On the Elimination of Dominated Strategies in Stochastic Models of Evolution with Large Populations

Christoph Kuzmics

J. L. Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2009
email: c-kuzmics@kellogg.northwestern.edu
phone: (847) 491 2978
fax: (847) 467 1220

January 28, 2007

Abstract

This paper analyzes a stochastic best reply evolutionary model with inertia in normal form games. The long-run behavior of individuals in this model is investigated in the limit where experimentation rates tend to zero, while the expected number of experimenters, and hence also population sizes, tend to infinity. Conditions on the learning-rate which are necessary and sufficient for the evolutionary elimination of weakly dominated strategies are found. The key determinant is found to be the sensitivity of the learning-rate to small payoff differences.

JEL classification: C62, C72, C73
Keywords: learning, experimentation, $S^{\infty}W$-procedure, weak dominance, iterated strict dominance

1 Introduction

The elimination of weakly dominated strategies is 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.

Evolutionary models so far, with the exception of Theorem 5 in Samuelson (1994), do not support the deletion of weakly dominated strategies. This is true for dynamic models such as the replicator dynamics of Taylor and Jonker (1978) as shown in Samuelson (1993), as well as for stochastic models such as that of Kandori, Mailath, and Rob (1993) as shown in Samuelson (1994, Theorem 3). This is somewhat surprising as evolutionary models directly or indirectly allow for random mutation or experimentation, which on the face of it 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, diminish much faster than the weakly dominated strategy and then vanish before the weakly dominated strategy does (see e.g. Example 3.4 in Weibull, 1995).

In a stochastic finite-population model a la Kandori, Mailath, and Rob (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 three conditions are satisfied. First, a given weakly dominated strategy is not played by anyone in the relevant player population. Second, opponents' strategies, against which the weakly dominated strategy performs worse than the strategy it is dominated by, are not present either. Third, the strategy which dominates the said weakly dominated strategy is a best reply in the given state. But then 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.

In this paper I take up Samuelson's (1994) model of stochastic evolution based on a best-reply dynamic with or without inertia and with random small probability experimentation. From the beginning I will make every individual's learning-rate dependent on the payoff difference between the strategy the individual is currently playing and a best reply. This assumption is immaterial for the finite population case, but becomes important, as Samuelson (1994) notes, when populations sizes tend to infinity. While Samuelson (1994) does not make this dependence explicit, he does discuss it in the aftermath of his Theorem 5.

Samuelson (1994) in Theorem 5 shows, for a particular 2-player game, that stochastically modelled evolution can eliminate weakly dominated strategies if population sizes tend to infinity. The proof of this theorem, based on tree arguments, introduced to Game Theory by Foster and Young (1990), Young (1993), and Kandori, Mailath, and Rob (1993) does extend to some degree to cover some other finite normal form games, see Samuelson's (1994) Footnote 14, but does not extend to all finite normal form games. In the discussion of his Theorem 5, Samuelson (1994) concludes that the assumptions needed to guarantee the elimination of weakly dominated strategies is that the learning-rate is a discontinuous function of the payoff difference between the weakly dominated strategy and the best-reply, which admittedly is a somewhat radical, if not implausible, assumption given individuals in an evolutionary model are usually somewhat boundedly rational and slow to adapt to change.

In this paper I investigate the same limit Samuelson (1994) does in his Theorem 5, but with the additional, and I believe reasonable, requirement that while the experimentation rate μ tends to zero, and population sizes m_i tend to infinity, their product μm_i, the expected number of experimenters in any given period, tends to infinity as well. This implies, and is in fact necessary to imply, that play, in this limit, will be in the interior of the strategy simplex with probability 1 as shown in Corollary 1. I believe this to be the appropriate limit if we want to have an evolutionary model of trembles. Recall that the presence of trembles also imply that play is in the interior of the simplex with probability 1, albeit in a very different sense.

In any case, under this limit I can avoid the tree arguments and use the properties of the invariant distribution more directly. In Theorems 4 and 5 I provide necessary and sufficient conditions on, what one might call, the sensitivity of the learning-rate to payoff differences, for the evolutionary elimination of weakly dominated strategies. This adds to Theorem 5 of Samuelson (1994) in two ways. First, these theorems hold for all finite normal form games. Second, these theorems show how exactly this required sensitivity to payoff differences depends on the number of players in the normal form game. The latter also demonstrates that Samuelson's (1994) conclusion in the discussion of his Theorem 5, that the learning-rate needs to be a discontinuous function of the pay-off differences, is not warranted.

In fact for 2-player games, while 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 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 infinity slope at a payoff-difference of 0. One interpretation one could give a power function with degree less than 1 is that learning individuals are risk-averse over **payoff-differences**. In 2-player games any small degree of such "risk-aversion" would lead to the evolutionary elimination of all weakly dominated strategies in these games. For games with more than 2 players the required "risk-aversion" would have to be higher. This implies that the more players, or better the more player-positions, there are in the game, the harder it is for evolution to eliminate all weakly dominated strategies. All this follows from Theorems 4 and 5. The proofs of Theorems 4 and 5 also suggest a taxonomy or at least a partial order of weakly dominated strategies with respect to the ease with which evolution eliminates them. This is discussed in Section 4.

The structure of this paper is as follows. Section 2 states the model. All of the main results are then presented in Section 3, which after some preliminary lemmas first proves the elimination of strictly dominated strategies in Theorems 1 and 2 before providing the main results on the evolutionary elimination of weakly dominated strategies, Theorems 3, 4, and 5. Finally Section 4 provides a discussion of these results as well as Theorem 7 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. I.e. Theorem 7 provides some evolutionary support for the so-called $S^\infty W$-procedure of Dekel and Fudenberg (1990).

2 Model

For finite populations sizes (m_i, see below) the following model is essentially the same as the stochastic best-reply model with (or without) inertia of Samuelson (1994) and a special case of the evolutionary model of Kandori, Mailath, and Rob (1993). The only difference to the model of Samuelson (1994) is that I will assume from the beginning that any individual's learning rate depends directly 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.

Let $\Gamma(N, S, u)$ be a normal form game, where $N = \{1, ..., n\}$ is the set of n players, $S = \times_{i\in N} S_i$ is the set of pure strategy profiles (S_i is player i's set of pure strategies) and u is the payoff function.

