshev's inequality and Theorem 1 and Lemma 4. A version of Chebyshev's inequality can be given as follows. If X is a random variable with mean $\mathbb{E}$ and variance V, then $P(|X-\mathbb{E}| \geqslant k) \leqslant \frac{V}{k^2}$. Here, $X = P^{i,s}_{\mu,m}$ with mean $\mu\lambda_s$ (given Theorem 1 and the assumption about f_i) and variance $V = \frac{\mu\lambda_s(1-\mu\lambda_s)}{2m_i}$ (by Lemma 4). And $P = \pi^m_\mu$. Then

$$\begin{aligned}
\pi^m_\mu\left(P^{i,s}_{\mu,m} \leqslant \frac{\mu\lambda_s}{2}\right) &= \pi^m_\mu\left(P^{i,s}_{\mu,m} - \mu\lambda_s \leqslant -\frac{\mu\lambda_s}{2}\right)\\
&\leqslant \pi^m_\mu\left(\big|P^{i,s}_{\mu,m} - \mu\lambda_s\big| \geqslant \frac{\mu\lambda_s}{2}\right)\\
&\leqslant 4\frac{\mu\lambda_s(1-\mu\lambda_s)}{2m_i}\frac{1}{\mu^2\lambda_s^2}\\
&= \frac{2(1-\mu\lambda_s)}{\mu\lambda_s m_i}\\
&\leqslant \frac{2}{\mu\lambda_s m_i},
\end{aligned}$$

where Chebyshev's inequality is used between the third and the fourth line. □

Lemma 5 together with Lemma 3 implies the following lemma, which is used in the proof of Lemma 7, which in turn is used in the proof of Theorem 4.

Lemma 6. *Let $s \in S_i$ be strictly dominated and let f_i be any learning function. Then it is true that*

$$\pi_\mu^m\left(P_{\mu,m}^{i,s} \leqslant \frac{\mu\lambda_s}{2}\right) \leqslant \frac{2}{\mu\lambda_s m_i}.$$

Proof. Immediate from Lemmas 6 and 3. □

The following lemma is used in the proof of Theorem 4.

Lemma 7. *Let $w \in S_i$ be a weakly dominated strategy for player i. Let f be arbitrary learning functions. Let $s_{-i} \in A_{-i}(w)$ be such that for every $j \neq i$ player j's component of s_{-i}, s_j, is strictly dominated for player j. Then, for $\beta \in \mathbb{R}$,*

$$\mathbb{E}\left[P_{\mu,m}^{i,w}\left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}\right)^\beta\right] \geqslant \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^\beta\left(\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \sum_{j\neq i}\frac{2}{\mu\lambda_{s_j}m_j}\right).$$

Proof. Let $\mathbf{1}_{(\cdot)}$ denote the indicator function, equal to 1 when the expression in the subscript $(\cdot)$ is true and zero otherwise. Then

$$\begin{aligned}
\mathbb{E}\left[P_{\mu,m}^{i,w}\left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}\right)^\beta\right] &\geqslant \mathbb{E}\left[P_{\mu,m}^{i,w}\left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}\right)^\beta \prod_{j\neq i}\mathbf{1}_{(P_{\mu,m}^{j,s_j}\geqslant\frac{\mu\lambda_{s_j}}{2})}\right] \\
&\geqslant \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^\beta \mathbb{E}\left[P_{\mu,m}^{i,w}\prod_{j\neq i}\mathbf{1}_{(P_{\mu,m}^{j,s_j}\geqslant\frac{\mu\lambda_{s_j}}{2})}\right] \\
&\geqslant \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^\beta \left\{\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \mathbb{E}\left[P_{\mu,m}^{i,w}\left(1-\prod_{j\neq i}\mathbf{1}_{(P_{\mu,m}^{j,s_j}\geqslant\frac{\mu\lambda_{s_j}}{2})}\right)\right]\right\} \\
&\geqslant \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^\beta \left\{\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \mathbb{E}\left[\left(1-\prod_{j\neq i}\mathbf{1}_{(P_{\mu,m}^{j,s_j}\geqslant\frac{\mu\lambda_{s_j}}{2})}\right)\right]\right\} \\
&\geqslant \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^\beta \left\{\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \sum_{j\neq i}\pi_\mu^m\left(P_{\mu,m}^{j,s_j}\leqslant\frac{\mu\lambda_{s_j}}{2}\right)\right\},
\end{aligned}$$

