ACADEMIC
PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Games and Economic Behavior 41 (2002) 227–264

GAMES and
Economic
Behavior

www.elsevier.com/locate/geb

Evolutionary dynamics and backward induction

Sergiu Hart

Center for Rationality and Interactive Decision Theory, Department of Economics,
Department of Mathematics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel

Received 14 June 2000

Abstract

The backward induction (or subgame-perfect) equilibrium of a perfect information game is shown to be the unique evolutionarily stable outcome for dynamic models consisting of selection and mutation, when the mutation rate is low and the populations are large.

JEL classification: C7; D7; C6

Keywords: Games in extensive form; Games of perfect information; Backward induction equilibrium; Subgame-perfect equilibrium; Evolutionary dynamics; Evolutionary stability; Mutation; Selection; Population games

1. Introduction

1.1. Background

A fascinating meeting of ideas has occurred in the last two decades between evolutionary biology and game theory. Now this may seem strange at first. The players in game-theoretic models are usually assumed to be fully rational, whereas genes and other vehicles of evolution are assumed to behave in ways that are entirely mechanistic. Nonetheless, once a player is replaced by a population of individuals, and a mixed strategy corresponds to the proportions of the various strategies in the population, the formal structures in the two fields turn out to be

---

E-mail address: hart@huji.ac.il.
URL address: http://www.ma.huji.ac.il/~hart.

0899-8256/02/$ – see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S0899-8256(02)00502-X

very closely related. This has led to many ideas flowing back and forth. On the one hand, game-theoretic constructs—at times quite sophisticated—find their way into evolutionary arguments; on the other, the basic paradigm of natural selection is used to justify and provide foundations for many aspects of rational behavior. For a discussion of these issues, including a historical overview, the reader is referred to Hammerstein and Selten (1994) and to Aumann (1998).

The basic analogous notions in the two fields are "strategic equilibrium" (introduced by Nash, 1950) and "evolutionarily stable strategy" (introduced by Maynard Smith and Price, 1973). Roughly speaking, when a game is played by populations of individuals (with identical payoff functions), then evolutionarily stable strategies in essence yield a Nash equilibrium point. This type of relation has been established in a wide variety of setups, both static and dynamic (see the books of Hofbauer and Sigmund, 1998; Weibull, 1995; and Vega-Redondo, 1997).

Evolutionary models are based on two main ingredients: selection and mutation. *Selection* is a process whereby better strategies prevail; in contrast, *mutation*, which is relatively rare, generates strategies at random, be they better or worse. It is the combination of the two that allows for natural adaptation: New mutants undergo selection, and only the better ones survive. Of course, selection includes many possible mechanisms, be they biological (the payoff determines the number of descendants, and thus the share of better strategies increases), social (imitation, learning), individual (experimentation, stimulus response), and so on. What matters is that the process is "adaptive" or "improving," in the sense that the proportion of better strategies is likely to increase.

Such (stochastic) dynamic evolutionary models have been extensively analyzed in various classes of games in strategic (or normal) form, starting with Kandori et al. (1993) and Young (1993) (see also Foster and Young, 1990 and the books of Young, 1998 and Fudenberg and Levine, 1998). It turns out that certain Nash equilibria—like the risk-dominant ones—are more stable than others.

Here we consider *games in extensive form*, where a most complete description of the game is given, exactly specifying the rules, the order of moves, the information of the players, and so on. Specifically, we look at the simplest such games: *finite games of perfect information*. In these games, an equilibrium point can always be obtained by a so-called "backward induction" argument: Starting from the final nodes, each player chooses a best reply given the (already determined) choices of all the players that move after him. This results in an equilibrium point also in each subgame (i.e., the game starting at any node of the original game), whether that subgame is reached or not. Such a point is called a *subgame-perfect equilibrium*, or a *backward induction equilibrium*, a notion introduced by Selten (1965, 1975).

Since mutations are essentially small perturbations that make everything possible (i.e., every pure strategy has positive probability), and, as the perturbations go to zero, this yields in the limit the subgame-perfect equilibrium points,[1] it is only natural to expect that evolutionary models with low mutation rates should lead to these same points. However, the literature until now has found the above claim to be false in general: Evolutionary models do not necessarily pick out the backward induction equilibria. Specifically, except for special classes of games, equilibria other than the backward induction ones also turn out to be "evolutionarily stable" (see Nöldeke and Samuelson, 1993; Gale et al., 1995; Cressman and Schlag, 1998; and the books of Samuelson, 1997 and Fudenberg and Levine, 1998).