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^m.

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 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. Let $u_i^*(\omega) = \max_{s_i \in S_i} u_i(s_i, \omega)$, where $u_i(s_i, \omega)$ is the payoff strategy s_i yields given the state is ω. Then the probability that this agent (currently playing s) switches to a best-reply given state ω is given by $\sigma(s, \omega) = f_i\left(u_i^*(\omega) - u_i(s_i, \omega)\right)$, where f_i is some function from the non-negative part of the real line into the unit interval. I will assume that, while all f_i's, for different $i \in N$, can be different, they all satisfy $f_i(0) = 0$, $f_i(x) > 0$ 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.

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 a fixed probability distribution with full support over all best replies. If the agent already plays a best reply she is assumed to continue playing it. 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 a (conditional) probability distribution $\lambda_i \in \text{int}[\Delta(S_i)]$, where generally $\Delta(D)$ denotes the set of all probability distributions over D, while $\text{int}[\Delta(D)]$ signifies that this distribution has full support. Hence, all strategies available to this agent (including the one she is playing at the moment) are possible realizations for an experimentation-draw. Let $\lambda = \{\lambda_1, ..., \lambda_n\} \in \times_{i=1}^n \text{int}[\Delta(S_i)]$ be the profile of these conditional mutation probabilities. In the absence of an experimentation-draw the agent does not change her strategy.

The above "mutation-selection" mechanism constitutes a Markov chain on the state space Ω^m with transition probability matrix denoted by Q_μ^m, indicating that it is different for different population sizes and different experimentation rates. The transition probabilities also vary with different choices of f. However, as I will study the limit of this process for any fixed $f = (f_1, ..., f_n)$ but taking μ to zero and m_i to infinity, I suppress the f 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 Results

In this paper I am interested in the long-run behavior of individuals for the limiting model with infinite population sizes. I first discuss, by means of a few lemmas, why I choose to study the limit in which, while population sizes tend to infinity, the experimentation rate tends to zero in such a way that their product, the expected number of experimenters (in any given period), tends to infinity as well. This not only makes proofs easier, I also think this is the most interesting case, as it represents, in my opinion, the evolutionary equivalent of the idea of trembling-hand perfection. Using these results I will then move on to prove the main theorems. First, I need some additional notation.

3.1 More Notation

Let $i \in N$ be an arbitrary player and let $s \in S_i$ be an arbitrary strategy available to individuals at population $M(i)$. 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 $\Phi_\tau^{i,s} = \bigcup_{k \leq \tau m_i} \Lambda_k^{i,s}$ denote the set of states in which not more than a proportion of τ individuals play s at player population $M(i)$. Let $P_{\mu,m}^{i,s} : \Omega^m \to \mathrm{I\!R}$ denote a random variable (given probability space (Ω^m, π_μ^m)) such that $P_{\mu,m}^{i,s}(\omega)$ denotes the proportion of s-players in population $M(i)$ given state ω. Note that $\pi_\mu^m \left(P_{\mu,m}^{i,s} \leq \epsilon \right) = \pi_\mu^m \left(\Phi_\epsilon^{i,s} \right)$. Throughout this section the conditional mutation-probability vector, $\lambda \in \times_{i=1}^n \mathrm{int}[\Delta(S_i)]$ is arbitrary. Hence, the results hold for any such λ. Then λ_s shall denote the probability λ puts on pure strategy $s \in S_i$.

3.2 Preliminary Results

Consider the most adversarial environment for a strategy to survive evolution. 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 $x \neq 0$. In this environment learning is as quick as it can be, given the assumptions of the model. Hence, evolutionary pressure against the strictly dominated strategy s is as severe as it can possibly be.

Lemma 1 *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 $\pi_\mu^m \left(\Lambda_0^{i,s}\right) = (1 - \lambda_s \mu)^{m_i}$.*

Proof: From the fact that π_μ^m is the invariant distribution it follows that

$$\pi_\mu^m \left(\Lambda_0^{i,s}\right) = \sum_{k=0}^{m_i} \pi_\mu^m \left(\Lambda_k^{i,s}\right) \left(Q_\mu^m\right)_{k0},$$

where $\left(Q_\mu^m\right)_{k0}$ is the transition probability of switching to any given state in $\Lambda_k^{i,s}$ (the same for all such states) to a state in $\Lambda_0^{i,s}$. This transition probability, given the assumptions of the lemma, is given by $\left(Q_\mu^m\right)_{k0} = (1 - \lambda_s \mu)^{m_i}$, the same for any k. This is due to the fact that under the severe learning assumption $f_i(x) = 1$ for all $x \neq 0$ and the fact that the payoff difference between s and the optimal strategy is strictly positive, in the learning phase every individual in population i who plays s will switch strategy away from s. To then stay in the set $\Lambda_0^{i,s}$ we need that no-one switches to s in the experimentation phase, the probability of which is given by $(1 - \lambda_s \mu)^{m_i}$. Plugging $\left(Q_\mu^m\right)_{k0} = (1 - \lambda_s \mu)^{m_i}$ into the above equation yields the result. QED

The limit I am investigating in this paper is the one in which population sizes m_i tend to infinity and the experimentation rate tends to zero such that their product tends to infinity as well. One definitely thinks of experimentation rates (or mutation rates in biological evolution) as small. This is the reason why Kandori, Mailath, and Rob (1993) and also Young (1993) investigate the limit in which μ tends to zero. Given that, however, the limit I am considering is the only limit in which we can guarantee for any finite normal form game and for any choice of learning functions $f = (f_1, ..., f_n)$, 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 $\lim (1 - \lambda_s \mu)^{m_i} = \lim \left(1 - \lambda_s \frac{\delta}{m_i}\right)^{m_i} = e^{-\lambda_s \delta} > 0$. Hence, under this limit, we cannot guarantee that a strictly dominated strategy s is always played by at least 1 person. Now, if we recall the idea behind introducing trembles to the "rational" formulation of a normal form game, it is exactly to guarantee that every pure strategy is used with at least some probability. I believe that the appropriate evolutionary counterpart of this is that the event that every strategy is played by at least 1 individual has limiting probability 1. Only then, do we have probability 1 in both cases that play is in the strict interior of the strategy simplex.