which, given Lemma 6, yields the result. □

Appendix E. Proof of Theorem 4

Eq. (9), in the proof of Theorem 3, still applies here. That is,

$$0 = \mathbb{E}\left[dP_{\mu,m}^{i,w}\right] = \pi_\mu^m\left(B^w\right)\mathbb{E}\left[dP_{\mu,m}^{i,w}\middle|B^w\right] + \left(1-\pi_\mu^m\left(B^w\right)\right)\mathbb{E}\left[dP_{\mu,m}^{i,w}\middle|B^{w,c}\right],$$

where the notation is the same as in the proof of Theorem 3. By Lemma 2, point (2), and the fact that $B^w \subset \Psi$ (defined in Corollary 1) we have that $\pi_\mu^m(B^w) \leqslant c(1-\tau\mu)^{m_i}$ for some constant $c > 0$ and some $\tau \in (0,1)$. By the fact that $dP_{\mu,m}^{i,w} \in [-1,1]$ we then have that $\pi_\mu^m(B^w)\mathbb{E}[dP_{\mu,m}^{i,w}|B^w] \geqslant -c(1-\tau\mu)^{m_i}$. Hence,

$$0 \geqslant -c(1-\tau\mu)^{m_i} + \left(1-c(1-\tau\mu)^{m_i}\right)\mathbb{E}\left[dP_{\mu,m}^{i,w}\middle|B^{w,c}\right]. \tag{14}$$

Again, as in the proof of Theorem 3, by definition $\sigma(w,\omega) = f_i(u_i^*(\omega) - u_i(w,\omega))$. Given the definition of $A_{-i}(w)$ we have that $\min_{s_{-i}\in S_{-i}} u_i(x,s_{-i}) - u_i(w,s_{-i}) = a > 0$. Using inequality (4), and the fact that $f_i(x) = \alpha x^\beta$, we obtain that $\sigma(w,\omega) \geqslant \alpha a^\beta(\prod_{j\neq i} P_{\mu,m}^{j,s_j}(\omega))^\beta$, where $s_j \in S_j$ is player j's part of some $s_{-i} \in A_{-i}$. Similarly to inequality (10), we here obtain

$$\begin{aligned}
\mathbb{E}\left[dP_{\mu,m}^{i,w}\middle|B^{w,c}\right] &= \mathbb{E}\left[\mathbb{E}\left[dP_{\mu,m}^{i,w}\middle|B^{w,c}\wedge\left(P_{\mu,m}^{i,w}\right)_t\wedge\left(P_{\mu,m}^{-i,A_{-i}}\right)_t\right]\right] \\
&\geqslant \mathbb{E}\left[P_{\mu,m}^{i,w}\left(\alpha a^\beta\left(\prod_{j\neq i}P_{\mu,m}^{j,s_j}\right)^\beta(1-\mu)+\mu\right)-\mu\lambda_w\right]
\end{aligned}$$

$$= \alpha a^{\beta}(1-\mu)\mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i} P^{j,s_j}_{\mu,m}\right)^{\beta}\right] + \mu\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mu\lambda_w. \tag{15}$$

Using this in inequality (14) we obtain

$$c(1-\tau\mu)^{m_i} \geqslant \left(1-c(1-\tau\mu)^{m_i}\right)\alpha a^{\beta}(1-\mu)\mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i} P^{j,s_j}_{\mu,m}\right)^{\beta}\right] + \left(1-c(1-\tau\mu)^{m_i}\right)\left(\mu\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mu\lambda_w\right).$$