1.2. Examples

Even without specifying exactly how selection and mutation operate, we can get some intuition by considering a few examples. The first one is the classical example of the two-person game Γ_1 of Fig. 1. It possesses two Nash equilibria in pure strategies: $b = (b^1, b^2)$ and $c = (c^1, c^2)$; the first, b, is the backward induction equilibrium. Assume that at each one of the two nodes 1 and 2 there is a population of individuals playing the game in that role. The populations are distinct, and each individual plays a pure action at his node.[2] If everyone at node 1 plays b^1 and everyone at node 2 plays b^2, then any mutant in population 1 that plays c^1 will get a payoff of 1 instead of 2, so selection will wipe him out; the same goes for any mutant at node 2. Therefore the backward induction equilibrium b is "stable." Now assume that we are in the c equilibrium: All the individuals at 1 play c^1 and all the individuals at 2 play c^2. Again, a mutant at 1 loses relative to his population: Instead of 1 he gets 0 (since the individuals at 2 that he will meet play c^2). But now a mutant at 2 that plays b^2 gets the same payoff as a c^2-individual, so selection has no effect at node 2. Since node 2 is not

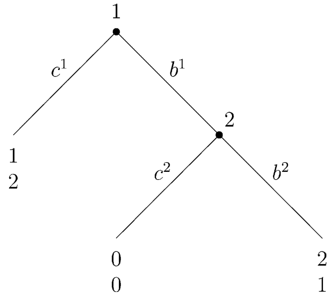

Fig. 1. The game Γ_1.

---

[1] Recall that these are games of perfect information, where "trembling-hand perfection" is the same as "subgame perfection."

[2] We say, for example, that an individual at node 2 "plays b^2" if he is programmed (by his "genes") to play b^2 whenever he is in a situation to choose (between c^2 and b^2).

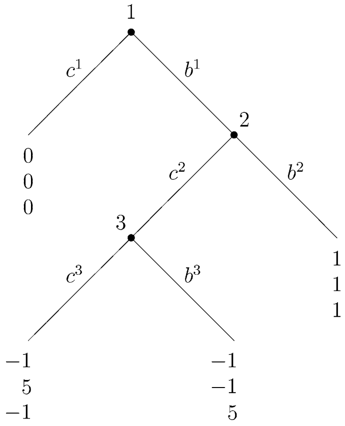

Fig. 2. The game Γ_2.

reached, all actions at 2 yield the same payoff; there is no "evolutionary pressure" at 2. Mutations in the population at 2, because they are not wiped out, keep accumulating (there is "genetic drift"). Eventually, a state will be reached where more than half the population at 2 consists of b^2-individuals.[3] At this point the action b^1 at 1 gets a higher expected payoff than the action c^1, and thus selection at 1 favors b^1. So the proportion of b^1 at node 1 becomes positive (and increases), which renders node 2 reachable. Once 2 is reached, evolutionary pressure there—i.e., selection—becomes effective, and it moves population 2 towards the better strategy b^2. This only increases the advantage of b^1 over c^1, and the whole system gets to the $b = (b^1, b^2)$ equilibrium.

To summarize: In Γ_1, evolutionary dynamics lead necessarily to b, the backward induction equilibrium; in other words, b is the evolutionarily stable equilibrium.

The next example is the three-player game Γ_2 of Fig. 2 (see Nöldeke and Samuelson, 1993; or Samuelson, 1997, Example 8.2). The backward induction equilibrium is $b = (b^1, b^2, b^3)$; the other pure Nash equilibrium—which is not perfect—is $c = (c^1, c^2, c^3)$. Let us start from a state where all individuals at each node i play their backward induction action b^i. Nodes 1 and 2 are reached, whereas node 3 is not. Therefore there is no selection operating at node 3, and mutations move the population at 3 randomly. As long as the proportion of b^3 is at least 2/3, the system is in equilibrium. Once it goes below 2/3—which, again, must happen eventually—the best reply of 2 becomes c^2; selection then moves the population at 2 toward c^2. But then node 3 is no longer unreached, so selection starts affecting the population at 3, moving it toward the best reply

---

[3] The assumption is that mutations have positive—though small—probability at each period. This yields a "random walk," and any proportion of b^2 and c^2 will occur eventually (with probability one).