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

Lemma 2 *Let $i \in \{1, ..., n\}$ be an arbitrary player and $s \in S_i$ an arbitrary strategy available to individuals at population $M(i)$.*

$$\lim_{\rho_\mu^m \to 0} \pi_\mu^m \left(\Lambda_0^{i,s}\right) = 0. \tag{2}$$

Proof: $\pi_\mu^m \left(\Lambda_0^{i,s}\right) \leq (1 - \lambda_s \mu)^{m_i}$ by the way we chose s and f_i in Lemma 1. Hence, $\lim_{\rho_\mu^m \to 0} \pi_\mu^m \left(\Lambda_0^{i,s}\right) \leq \lim_{\rho_\mu^m \to 0} (1 - \lambda_s \mu)^{m_i} = 0.$ QED

The following corollary is immediate from Lemma 2.

Corollary 1 *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.*

$$\Psi = \bigcup_{i=1}^{n} \bigcup_{x \in S_i} \Lambda_0^{i,x}. \tag{3}$$

Then

$$\lim_{\rho_\mu^m \to 0} \pi_\mu^m (\Psi) = 0. \tag{4}$$

3.3 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 (or 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. I.e. $u_i^*(\omega) - u_i(s_i, \omega) > 0$. In fact we must have that $\min_{\omega \in \Omega} \left(u_i^*(\omega) - u_i(s_i, \omega)\right) = a > 0$. But then under the assumptions about f_i we must have that there is a $\tilde{\sigma}$ such that $\sigma(s, \omega) = f_i \left(u_i^*(\omega) - u_i(s_i, \omega)\right) \geq \tilde{\sigma}$ for all $\omega \in \Omega$. On the other hand, we, of course, have that $\sigma(s, \omega) \leq 1$ 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.

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

$$\mu\lambda_s \leq \mathbb{E}\left[P^{i,s}_{\mu,m}\right] \leq \frac{\mu\lambda_s}{\tilde{\sigma}(1-\mu)+\mu}.$$

Proof: Let $\{\Omega \times \Omega, \Pr\}^{\dagger}$ denote a probability space, where Pr is such that‡ $\Pr(\omega, \omega') = \pi^m_\mu(\omega)\left(Q^m_\mu\right)_{\omega,\omega'}$ for all $(\omega, \omega') \in \Omega \times \Omega$.

Let $\left(P^{i,s}_{\mu,m}\right)_t$ 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 $t+1$. I.e.

$$\left(P^{i,s}_{\mu,m}\right)_{t+1} = \left(P^{i,s}_{\mu,m}\right)_t + dP^{i,s}_{\mu,m}. \tag{5}$$

If $\left(P^{i,s}_{\mu,m}\right)_t$ is distributed according to the invariant distribution π^m_μ then so is $\left(P^{i,s}_{\mu,m}\right)_{t+1}$ and, hence, the expected value $\mathbb{E}\left[dP^{i,s}_{\mu,m}\right] = 0$. 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 $\mathbb{E}\left[dP^{i,s}_{\mu,m}\right] = \mathbb{E}\left[\mathbb{E}\, dP^{i,s}_{\mu,m}\Big|\left(P^{i,s}_{\mu,m}\right)\right]$, and hence

$$0 = \mathbb{E}\left[dP^{i,s}_{\mu,m}\right] = \sum_{k=0}^{m_i} \pi^m_\mu\left(\Lambda^{i,s}_k\right)\mathbb{E}\left(dP^{i,s}_{\mu,m}\Big|\left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}\right). \tag{6}$$

Conditional on $\left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}$, 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 $\left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}$, X and Y are binomially distributed, i.e. $X \sim \text{Bin}\left(m_i - k, \mu\lambda_x\right)$ and $Y \sim$

*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))

†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.

‡Given the axioms of a probability measure this is sufficient to uniquely define Pr.

$\operatorname{Bin}(k, \sigma(s,\omega)(1-\mu)+\mu(1-\lambda_x))$. Given that s is a strictly dominated strategy we know that $\tilde{\sigma} \leq \sigma(s,\omega) \leq 1$. 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 below by

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

and above by

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

Plugging the lower bound back into Equation 6 we obtain

$$0 \geq [\tilde{\sigma}(1-\mu)+\mu] \sum_{k=0}^{m_i} \frac{k}{m_i} \pi^m_\mu \left(\Lambda^{i,s}_k\right) - \mu\lambda_s, \tag{7}$$

which by the assumptions of the lemma and by the fact that $\sum_{k=0}^{m_i} \frac{k}{m_i} \pi^m_\mu \left(\Lambda^{i,s}_k\right) = \mathbb{E}\left(P^{i,s}_{\mu,m}\right)$ yields $\mathbb{E}\left(P^{i,s}_{\mu,m}\right) \leq \frac{\mu\lambda_s}{\tilde{\sigma}(1-\mu)+\mu}$. Doing same with the upper bound yields $\mathbb{E}\left(P^{i,s}_{\mu,m}\right) \geq \mu\lambda_s$. QED

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. While I do not know whether this has been put on record quite like this, it is clear, from reading e.g. Samuelson (1994) that the evolutionary elimination of strictly dominated strategies in such models is well understood. In any case 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*

$$\lim_{\mu \to 0} \mathbb{E}\left[P^{i,s}_{\mu,m}\right] = 0.$$

Proof: Immediate from Theorem 1. QED

I would like to draw attention to one more observation. Theorem 1 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.

The following result about the variance of $P^{i,s}_{\mu,m}$ when $s \in S_i$ is a strictly dominated strategy will become useful later. As in Lemma 1 I will again consider the most adversarial environment for a strictly dominated strategy to survive evolution. Let again $s \in S_i$ be strictly dominated and let f_i be such that $f_i(0) = 0$ and $f_i(x) = 1$ for all $x \neq 0$.