Now using Lemma 7 we obtain

$$\begin{aligned} c(1-\tau\mu)^{m_i} \geqslant{}& \left(1-c(1-\tau\mu)^{m_i}\right)\alpha a^{\beta}(1-\mu)\left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\mathbb{E}\left[P^{i,w}_{\mu,m}\right] \\ &- \left(1-c(1-\tau\mu)^{m_i}\right)\alpha a^{\beta}(1-\mu)\left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\sum_{j\neq i}\frac{2}{\mu\lambda_{s_j}m_j} \\ &+ \left(1-c(1-\tau\mu)^{m_i}\right)\left(\mu\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mu\lambda_w\right). \end{aligned}$$

Rearranging and letting $d = \alpha a^{\beta}(\prod_{j\neq i}\frac{\lambda_{s_j}}{2})^{\beta} > 0$, we obtain that

$$\mathbb{E}\left[P^{i,w}_{\mu,m}\right] \leqslant \frac{c(1-\tau\mu)^{m_i} + (1-c(1-\tau\mu)^{m_i})\left(d(1-\mu)\mu^{(n-1)\beta}\sum_{j\neq i}\frac{2}{\mu\lambda_{s_j}m_j} + \mu\lambda_w\right)}{d(1-\mu)\mu^{(n-1)\beta} + \mu},$$

or alternatively

$$\frac{\frac{c(1-\tau\mu)^{m_i}}{\mu^{(n-1)\beta}} + (1-c(1-\tau\mu)^{m_i})\left(d(1-\mu)\sum_{j\neq i}\frac{2}{\mu\lambda_{s_j}m_j} + \frac{\mu\lambda_w}{\mu^{(n-1)\beta}}\right)}{d(1-\mu) + \frac{\mu}{\mu^{(n-1)\beta}}}.$$

Now as ρ^m_μ tends to 0, and under the assumption that $\beta < \frac{1}{n-1}$, the denominator tends to d, while the numerator tends to 0. To see that the numerator indeed tends to 0, note that under this limit, $(1-\tau\mu)^{m_i}$ tends to 0 at a faster rate than $\mu^{(n-1)\beta}$, $(1-c(1-\tau\mu)^{m_i})$ tends to 1, and both $\sum_{j\neq i}\frac{2}{\mu\lambda_{s_j}m_j}$ as well as $\frac{\mu\lambda_w}{\mu^{(n-1)\beta}}$ tend to zero (the last because $(n-1)\beta < 1$). □

Appendix F. Proof of Theorem 5

Let $\Gamma = (N, S, u)$ be such that $|N| = n$, $S_j = \{A_j, B_j\}$ for all $j \in N$, $u_i(A_i, s_{-i}) = 1$ for all $s_{-i} \in S_{-i}$, $u_i(B_i, B_{-i}) = 0$ where B_{-i} is the strategy combination where each player $j \neq i$ plays B_j, $u_i(B_i, s_{-i}) = 1$ for all $s_{-i} \neq B_{-i}$, $u_j(A_j, s_{-j}) = 1$ for all $s_{-j} \in S_{-j}$ for all $j \neq i$, $u_j(B_j, s_{-j}) = 0$ for all $s_{-j} \in S_{-j}$ for all $j \neq i$, $f_j(0) = 0$ for all $j \neq i$, and $f_j(x) = 1$ for all $x > 0$ for all $j \neq i$. Then $w = B_i$ is weakly dominated by A_i for player i, while B_j is strictly dominated by A_j for all j. We will show that the theorem holds for $\beta = \frac{1}{n-1}$. Given that $x^{\beta} \leqslant x^{\frac{1}{n-1}}$ for all $x \in [0, 1]$ for all $\beta > \frac{1}{n-1}$ the theorem must then also be true for all $\beta > \frac{1}{n-1}$.