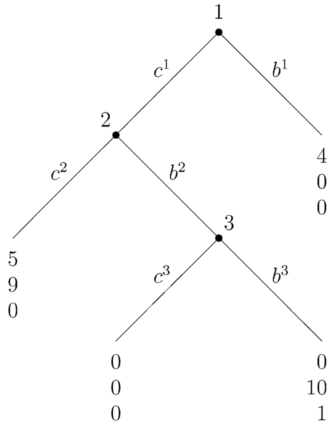

Fig. 3. The game Γ_3.

there, b^3. Thus, as soon as the proportion of b^3 drops below $2/3$, the evolutionary dynamic immediately pushes it back up; it is as if there were a "reflecting barrier" below the $2/3$ mark. Selection at 2 then moves back toward b^2. Meanwhile the population at 1, which is playing b^1, can move only a little, if at all.[4] Therefore we have essentially shown that the equilibrium component[5] of b—where b^i is played at $i = 1, 2$ and b^3 is played at 3 with proportion at least $2/3$—is evolutionarily stable. Moreover, since b and its component must eventually be reached from any state—by appropriate mutations—it follows that other equilibria, in particular c, are *not* stable. This conclusion differs from the result of Nöldeke and Samuelson (1993): In their model, the non-subgame-perfect equilibrium c also belongs to the stable set; see (a) in Section 5.3 for a more extensive discussion.

Consider now another three-player game: the game Γ_3 given in Fig. 3. The backward induction equilibrium is b, at which both nodes 2 and 3 are unreached. The populations at 2 and 3 therefore move by mutations. Eventually, when the proportion of b^2 at node 2 gets below $1/5$, selection at 1 will move the population at 1 from b^1 to c^1. At that point both 2 and 3 are reached, and which action of 2 is the best reply at 2 depends on the composition of the population at 3. If less than $9/10$ of them play b^3 (which is possible, and even quite probable,[6] given that only random mutations have affected 3 until now), then c^2 is the best reply at 2, and selection keeps decreasing the proportion of b^2. Again, it is quite probable

---

[4] Only when the proportion of b^2 drops below $1/2$ will selection affect node 1.

[5] We say that two (mixed) Nash equilibria belong to the same *equilibrium component* if their equilibrium paths coincide, and they differ only at unreached nodes (for generic games, this corresponds to a "connected component").

[6] We take "quite probable" to mean that the probability of its happening is positive and bounded away from zero (as the rate of mutation goes to zero).

for the proportion of b^2 to get all the way down to 0 (from $1/5$) long before the proportion of b^3 at 3 will have increased to $9/10$. What this discussion shows is that, in the game Γ_3, the non-subgame-perfect equilibrium c (together with its equilibrium component) cannot be ruled out; evolutionary dynamic systems may well be in such states a positive fraction of the time.

However, we claim that such behavior *cannot occur if the populations are large enough.*[7]

1.3. *This paper*

As stated above, the games studied in this paper are finite extensive-form games with perfect information. We assume that the backward induction equilibrium is unique; this holds when the game is generic (i.e., in almost every game). At each node there is a distinct population of individuals that play the game in the role of the corresponding player. Each individual is fully characterized by his action, i.e., by the pure choice that he makes at his node (of course, this goes into effect only if his node is reached[8]). We will refer to such a population game as a "gene-normal form" (it parallels the "agent-normal form").

The games are analyzed in a dynamic framework. The model is as follows: At each period, one individual is chosen at random[9] in each population. His current action, call it a^i, may then change by selection, by mutation, or it may not change at all. Selection replaces a^i by another action which, against the other populations currently playing the game (i.e., "against the field"), yields a higher payoff than a^i. Of course, this can only be if such a better action exists (if there are many, one of them is chosen at random). Mutation replaces a^i by an arbitrary action, chosen at random. Finally, all the choices at each node are made independently. This model—which we refer to as the "basic model"—is essentially a most elementary process that provides for both adaptive selection and mutation. It turns out that the exact probabilities of all the above choices do not matter; what is essential is that all of them be bounded away from zero (this is the "general model").

Such dynamics yield an ergodic system[10] whose long-run behavior is well described by the corresponding unique invariant distribution, which, for each state, gives the (approximate) frequency of that state's occurrence during any large time interval. The mutations are rare; we are therefore interested in those

---

[7] How large may depend on the mutation rate.

[8] This action is thus the individual's "genotype"—the hard-wired programming by the genes; it becomes his "phenotype"—his actual revealed behavior—when his node is reached and it is his turn to play.