Lemma 3 *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 $V\left(P^{i,s}_{\mu,m}\right) = \frac{\mu\lambda_s(1-\mu\lambda_s)}{2m_i}$.*

Proof: From equation 5 we obtain

$$V\left[\left(P^{i,s}_{\mu,m}\right)_{t+1}\right] = V\left[\left(P^{i,s}_{\mu,m}\right)_t\right] + 2\mathrm{Cov}\left[\left(P^{i,s}_{\mu,m}\right)_t, dP^{i,s}_{\mu,m}\right] + V\left[dP^{i,s}_{\mu,m}\right].$$

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

$$2\mathrm{Cov}\left[\left(P^{i,s}_{\mu,m}\right)_t, dP^{i,s}_{\mu,m}\right] + V\left[dP^{i,s}_{\mu,m}\right] = 0. \tag{8}$$

By definition

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

By Theorem 1 and the given assumption about f_i we have that $\mathbb{E}\left[\left(P^{i,s}_{\mu,m}\right)_t\right] = \mu\lambda_s$. Given the assumption that time t behavior is governed by the stationary invariant distribution we have that $\mathbb{E}\left[dP^{i,s}_{\mu,m}\right] = 0$. Hence,

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

By the law of iterated expectation we have

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

Recall the argument given in the proof of Theorem 1 that $dP^{i,s}_{\mu,m}$ conditional on $\left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}$ 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 $\mathbb{E}\left[dP^{i,s}_{\mu,m}\middle| \left(P^{i,s}_{\mu,m}\right)_t = \frac{k}{m_i}\right] = \mu\lambda_s - \frac{k}{m_i}$ and, hence

$\mathbb{E}\left[\left(P_{\mu,m}^{i,s}\right)_t \mathbb{E}\left[dP_{\mu,m}^{i,s}\middle| \left(P_{\mu,m}^{i,s}\right)_t\right]\right] = \mu\lambda_s \mathbb{E}\left[\left(P_{\mu,m}^{i,s}\right)_t\right] - \mathbb{E}\left[\left(P_{\mu,m}^{i,s}\right)_t^2\right]$. Given that $\mathbb{E}\left[\left(P_{\mu,m}^{i,s}\right)_t\right] = \mu\lambda_s$ we finally have that

$$\mathrm{Cov}\left[\left(P_{\mu,m}^{i,s}\right)_t, dP_{\mu,m}^{i,s}\right] = \mathbb{E}\left[\left(P_{\mu,m}^{i,s}\right)_t\right]^2 - \mathbb{E}\left[\left(P_{\mu,m}^{i,s}\right)_t^2\right] = -V\left[\left(P_{\mu,m}^{i,s}\right)_t\right].$$

Turning to the second term in equation 8 note that

$$V\left[dP_{\mu,m}^{i,s}\right] = \mathbb{E}\left[V\left[dP_{\mu,m}^{i,s}\middle| \left(P_{\mu,m}^{i,s}\right)_t\right]\right],$$

again, by the law of iterated expectation. Recall again that $dP_{\mu,m}^{i,s}$ conditional on $\left(P_{\mu,m}^{i,s}\right)_t = \frac{k}{m_i}$ 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 $\left(P_{\mu,m}^{i,s}\right)_t = \frac{k}{m_i}$, 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

$$V\left[dP_{\mu,m}^{i,s}\right] = \frac{1}{m_i}(1-\mu\lambda_s)\mu\lambda_s. \tag{9}$$

Using both intermediate results in equation 8 we obtain the desired result. QED

The next lemma will also be useful later.

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 it is true that*

$$\pi_\mu^m\left(P_{\mu,m}^{i,s} \leq \frac{\mu\lambda_s}{2}\right) \leq \frac{4}{\mu\lambda_s m_i}.$$

Proof: This is immediate from Chebyshev's inequality, Theorem 1, and Lemma 3. QED

3.4 Weakly Dominated Strategies

Let $w \in S_i$ be a weakly dominated strategy which is not strictly dominated. We then have that $u_i(w,\omega) \leq u_i^*(\omega)$. Let w be in fact weakly dominated by some mixed strategy $x \in \Delta(S_i)$. We then have that $u_i^*(\omega) - u_i(w,\omega) \geq$

$u_i(x,\omega) - u_i(w,\omega) \geq 0$. Let $S_{-i} = \times_{j\neq i} S_j$. 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

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

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

$$u_i^*(\omega) - u_i(w,\omega) \geq \left(u_i(x, s_{-i}) - u_i(w, s_{-i})\right) P_{\mu,m}^{-i,s_{-i}}(\omega) \tag{10}$$

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 modelled 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. $A_{-i} = \{s_{-i} \in S_{-i} | u_i(x, s_{-i}) > u_i(w, s_{-i})$. 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 $w \in S_i$ be weakly dominated. Suppose $\lim_{\rho_\mu^m\to 0} \mathbb{E} P_{\mu,m}^{-i,A_{-i}} > 0$. Then $\lim_{\rho_\mu^m\to 0} \mathbb{E}\left[P_{\mu,m}^{i,w} P_{\mu,m}^{-i,A_{-i}}\right] = 0$ for any choice of learning function f_i.*

Proof: Reconsider equation 5, now for strategy w, $\left(P_{\mu,m}^{i,w}\right)_{t+1} = \left(P_{\mu,m}^{i,w}\right)_t + dP_{\mu,m}^{i,w}$. 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 $\mathbb{E}\left[dP_{\mu,m}^{i,w}\right]$, which as in the proof of Theorem 1 must be zero, using the law of iterated expectations, can be written as

$$\mathbb{E}\left[dP_{\mu,m}^{i,w}\right] = \pi_\mu^m(B^w)\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^w\right] + (1-\pi_\mu^m(B^w))\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^{w,c}\right], \tag{11}$$

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

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

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 $P^{i,w}_{\mu,m} = \frac{k}{m_i}$ we still have $X \sim \text{Bin}(m_i - k, \mu\lambda_w)$ as well as $Y \sim \text{Bin}(k, \sigma(w,\omega)(1-\mu)+\mu(1-\lambda_w))$. 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 10, and the fact that f_i is weakly increasing, we obtain that $\sigma(w,\omega) \geq f_i\left(aP^{-i,A_{-i}}_{\mu,m}(\omega)\right)$. Putting all this together we obtain

$$\begin{aligned} &\mathbb{E}\left[\mathbb{E}\left[dP^{i,w}_{\mu,m}|B^{w,c} \wedge \left(P^{i,w}_{\mu,m}\right)_t \wedge \left(P^{-i,A_{-i}}_{\mu,m}\right)_t\right]\right] \geq \\ &\quad \mathbb{E}\left[P^{i,w}_{\mu,m}\left(f_i\left(aP^{-i,A_{-i}}_{\mu,m}\right)(1-\mu)+\mu\right) - \mu\lambda_w\right] \end{aligned} \tag{12}$$