Eq. (9) still applies here:

$$0 = \mathbb{E}\left[dP^{i,w}_{\mu,m}\right] = \pi^m_\mu\left(B^w\right)\mathbb{E}\left[dP^{i,w}_{\mu,m}\middle|B^w\right] + \left(1-\pi^m_\mu\left(B^w\right)\right)\mathbb{E}\left[dP^{i,w}_{\mu,m}\middle|B^{w,c}\right],$$

where the notation is as in the proof of Theorem 3. By Lemma 2, point (2), and the fact that $B^w \subset \Psi$ (defined in Corollary 1) we have that $\pi^m_\mu(B^w) \leqslant c(1-\tau\mu)^{m_i}$ for some constant $c > 0$ and some $\tau \in (0, 1)$. By the fact that $dP^{i,w}_{\mu,m} \in [-1, 1]$ we then have that $\pi^m_\mu(B^w)\mathbb{E}[dP^{i,w}_{\mu,m}|B^w] \leqslant c(1-\tau\mu)^{m_i}$. Hence,

$$0 \leqslant c(1-\tau\mu)^{m_i} + \left(1-c(1-\tau\mu)^{m_i}\right)\mathbb{E}\left[dP^{i,w}_{\mu,m}\middle|B^{w,c}\right]. \tag{16}$$

Inequality (15) here holds as an equality,

$$\mathbb{E}\left[dP^{i,w}_{\mu,m}\middle|B^{w,c}\right] = (1-\mu)\mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i} P^{j,B_j}_{\mu,m}\right)^{\beta}\right] + \mu\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mu\lambda_w. \tag{17}$$

As the covariance between $P^{i,w}_{\mu,m}$ and $\prod_{j\neq i} P^{j,B_j}_{\mu,m}$ must be non-positive, we have that

$$\mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i}P^{j,B_j}_{\mu,m}\right)^{\beta}\right] \leqslant \mathbb{E}\left[P^{i,w}_{\mu,m}\right]\mathbb{E}\left[\left(\prod_{j\neq i}P^{j,B_j}_{\mu,m}\right)^{\beta}\right].$$

By the obvious independence of the $P^{j,B_j}_{\mu,m}$ for all $j \neq i$, we have that

$$\mathbb{E}\left[\left(\prod_{j\neq i}P^{j,B_j}_{\mu,m}\right)^{\beta}\right] = \prod_{j\neq i}\mathbb{E}\left[\left(P^{j,B_j}_{\mu,m}\right)^{\beta}\right].$$

Jensen's inequality (given $\beta = \frac{1}{n-1} \leqslant 1$) then implies that

$$\mathbb{E}\left[\left(P^{j,B_j}_{\mu,m}\right)^{\beta}\right] \leqslant \left(\mathbb{E}\left[P^{j,B_j}_{\mu,m}\right]\right)^{\beta}.$$

Given the particular choice of f_j's here, by Theorem 1 we have that $\mathbb{E}[P^{j,B_j}_{\mu,m}] = \mu\lambda_{B_j}$. Putting all this together into inequality (16), we have

$$0 \leqslant c(1-\tau\mu)^{m_i} + \left(1-c(1-\tau\mu)^{m_i}\right)\left((1-\mu)\mathbb{E}\left[P^{i,w}_{\mu,m}\right]\prod_{j\neq i}(\mu\lambda_{B_j})^{\beta} + \mu\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mu\lambda_w\right).$$