[9] Uniformly, i.e., each individual has the same probability of being chosen.

[10] Irreducible (mutations make every state reachable from any other state) and aperiodic.

states which occur with positive frequency, however low the mutation rate.[11] We call such states “evolutionarily stable.” A preliminary result is that the backward induction equilibrium is always evolutionarily stable. However, as the examples show, other Nash equilibria may be evolutionarily stable as well.

We therefore add another factor: The populations are large. This yields:

> **Main Result**: *The backward induction equilibrium becomes in the limit the only evolutionarily stable outcome as the mutation rate decreases to zero and the populations increase to infinity* (*provided that the expected number of mutations per generation does not go to zero*).

In other words: Evolutionary dynamic systems, consisting of adaptive selection and rare mutations, lead in large populations to most of the individuals most of the time playing their backward induction strategy. Observe that this applies to reached as well as to unreached nodes; for example, in game Γ_2 we have most of the individuals at all nodes i—*including* node 3—playing b^i. Evolutionary stability in large populations picks out not merely the equilibrium component of b, but b itself.[12] The intuition for the role of the large population assumption will be provided in Section 4.1 below. Suffice it to say here that it has to do with a change of action (whether by mutation or selection) being much less likely for a *specific* individual than for an *arbitrary* individual in a large population. This leads to considerations of “sequential” rather than “simultaneous” mutations. As a further consequence, unlike in most of the evolutionary game-theoretic literature,[13] our result does *not* rely on comparing different powers of the infinitesimal mutation rate (which require extremely long waiting times); single mutations suffice.[14]

We conclude this introduction with two comments. First, two almost diametrically opposed approaches lead to the backward induction equilibrium. One approach (Aumann, 1995), in the realm of full rationality, assumes that all players are rational (i.e., they never play anything which they know is not optimal), and moreover that this fact is commonly known to them (i.e., each player knows that everyone is rational, and also knows that everyone knows that everyone is rational, and so on). The other approach (of this paper), in the realm of evolutionary

---

[11] More precisely, states whose probability, according to the invariant distribution, is bounded away from zero as the probability of mutation goes to zero; these are called “stochastically stable states” by Young (1993, 1998).

[12] Actually, an arbitrarily small neighborhood of b.

[13] Exceptions are Nöldeke and Samuelson (1993) and the “modified co-radius” of Ellison (2000).

[14] The static notion of an “evolutionarily stable strategy” is also based on single mutations.

dynamics, is essentially machine-like and requires no conscious optimization or rationality.[15] It is striking that such disparate models converge.[16]

Second, note that the backward induction equilibrium is by no means the conclusive outcome: Substantial assumptions are needed, like large populations, or common knowledge of rationality.

The paper is organized as follows: Section 2 presents the model: the extensive-form game (in Section 2.1), the associated population game (in Section 2.2), and the evolutionary dynamics (in Section 2.3). The results are stated in Section 3, which also includes the proof of some preliminary results for fixed populations (in Section 3.1). The Main Result, stated in Section 3.2, is proved in Section 4. The intuition behind our result is presented in Section 4.1, followed by an informal outline of the proof (in Section 4.2). We conclude in Section 5 with a discussion of various issues, including relations to the existing literature and possible extensions and generalizations of our results.

2. The model

2.1. The game

Let Γ be a finite extensive-form game with perfect information. We are thus given a rooted tree; each non-terminal vertex corresponds to a *move*. It may be a chance move, with fixed positive probabilities for its outgoing branches; or a move of one of the players, in which case the vertex is called a *node*. The set of nodes is denoted N. It is convenient to view the game in "agent-normal form:" At each node there is a different agent, and a player consists of a number of agents with identical payoff functions. For each node $i \in N$, the agent there—called "agent i"—has a set of choices A^i, which is the set of outgoing branches at i. We refer to a^i in A^i as an *action* of i, and we put $A := \prod_{i \in N} A^i$ for the set of N-tuples of actions. At each terminal vertex (a *leaf*) there are associated payoffs to all agents; let[17] $u^i : A \to \mathbb{R}$ be the resulting payoff function of agent i (i.e., for each $a = (a^j)_{j \in N} \in A$: if there are no chance moves, then $u^i(a)$ is the payoff of i at the leaf that is reached when every agent $j \in N$ chooses a^j; if there are chance moves, it is the appropriate expectation). Of course, if i and j are agents of the same player, then $u^i \equiv u^j$. As usual, the payoff functions are extended multilinearly to *randomized* (or *mixed*) actions; thus $u^i : X \to \mathbb{R}$, where $X := \prod_{i \in N} X^i$ and $X^i := \Delta(A^i) = \{x^i \in \mathbb{R}_+^{A^i} : \sum_{a^i \in A^i} x^i_{a^i} = 1\}$, the unit simplex on A^i, is the set of probability distributions over A^i.