By the fact that w is weakly dominated we have that $B^w \subset \Psi$, and, hence, by Corollary 1 we have that $\lim_{\rho^m_\mu \to 0} \pi^m_\mu(B^w) = 0$. Hence, from equation 11 we have that $\lim_{\rho^m_\mu \to 0} \mathbb{E}\left[dP^{i,w}_{\mu,m}|B^{w,c}\right] = 0$. But then, by inequality 12, we have that

$$\lim_{\rho^m_\mu \to 0} \mathbb{E}\left[P^{i,w}_{\mu,m}\left(f_i\left(aP^{-i,A_{-i}}_{\mu,m}\right)(1-\mu)+\mu\right) - \mu\lambda_w\right] \leq 0,$$

which, given the assumption that $\lim_{\rho^m_\mu \to 0} \mathbb{E}P^{-i,A_{-i}}_{\mu,m} > 0$ and, hence, that $\lim_{\rho^m_\mu \to 0} \mathbb{E}f_i\left(aP^{-i,A_{-i}}_{\mu,m}\right) > 0$ implies that

$$\lim_{\rho^m_\mu \to 0} \mathbb{E}\left[P^{i,w}_{\mu,m} f_i\left(aP^{-i,A_{-i}}_{\mu,m}\right)\right] \leq 0. \tag{13}$$

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

$$\lim_{\rho^m_\mu \to 0} \mathbb{E}\left[P^{i,w}_{\mu,m} f_i\left(aP^{-i,A_{-i}}_{\mu,m}\right)\right] = 0.$$

Given the assumption that any $f_i(x) > 0$ for all $x > 0$ and that $\mathbb{E}\left[P^{-i,A_{-i}}_{\mu,m}\right] > 0$ this implies the result. QED

Suppose now that for all $j \neq i$ f_j is as severe as possible, i.e. $f_j(x) = 0$ if $x = 0$ and $f_j(x) = 1$ for all $x > 0$. Suppose furthermore that there is an $s_{-i} \in A_{-i}$ such that for every $j \neq i$ player j's component of s_{-i}, s_j, is strictly dominated for player j.

Lemma 5 *Let $w \in S_i$ be a weakly dominated strategy for player i. For all $j \neq i$ let f_j be such that $f_j(x) = 0$ if $x = 0$ and $f_j(x) = 1$ for all $x > 0$. 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 $\beta \in \mathbb{R}$,

$$\mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i} P^{j,s_j}_{\mu,m}\right)^{\beta}\right] \geq \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\left(\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \sum_{j\neq i}\frac{4}{\mu\lambda_{s_j}m_j}\right).$$

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

$$\mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i} P^{j,s_j}_{\mu,m}\right)^{\beta}\right] \geq \mathbb{E}\left[P^{i,w}_{\mu,m}\left(\prod_{j\neq i} P^{j,s_j}_{\mu,m}\right)^{\beta}\prod_{j\neq i}\mathbf{1}_{\left(P^{j,s_j}_{\mu,m}\geq\frac{\mu\lambda_{s_j}}{2}\right)}\right] \tag{14}$$

$$\geq \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\mathbb{E}\left[P^{i,w}_{\mu,m}\prod_{j\neq i}\mathbf{1}_{\left(P^{j,s_j}_{\mu,m}\geq\frac{\mu\lambda_{s_j}}{2}\right)}\right] \tag{15}$$

$$\geq \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\left\{\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mathbb{E}\left[P^{i,w}_{\mu,m}\left(1-\prod_{j\neq i}\mathbf{1}_{\left(P^{j,s_j}_{\mu,m}\geq\frac{\mu\lambda_{s_j}}{2}\right)}\right)\right]\right\} \tag{16}$$

$$\geq \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\left\{\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \mathbb{E}\left[\left(1-\prod_{j\neq i}\mathbf{1}_{\left(P^{j,s_j}_{\mu,m}\geq\frac{\mu\lambda_{s_j}}{2}\right)}\right)\right]\right\} \tag{17}$$

$$\geq \left(\prod_{j\neq i}\frac{\mu\lambda_{s_j}}{2}\right)^{\beta}\left\{\mathbb{E}\left[P^{i,w}_{\mu,m}\right] - \sum_{j\neq i}\pi^m_\mu\left(P^{j,s_j}_{\mu,m}\leq\frac{\mu\lambda_{s_j}}{2}\right)\right\}, \tag{18}$$

which, given Lemma 4, yields the result. QED

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 $\beta < \frac{1}{n-1}$ then $\lim_{\rho^m_\mu\to 0}\mathbb{E}\left[P^{i,w}_{\mu,m}\right] = 0$.*

Proof: Equation 11, in the proof of Theorem 3, still applies here. I.e.

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

where the notation is the same as in the proof of Theorem 3. By Lemma 1, and the fact that $B^w \subset \Psi$ (defined in Corollary 1) we have that $\pi^m_\mu(B^w) \leq c(1-\tau\mu)^{m_i}$ 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 $\pi^m_\mu(B^w)\mathbb{E}\left[dP^{i,w}_{\mu,m}|B^w\right] \geq -c(1-\tau\mu)^{m_i}$. Hence,

$$0 \geq -c(1-\tau\mu)^{m_i} + (1-c(1-\tau\mu)^{m_i})\mathbb{E}\left[dP^{i,w}_{\mu,m}|B^{w,c}\right]. \tag{19}$$