Rearranging leads to

$$\mathbb{E}\left[P^{i,w}_{\mu,m}\right] \geqslant \frac{(1-c(1-\tau\mu)^{m_i})\mu\lambda_w - c(1-\tau\mu)^{m_i}}{(1-c(1-\tau\mu)^{m_i})((1-\mu)\mu^{(n-1)\beta}\prod_{j\neq i}(\lambda_{B_j})^{\beta} + \mu)},$$

or equivalently

$$\mathbb{E}\left[P^{i,w}_{\mu,m}\right] \geqslant \frac{(1-c(1-\tau\mu)^{m_i})\lambda_w - c\frac{(1-\tau\mu)^{m_i}}{\mu}}{(1-c(1-\tau\mu)^{m_i})((1-\mu)\frac{\mu^{(n-1)\beta}}{\mu}\prod_{j\neq i}(\lambda_{B_j})^{\beta} + 1)}.$$

Given $\beta = \frac{1}{n-1}$ the right-hand side of the last inequality converges to $\frac{\lambda_w}{\prod_{j\neq i}(\lambda_{B_j})^{\beta}+1} > 0$. □

References

Ben Porath, E., 1997. Rationality, Nash equilibrium and backwards induction in perfect information games. Rev. Econ. Stud. 64, 23–46.
Bernheim, B.D., 1984. Rationalizable strategic behavior. Econometrica 52, 1007–1029.
Börgers, T., 1994. Weak dominance and approximate common knowledge. J. Econ. Theory 64, 265–276.
Brandenburger, A., 1992. Lexicographic probabilities and iterated admissibility. In: Dasgupta, P., Gale, D., Hart, O., Maskin, E. (Eds.), Economic Analysis of Markets and Games. MIT Press, Cambridge.
Dekel, E., Fudenberg, D., 1990. Rational behavior with payoff uncertainty. J. Econ. Theory 52, 243–267.
Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics. Theoret. Population Biol. 38, 219–232.
Gul, F., 1996. Rationality and coherent theories of strategic behavior. J. Econ. Theory 70, 1–31.
Hart, S., 2002. Evolutionary dynamics and backward induction. Games Econ. Behav. 41, 227–264.
Kandori, M., Mailath, G., Rob, R., 1993. Learning, mutation, and long-run equilibria in games. Econometrica 61, 29–56.
Kohlberg, E., Mertens, J.-F., 1986. On the strategic stability of equilibria. Econometrica 54, 1003–1037.
Kuzmics, C.A., 2004. Stochastic evolutionary stability in extensive form games of perfect information. Games Econ. Behav. 83 (2), 321–336.
Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. Nature 246, 15–18.
Myerson, R.B., 1978. Refinements of the Nash equilibrium concept. Int. J. Game Theory 7, 73–80.
Nachbar, J., 1990. "Evolutionary" selection dynamics in games: Convergence and limit properties. Int. J. Game Theory 19, 59–89.
Nöldeke, G., Samuelson, L., 1993. An evolutionary analysis of backward and forward induction. Games Econ. Behav. 5, 425–454.
Pearce, D.G., 1984. Rationalizable strategic behavior and the problem of perfection. Econometrica 52, 1029–1051.
Ritzberger, K., 2002. Foundations of Non-Cooperative Game Theory. Oxford University Press.
Samuelson, L., 1993. Does evolution eliminate dominated strategies? In: Binmore, K., Kirman, A., Tani, P. (Eds.), Frontiers of Game Theory. MIT Press, Cambridge, pp. 213–235.
Samuelson, L., 1994. Stochastic stability in games with alternative best replies. J. Econ. Theory 64, 35–65.
Samuelson, L., Zhang, J., 1992. Evolutionary stability in asymmetric games. J. Econ. Theory 57, 363–391.
Selten, R., 1975. Re-examination of the perfectness concept for equilibrium points in extensive games. Int. J. Game Theory 4, 25–55.
Taylor, P., Jonker, L., 1978. Evolutionary stable strategies and game dynamics. Math. Biosci. 40, 145–156.
van Damme, E.E.C., 1991. Stability and Perfection of Nash Equilibria. Springer-Verlag, Berlin, Heidelberg.
Weibull, J.W., 1995. Evolutionary Game Theory. MIT Press, Cambridge.