---

[15] The biological mechanisms of selection are entirely automatic; other selection processes (like learning, imitation, and so on) may well use some form of rationality or "bounded rationality."

[16] For an interesting discussion of these matters, see Aumann (1998) (in particular, pages 191–195).

[17] $\mathbb{R}$ is the real line.

For each node $i \in N$, let $N(i)$ be the set of nodes that are successors (not necessarily immediate) of i in the tree, and let $\Gamma(i)$ be the subgame of Γ starting at the node i. For example, if $1 \in N$ is the root then $N(1) = N \backslash \{1\}$ and $\Gamma(1) = \Gamma$; in general, $j \in N(i)$ if and only if the unique path from the root to j goes through i, and the set of nodes of $\Gamma(i)$ is $N(i) \cup \{i\}$.

An N-tuple of randomized actions $x = (x^i)_{i \in N} \in X$ is a *Nash equilibrium* of Γ if[18] $u^i(x) \geqslant u^i(y^i, x^{-i})$ for every $i \in N$ and every $y^i \in X^i$. It is moreover a *subgame-perfect* (or *backward induction*) *equilibrium* of Γ if it is a Nash equilibrium in each subgame $\Gamma(i)$, for all $i \in N$. This is equivalent to each x^i being a best reply of i in $\Gamma(i)$ when every $j \in N(i)$ plays x^j. Such an equilibrium is therefore obtained by *backward induction*, starting from the final nodes (those nodes i with no successors, i.e., with $N(i) = \emptyset$) and going towards the root. We will denote by *EQ* and *BI* the set of Nash equilibria and the set of backward induction equilibria, respectively, of the game Γ; thus $BI \subset EQ \subset X$.

At this point it is useful to point out the distinction between a best reply of i in the whole game Γ—which we call a *global best reply*—and a best reply of i in the subgame $\Gamma(i)$—which we call a *local best reply.* Thus a local best reply is always a global best reply, but the converse is not necessarily true.[19] If i is reached (i.e., when all agents on the path from the root to i make the choice along the path with positive probability), then the two notions coincide. If i is not reached, then the payoff of i in Γ is independent of his action, and thus every action in A^i (and every mixed action in X^i) is a global best reply of i—but not necessarily a local best reply. The difference between a Nash equilibrium and a subgame-perfect equilibrium is precisely that in the former each action of an agent that is played with positive probability is a global best reply to the others' (mixed) actions, whereas in the latter it is additionally a local best reply.

The classical result of Kuhn (1953) states that there always exists a *pure* backward induction equilibrium; the proof constructs it by backward induction. We assume here that the game Γ has a *unique* backward induction equilibrium, which must therefore be pure; we denote it $b = (b^i)_{i \in N} \in A$, and refer to b^i as the "backward induction action of i." This uniqueness is true generically, i.e., for almost every game. For instance, when there are no chance moves, it suffices for each player to have different payoffs at different leaves.

2.2. The gene-normal form

We now consider a *population game* associated to Γ: At each node $i \in N$ there is a non-empty population $M(i)$ of individuals playing the game in the role of agent i. We assume that *the populations at different nodes are distinct*:

$$M(i) \cap M(j) = \emptyset \quad \text{for all } i \neq j. \tag{2.1}$$

---

[18] We write x^{-i} for the $(|N| - 1)$-tuple of actions of the other agents, i.e., $x^{-i} = (x^j)_{j \in N \backslash \{i\}}$.

[19] One should not confuse these with the parallel notions for optima (where global implies local).

This assumption is not completely innocuous; see the discussion below and also (b) in Section 5.2. Each individual $q \in M(i)$ is characterized by a pure action in A^i, which we denote by $\omega_q^i \in A^i$; put $\omega^i = (\omega_q^i)_{q\in M(i)}$ and $\omega = (\omega^i)_{i\in N}$. For each $a^i \in A^i$, let[20]

$$x^i_{a^i} \equiv x^i_{a^i}\big(\omega^i\big) := \frac{|\{q \in M(i)\colon \omega_q^i = a^i\}|}{|M(i)|} \tag{2.2}$$

be the proportion of population $M(i)$ that plays the action a^i; then $x^i \equiv x^i(\omega^i) := (x^i_{a^i}(\omega^i))_{a^i\in A^i} \in X^i$ may be viewed as a mixed action of i. The payoff of an individual $q \in M(i)$ is defined as his average payoff against the other populations, i.e., $u^i(\omega_q^i, x^{-i})$; we shall slightly abuse notation by writing this as $u^i(\omega_q^i, \omega^{-i})$.