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 10, and the fact that $f_i(x) = \alpha x^\beta$, we obtain that $\sigma(w,\omega) \geq \alpha a^\beta \left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}(\omega)\right)^\beta$, where $s_j \in S_j$ is player j's part of some $s_{-i} \in A_{-i}$. Similarly to inequality 12, we here obtain

$$\begin{aligned}
&\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^{w,c}\right] = \\
&\mathbb{E}\left[\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^{w,c} \wedge \left(P_{\mu,m}^{i,w}\right)_t \wedge \left(P_{\mu,m}^{-i,A_{-i}}\right)_t\right]\right] \geq \\
&\mathbb{E}\left[P_{\mu,m}^{i,w}\left(\alpha a^\beta \left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}\right)^\beta (1-\mu) + \mu\right) - \mu\lambda_w\right] = \\
&\alpha a^\beta (1-\mu)\mathbb{E}\left[P_{\mu,m}^{i,w}\left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}\right)^\beta\right] + \mu\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \mu\lambda_w
\end{aligned} \tag{20}$$

Using this in Inequality 19 we obtain

$$\begin{aligned}
c(1-\tau\mu)^{m_i} \geq{}& (1 - c(1-\tau\mu)^{m_i})\, \alpha a^\beta (1-\mu)\mathbb{E}\left[P_{\mu,m}^{i,w}\left(\prod_{j\neq i} P_{\mu,m}^{j,s_j}\right)^\beta\right] \\
&+ (1 - c(1-\tau\mu)^{m_i})\left(\mu\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \mu\lambda_w\right).
\end{aligned}$$

Now using Lemma 5 we obtain

$$\begin{aligned}
c(1-\tau\mu)^{m_i} \geq{}& (1 - c(1-\tau\mu)^{m_i})\, \alpha a^\beta (1-\mu)\left(\prod_{j\neq i} \frac{\mu\lambda_{s_j}}{2}\right)^\beta \mathbb{E}\left[P_{\mu,m}^{i,w}\right] \\
&- (1 - c(1-\tau\mu)^{m_i})\, \alpha a^\beta (1-\mu)\left(\prod_{j\neq i} \frac{\mu\lambda_{s_j}}{2}\right)^\beta \sum_{j\neq i} \frac{4}{\mu\lambda_{s_j} m_j} \\
&+ (1 - c(1-\tau\mu)^{m_i})\left(\mu\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \mu\lambda_w\right).
\end{aligned}$$

Rearranging and letting $d = \alpha a^\beta \left(\prod_{j\neq i} \frac{\lambda_{s_j}}{2}\right)^\beta > 0$, we obtain that $\mathbb{E}\left[P_{\mu,m}^{i,w}\right] \leq$

$$\frac{c(1-\tau\mu)^{m_i} + (1 - c(1-\tau\mu)^{m_i})\left(d(1-\mu)\mu^{(n-1)\beta}\sum_{j\neq i}\frac{4}{\mu\lambda_{s_j}m_j} + \mu\lambda_w\right)}{d(1-\mu)\mu^{(n-1)\beta} + \mu},$$

or alternatively

$$\frac{\frac{c(1-\tau\mu)^{m_i}}{\mu^{(n-1)\beta}} + (1 - c(1-\tau\mu)^{m_i})\left(d(1-\mu)\sum_{j\neq i}\frac{4}{\mu\lambda_{s_j}m_j} + \frac{\mu\lambda_w}{\mu^{(n-1)\beta}}\right)}{d(1-\mu) + \frac{\mu}{\mu^{(n-1)\beta}}}.$$

Now as ρ_μ^m 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 the last statement, 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 $\sum_{j \neq i} \frac{4}{\mu \lambda_{s_j} m_j}$ as well as $\frac{\mu \lambda_w}{\mu^{(n-1)\beta}}$ tend to zero. QED

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 $f_i(x) = x^\beta$ with $\beta \geq \frac{1}{n-1}$. 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 $\lim_{\rho_\mu^m \to 0} \mathbb{E}\left[P_{\mu,m}^{i,w}\right] > 0$.*

Proof: Let $\Gamma = (N, S, u)$ be such that $|N| = n$, $S_i = \{A_i, B_i\}$ for all $i \in N$, $u_i(A_i, s_{-i}) = 1$ 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 $j \neq i$, $u_j(A_j, s_{-j}) = 1$ 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 the Theorem for $\beta = \frac{1}{n-1}$. Given that $x^\beta \leq x^{\frac{1}{n-1}}$ for all $x \in [0,1]$ for all $\beta > \frac{1}{n-1}$ the Theorem must then clearly also be true for all such $\beta > \frac{1}{n-1}$.

Equation 11 still applies here:

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

where the notation is as in the proof of Theorem 3. By Lemma 1 and the fact that $B^w \subset \Psi$ (defined in Corollary 1) we have that $\pi_\mu^m(B^w) \leq c(1-\tau\mu)^{m_i}$ 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 $\pi_\mu^m(B^w)\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^w\right] \leq c(1-\tau\mu)^{m_i}$. Hence,

$$0 \leq c(1-\tau\mu)^{m_i} + (1 - c(1-\tau\mu)^{m_i})\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^{w,c}\right]. \tag{21}$$

Inequality 20 here holds as an equality,

$$\mathbb{E}\left[dP_{\mu,m}^{i,w}|B^{w,c}\right] = (1-\mu)\mathbb{E}\left[P_{\mu,m}^{i,w}\left(\prod_{j \neq i} P_{\mu,m}^{j,B_j}\right)^\beta\right] + \mu\mathbb{E}\left[P_{\mu,m}^{i,w}\right] - \mu\lambda_w. \tag{22}$$

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

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

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 $\beta = \frac{1}{n-1} \leq 1$) then implies that

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

Given the particular choice of f_j's here, by Theorem 1 we have that $\mathbb{E}\left[P^{j,B_j}_{\mu,m}\right] = \mu\lambda_{B_j}$. Putting all this together into Inequality 21, we have

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

Rearranging leads to

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

or equivalently

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

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

4 Discussion

4.1 The sensitivity to payoff differences

Restricting attention to learning functions of the power form, i.e. $f_i(x) = \alpha x^{\beta}$, for $\beta \in (0,1]$, one could call $\frac{1}{\beta}$ the sensitivity to the payoff-difference between the individual's current strategy and the best option available for an individual with this learning function. The limit when β tends to zero yields the extreme learning function $f_i(0) = 0$ and $f_i(x) = \alpha$ for all $x > 0$. This learning function then has an infinity sensitivity to payoff-differences.

Using the tree arguments introduced to Game Theory by Foster and Young (1990), Young (1993), and Kandori, Mailath and Rob (1993), Samuelson (1994) in his Theorem 5 showed that taking the population size to infinity can lead to the elimination of weakly dominated strategies. Samuelson's (1994) Theorem 5 is for a particular 2-player example only. The proof, however, see Samuelson's (1994) footnote 14, does extend to all cases in which the invariant distribution eventually puts limiting probability 1 on any small neighborhood of a single state. The proof technique based on these tree arguments, however, does not extend beyond that, i.e. does not cover all finite normal form games. Under the additional assumption that the expected number of experimenters in each period, μm_i, tends to infinity while μ tends to zero, and, hence, m_i tends to infinity, I was able to avoid the usual tree arguments by using the properties of the invariant distribution more directly. My Theorem 4 then gives sufficient conditions, depending on the number of players n, under which evolution does indeed eliminate all weakly dominated strategies in any finite normal form game with up to n players.

Samuelson's (1994) conclusion, on page 61, that in order for evolution to eliminate weakly dominated strategies the learning rate must be discontinuous, is not quite warranted. Theorem 4 shows that in order to guarantee the evolutionary elimination of all weakly dominated strategies in an arbitrary finite n-player game the individuals' sensitivity to payoff differences in their learning functions must be above a certain threshold, in fact above $n - 1$, but does not need to be infinite.

I find it quite interesting that this threshold depends on the number of players. For 2-player games it means that while (by Theorem 5) if individuals' sensitivity to payoff differences is 1 (i.e. the learning rate is a linear function of the payoff-difference) the evolutionary elimination of all weakly dominated strategies is not guaranteed, any sensitivity greater than 1 would guarantee it. This does mean that the derivative of the learning function with respect to the payoff difference must be infinite, but the learning function need not be discontinuous. While I agree with Samuelson (1994) that a discontinuous learning-function is somewhat counter-intuitive to the idea of somewhat boundedly rational individuals slowly learning to play the game, an infinite derivative at zero I do not find so implausible. In fact, it is as if these individuals have a certain degree of risk-aversion over **payoff differences**, not over payoffs, when choosing to switch strategies or not.

Theorems 4 and 5 indicate that in general weakly dominated strategies are more readily eliminated by evolution the fewer the number of players, or player positions, n. While it may require a high degree of sensitivity to payoff differences to guarantee the elimination of all weakly dominated strategies in an arbitrary finite normal form game with a large number of players n, any given weakly dominated strategy may not require such high sensitivity. Consider the case where a certain weakly dominated strategy w for player i is only worse than the strategy it is dominated by for one specific strategy combination of the other players. Suppose now that every other players's part of this strategy combination is strictly dominant for this player. But then this one strategy combination will be played with limiting probability 1, given Theorem 2. In this case, however, it should be obvious that any degree of sensitivity to payoff differences is sufficient to eliminate the weakly dominated strategy w. In fact, the results in the previous section suggest a taxonomy or at least a partial order of weakly dominated strategies with respect to the sensitivity required for evolution to eliminate them. Of all strategy combinations against which a certain weakly dominated strategy is not a best reply there must be one which is most prevalent (in expectation) among them. Suppose now that this strategy combination is such that for n_1 of the $n-1$ other player positions their component of this strategy is **strictly dominant**. Suppose further that for another n_2 of the $n-1$ other player positions their component of this strategy is **strictly dominated**. Of course $n_1 + n_2 \leq n - 1$. The required sensitivity to payoff differences needed for evolution to eliminate this weakly dominated strategy, according to the arguments made in the proofs of Theorems 4 and 5 must then lie between n_2 and $n - 1 - n_1$.

4.2 $\mathbf{S^{\infty}W}$

I now turn to a brief discussion about which other strategies will have to be eliminated by evolution as modelled 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 $\mathbb{E}\left[P^{i,w}_{\mu,m}\right]$ tend to zero as ρ^m_μ tends to 0. Lemma 6, given in the appendix, shows that whenever $\lim_{\rho^m_\mu \to 0} \mathbb{E}\left[P^{i,w}_{\mu,m}\right] = 0$ then it must be true that, for any $\epsilon \in (0,1)$ we have $\pi^m_\mu\left(P^{i,x}_{\mu,m} \leq \epsilon\right) = \pi^m_\mu\left(\Phi^{i,x}_\epsilon\right)$ 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. I.e. 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. The next theorem states that, 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.

Theorem 6 *For $i \in N$ let $s \in S_i^1$ be a strategy which is strictly dominated*[§] *in Γ^1. Whenever $\lim_{\rho_\mu^m \to 0} \mathbb{E}\left(P_{\mu,m}^{j,w}\right) = 0$ for every weakly dominated strategy $w \in S_j$ for every player j, then*

$$\lim_{\rho_\mu^m \to 0} \mathbb{E}\left(P_{\mu,m}^{i,s}\right) = 0. \tag{23}$$

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. QED

In fact, the above 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 7 *For $i \in N$, let $d_i \in S_i$, be a strategy which is not strongly rationalizable. Whenever $\lim_{\rho_\mu^m \to 0} \mathbb{E}\left(P_{\mu,m}^{j,w}\right) = 0$ for every weakly dominated strategy $w \in S_j$ for every player j, then*

$$\lim_{\rho_\mu^m \to 0} \mathbb{E}\left(P_{\mu,m}^{i,d_i}\right) = 0. \tag{24}$$

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. 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.

4.3 Other related papers

The papers by Hart (2002) and more so Kuzmics (2004) are related to this paper in terms of the techniques of proof. Both of these papers deal, however, with a stochastic model of evolution in generic extensive form games

[§] Again, as in Theorem 1, strictly dominated can be replaced be never a best reply.

of perfect information. Hart's (2002) model is perhaps more biologically flavored and is such that at any given point in time only 1 individual in each population can change strategy and will typically do so to a better reply. The limit Hart (2002) considers is one in which the product of experimentation or mutation rate and the population sizes are bounded from below. The interpretation of this product, in Hart (2002), is, however, not quite the same as here, given that in his model only 1 individual in each population changes strategy at any given point in time. The proofs in Hart (2004) are quite different from the ones given here, primarily because in Hart's (2002) limit one cannot rule out that play is on the boundary of the strategy simplex.