We refer to the above model as the *gene-normal form of Γ* (by (2.1), it is the counterpart, in population games, of the “agent-normal form”).

This model is clear and needs no explanation when all the players in Γ are distinct (i.e., when each player plays at most once in Γ). When however a player may play more than once (and thus have more than one agent), then a “biological” interpretation is as follows: Each one of the player’s decisions (i.e., each one of his agents i) is controlled by a “gene,” whose various “alleles” correspond to the possible choices at node i (i.e., the set of alleles of gene i is precisely A^i). The genes of different nodes i and j of the same player are distinct (i.e., at different locations, or “loci”); were it the same gene, then the player would behave identically at the two nodes—and the appropriate representation would place the two nodes i and j in the same information set.[21,22]

2.3. The dynamics

We come now to the dynamic model. A *state* ω of the system specifies the pure action of each individual in each population; i.e., $\omega = (\omega^i)_{i\in N}$, where $\omega^i = (\omega_q^i)_{q\in M(i)}$ and $\omega_q^i \in A^i$ for each $i \in N$ and each $q \in M(i)$. Let $\Omega := \prod_{i\in N}(A^i)^{M(i)}$ be the state space. We consider discrete-time stochastic systems: Starting with an initial state[23] $\omega_1 \in \Omega$, a sequence of states $\omega_1, \omega_2, \ldots, \omega_t, \ldots$ in Ω is generated, according to certain probabilistic rules. These so-called “transition probabilities” specify, for each $t = 1, 2, \ldots,$ the probabilities $P[\omega_{t+1} = \tilde{\omega} \mid \omega_1, \ldots, \omega_t]$ that ω_{t+1} equals a state $\tilde{\omega} \in \Omega$, given the history

---

[20] $|Z|$ denotes the number of elements of a finite set Z.

[21] One would thus get a game of imperfect information, where information sets are not necessarily singletons. Moreover, observe that here a path may intersect an information set more than once; we are thus led naturally to “games of imperfect recall.”

[22] A case where, say, the decision at i is controlled by the two genes in locations 1 and 2, and the decision at j by the two genes in locations 2 and 3, is not considered here.

[23] As we shall see below, the process is irreducible and aperiodic; thus, in the long run, the starting state does not matter. Hence there will be no need to specify it.

$\omega_1, \omega_2, \ldots, \omega_t$. Our processes will be *stationary Markov chains*: The transition probabilities depend only on ω_t, the state in the previous period (and depend neither on the other past states $\omega_1, \ldots, \omega_{t-1}$, nor on the "calendar time" t). That is, there is a stochastic matrix[24] $Q = (Q[\tilde{\omega} \mid \omega])_{\tilde{\omega}, \omega \in \Omega}$ such that $P[\omega_{t+1} = \tilde{\omega} \mid \omega_1, \ldots, \omega_t] = Q[\tilde{\omega} \mid \omega_t]$ for every $\omega_1, \ldots, \omega_t,\ \tilde{\omega} \in \Omega$ and $t = 1, 2, \ldots$. The matrix Q is called the *one-step transition probability matrix*.

We present first a simple dynamic model, which we call the *basic model*. Assume that all populations are of equal size, say $m = |M(i)|$ for each $i \in N$. Let $\mu > 0$ and $\sigma > 0$ be given, such that $\mu + \sigma \leqslant 1$. The one-step transition probabilities $Q[\tilde{\omega} \mid \omega]$ are given by the following process, performed independently for each $i \in N$:

- Choose an individual $q(i) \in M(i)$ at random: All m individuals in $M(i)$ have the same probability $1/m$ of being chosen.
- Put $\tilde{\omega}^i_q := \omega^i_q$ for each $q \in M(i),\ q \neq q(i)$; i.e., all individuals in $M(i)$ except $q(i)$ do not change their actions.
- Choose one of $SE(i)$ ("selection"), $MU(i)$ ("mutation") and $NC(i)$ ("no change"), with probabilities σ, μ, and $1 - \mu - \sigma$, respectively.
- If selection $SE(i)$ was chosen, then define

  $$B^i \equiv B^i\big(q(i), \omega\big) := \big\{a^i \in A^i\colon\ u^i\big(a^i, \omega^{-i}\big) > u^i\big(\omega^i_{q(i)}, \omega^{-i}\big)\big\}; \tag{2.3}$$

  this is the set of "better actions"—those actions at node i that are strictly better in Γ, against the populations at the other nodes, than the action $\omega^i_{q(i)}$ of the chosen individual $q(i)$. If B^i is not empty, then the new action $\tilde{\omega}^i_{q(i)}$ of $q(i)$ is a randomly chosen better action: $\tilde{\omega}^i_{q(i)} := a^i$ with probability $1/|B^i|$ for each $a^i \in B^i$. If B^i is empty, then there is no change in $q(i)$'s action: $\tilde{\omega}^i_{q(i)} := \omega^i_{q(i)}$.
- If mutation $MU(i)$ was chosen, then $\tilde{\omega}^i_{q(i)}$ is a random action in A^i; i.e., $\tilde{\omega}^i_{q(i)} := a^i$ with probability $1/|A^i|$ for each $a^i \in A^i$.
- If no-change $NC(i)$ was chosen, then the action of $q(i)$ does not change: $\tilde{\omega}^i_{q(i)} := \omega^i_{q(i)}$.

For example, in the game Γ_1 of Section 1.2 (see Fig. 1; here $N = \{1, 2\}$, $A^1 = \{c^1, b^1\}$, $A^2 = \{c^2, b^2\}$), with populations of size $m = 3$, let $\omega = ((c^1, c^1, c^1), (b^2, b^2, c^2))$ and $\tilde{\omega} = ((b^1, c^1, c^1), (b^2, b^2, b^2))$, then $Q[\tilde{\omega} \mid \omega] = (1/3) \cdot (\mu/2 + \sigma) \cdot (1/3) \cdot (\mu/2)$. Indeed, the probability that $q(1) = 1$ is $1/3$; then c^1 changes to b^1 either by mutation, with probability $\mu \cdot (1/2)$, or by selection (since $B^1 = \{b^1\}$), with probability σ; similarly, the probability that $q(2) = 3$ is

---

[24] I.e., $Q[\tilde{\omega} \mid \omega] \geqslant 0$ for all $\tilde{\omega}, \omega \in \Omega$ and $\sum_{\tilde{\omega} \in \Omega} Q[\tilde{\omega} \mid \omega] = 1$ for every $\omega \in \Omega$.

$1/3$, and then c^2 changes to b^2 by mutation only (since $B^2 = \emptyset$), with probability $\mu \cdot (1/2)$.

A few remarks are now in order.

Remarks.

1. We have assumed that in each period there is at most one individual in each population that may change his action. This defines what is meant by "one period:" It is a time interval which is small enough that the probability of more than one individual changing his action in the same period is (relatively) negligible.[25] This is a standard construct in stochastic setups (recall, for instance, the construction of Poisson processes). Our assumption may thus be viewed as essentially nothing more than a convenient rescaling of time.[26] As we shall see below (in particular in Section 4.1), our arguments are based on *comparing* occurrence times, and are thus independent of the units in which time is measured.
2. The difference between mutation and selection is that mutation is "blind"—in the sense that *all* actions are possible—whereas selection is "directional"—*only better* actions are possible. We emphasize that "better" is understood, as it should be, with respect to the payoffs in the whole game, i.e., "globally better." Of course, "selection" may stand for various processes of adaptation, imitation, learning, experimentation, and so on. Our selection thus assumes that better actions fare better; it is a "*better-reply selection dynamic*." It is also a "strictly aggregate monotonic" mechanism,[27] in the sense that, for each i, if we hold all populations except i fixed, then selection at i never decreases the average payoff of population i, and it has a positive probability of strictly increasing it whenever there are individuals who are not playing a best-reply action. See also the general model below, where the selection probability of a better action may be proportional to its current proportion in the population.
3. Actions are compared (see (2.3)) according to their *average payoffs* against the other populations. This is a standard assumption in the literature. It is correct, for instance, if everyone plays against everyone else; i.e., all $m^{|N|}$ combinations (of one individual from each population) play the game. When the populations are large—as is the case here—it is also approximately correct when each individual plays against random samples from the other populations.[28]

---

[25] Our results do not need simultaneous mutations; these may indeed be ignored.