The model in Kuzmics (2004) is essentially the same as the finite population model in Nöldeke and Samuelson (1993). Kuzmics (2004) investigates the same limit as in this paper, in which the expected number of experimenters μm_i tends to infinity, while μ tends to zero. The learning-rate in Kuzmics (2004), however, does not depend on the payoff differences. Yet, the proofs of Lemma 1, and Corollaries 1 and 2 in Kuzmics (2004) are similar, but not identical to the proofs of Lemmas 1 and 2, and Corollary 1 in this paper. The combined proofs of Theorems 1 and 2 are reminiscent of the proof of Lemma 4 in Kuzmics (2004). The main Theorems 3, 4 and 5, although of course in a somewhat similar vein, do not have a counterpart to any statement in Kuzmics (2004).

A Lemma 6

Lemma 6 *Let $\{X_t\}_{t\in\mathbb{N}} \in [0,1]$ be a sequence of random variables (defined on a sequence of probability spaces $\mathcal{P}_t$), each of which realizes into the unit interval. Then the following two statements are equivalent.*

1. $\lim_{t\to\infty} \mathbb{E}[X_t] = 0$
2. $\lim \mathit{Prob}(X_t \leq \epsilon) = 1$ *for all* $\epsilon \in (0,1)$.

Proof: First, suppose $\lim_{t\to\infty} \mathbb{E}[X_t] = 0$. Suppose $\lim \mathrm{Prob}(X_t \leq \epsilon) < 1$ for some $\epsilon \in (0,1)$. Then there is an $\epsilon > 0$ and a $\delta > 0$ such that $\mathrm{Prob}(X_t \leq \epsilon) \leq 1-\delta$ for all t greater than some T. But then

$$\begin{aligned}\mathbb{E}[X_t] &= \mathbb{E}[X_t|X_t \leq \epsilon]\cdot \mathrm{Prob}(X_t \leq \epsilon) + \mathbb{E}[X_t|X_t > \epsilon]\cdot \mathrm{Prob}(X_t > \epsilon) \\ &\geq 0\cdot 0 + \delta\cdot\epsilon \end{aligned} \tag{25}$$

for all $t \geq T$ which provides a contradiction.

Second, suppose $\lim \text{Prob}\,(X_t \le \epsilon) = 1$ for all $\epsilon \in (0,1)$. Then

$$\begin{aligned}\lim_{t\to\infty} \mathbb{E}\,[X_t] &= \lim_{t\to\infty} \mathbb{E}\,[X_t|X_t \le \epsilon] \cdot \lim_{t\to\infty} \text{Prob}\,(X_t \le \epsilon) + \\ &+ \lim_{t\to\infty} \mathbb{E}\,[X_t|X_t > \epsilon] \cdot \lim_{t\to\infty} \text{Prob}\,(X_t > \epsilon) \\ &\le 1 \cdot \epsilon + 0 \cdot 1. \end{aligned} \tag{26}$$

As this is true for any $\epsilon \in (0,1)$ we arrive at the wanted result. QED

References

[1] Ben Porath, E. (1997). "Rationality, Nash equilibrium and backwards induction in perfect information games," *Review of Economic Studies*, **64**, 23-46.

[2] Bernheim, B. D. (1984). "Rationalizable strategic behavior", *Econometrica*, **52**, 1007-29.

[3] Börgers, T. (1994). "Weak dominance and approximate common knowledge", *Journal of Economic Theory*, **64**, 265-276.

[4] Brandenburger, A. (1992). "Lexicographic probabilities and iterated admissibility", in *Economic Analysis of Markets and Games*, Eds. Dasgupta, P., Gale, D., Hart, O., and Maskin, E., MIT Press, Cambridge.

[5] Dekel, E. and Fudenberg, D. (1990). "Rational behavior with payoff uncertainty", *Journal of Economic Theory*, **52**, 243-67.

[6] Foster, D., and Young, P. (1990). "Stochastic evolutionary game dynamics," *Theoretical Population Biology*, **38**, 219-32.

[7] Gul, F. (1996). "Rationality and coherent theories of strategic behavior," *Journal of Economic Theory*, **70**, 1-31.

[8] Hart, S. (2002). "Evolutionary dynamics and backward induction," *Games and Economic Behavior* **41**, 227-64.

[9] Kandori, M., Mailath, G., and Rob, R. (1993). "Learning, mutation, and long-run equilibria in games", *Econometrica*, **61**, 29-56.

[10] Kohlberg, E. and Mertens, J.-F. (1986). "On the strategic stability of equilibria", *Econometrica*, **54**, 1003-37.

[11] Kuzmics, C. A. (2004). "Stochastic evolutionary stability in extensive form games of perfect information", *Games and Economic Behavior* **48**, 321-336.

[12] Myerson, R. B. (1978). "Refinements of the Nash equilibrium concept", *International Journal of Game Theory*, **7**, 73-80.

[13] Nöldeke, G. and Samuelson, L. (1993). "An evolutionary analysis of backward and forward induction", *Games and Economic Behavior*, **5**, 425-54.

[14] Pearce, D. G. (1984). "Rationalizable strategic behavior and the problem of perfection", *Econometrica*, **52**, 1029-51.

[15] Ritzberger, K. (2002). *Foundations of Non-Cooperative Game Theory*, Oxford University Press.

[16] Samuelson, L., and Zhang, J. (1992). "Evolutionary stability in asymmetric games", *Journal of Economic Theory*, **57**, 363-91.

[17] Samuelson, L. (1993). "Does evolution eliminate dominated strategies?", in *Frontiers of Game Theory*, Eds. Binmore, K., Kirman, A., Tani, P., MIT Press, Cambridge, pp. 213-35.

[18] Samuelson, L. (1994). "Stochastic Stability in Games with Alternative Best Replies", *Journal of Economic Theory*, **64**, 35-65.

[19] Selten, R. (1975). Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* **4**, 25-55.

[20] Taylor, P. and Jonker, L. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, **40**, 145-56.

[21] Weibull, J. W. (1995). *Evolutionary Game Theory*, Cambridge, MIT Press.

[22] Young, H. P. (1993). "The evolution of conventions", *Econometrica*, **61**, 57-84.