[26] Or, alternatively, as an appropriate discretization of a continuous-time process.

[27] This is called a "strict myopic adjustment dynamic" in Swinkels (1993).

[28] The larger the population, the shorter the period (see Remark 1 above); it may thus be difficult to play against everyone in one period. See also (g) in Section 5.1.

4. The basic dynamic is determined by two parameters,[29] μ and σ. As we shall see below, what really matters is that μ be small relative to σ; formally, $\mu \to 0$ while $\sigma > 0$ is fixed: Mutations are rare relative to selection. Equivalently, we could well take $\sigma = 1 - \mu$, and thus have only one parameter. We have preferred to add the no-change case since it allows for more general interpretations. For instance, the no-change periods may be viewed as "payoff accumulation" periods, or as "selection realization" periods (i.e., periods during which actual selection occurs[30]).
5. The one-step transition probabilities are defined to be independent over the nodes; this just means that the transitions are *conditionally independent*. In general, the evolution of one population will depend substantially on that of other populations.

The basic model is essentially a most simple model that captures the evolutionary paradigm of selection and mutation. It may appear however as too specific. Therefore we now present a general class of dynamic models, which turn out to lead to the same results.

The *general model* is as follows: We are given a *mutation rate* parameter $\mu > 0$ and populations $M(i)$ at all nodes $i \in N$, which may be a different size at different nodes. The process is a stationary Markov chain, whose one-step transition probability matrix $Q = (Q[\tilde{\omega} \mid \omega])_{\tilde{\omega},\omega\in\Omega}$ satisfies:

- Conditional independence over $i \in N$, i.e.,[31]

$$Q[\tilde{\omega} \mid \omega] = \prod_{i\in N} Q\big[\tilde{\omega}^i \bigm| \omega\big]. \tag{2.4}$$

- For each $i \in N$, one individual $q(i) \in M(i)$ is chosen, such that there exist constants $\gamma_1, \gamma_2 > 0$ with

$$\frac{\gamma_1}{|M(i)|} \leqslant Q\big[q(i) = q \bigm| \omega\big] \leqslant \frac{\gamma_2}{|M(i)|} \quad \text{for each } q \in M(i); \quad \text{and} \tag{2.5}$$

$$Q\big[\tilde{\omega}^i_q = \omega^i_q \text{ for all } q \in M(i)\backslash\{q(i)\} \bigm| \omega\big] = 1. \tag{2.6}$$

- There exists a constant $\beta > 0$ such that, for each $i \in N$,

$$Q\big[\tilde{\omega}^i_{q(i)} = a^i \bigm| \omega\big] \geqslant \beta x^i_{a^i}(\omega) \quad \text{for each } a^i \in B^i, \tag{2.7}$$

where $B^i \equiv B^i(q(i), \omega)$ is the set of strictly better actions, as defined in (2.3).

---

[29] Once the game Γ and the population size m are given.

[30] This may help to justify the fact that our selection mechanism is not continuous (any "better" action has probability bounded away from zero, whereas an "equally good" action has zero probability): Indeed, selection makes even a slightly better action "win," given enough time.

[31] For each $\omega \in \Omega$, we view $Q[\cdot \mid \omega]$ as a probability distribution over Ω; derived probabilities—like its marginals, etc.—will also be denoted by $Q[\cdot \mid \omega]$.

- There exist constants $\alpha_1, \alpha_2 > 0$ such that, for each $i \in N$,

$$Q\big[\tilde{\omega}^i_{q(i)} = a^i \bigm| \omega\big] \geqslant \alpha_1 \mu \quad \text{for each } a^i \in A^i; \quad \text{and} \tag{2.8}$$

$$Q\big[\tilde{\omega}^i_{q(i)} = a^i \bigm| \omega\big] \leqslant \alpha_2 \mu \quad \text{for each } a^i \notin B^i,\ a^i \neq \omega^i_{q(i)}. \tag{2.9}$$

Without loss of generality, all parameters α_1, α_2, β, γ_1, γ_2 are taken to be the same for all $i \in N$ (if needed, replace them by the appropriate maximum or minimum over i). To see that the basic model is a special case of the general model, take $\gamma_1 = \gamma_2 = 1$, $\beta = \sigma/|A^i|$, and $\alpha_1 = \alpha_2 = 1/|A^i|$.

Notice that we now allow selection to switch to better actions with probabilities that are proportional to their current proportions in the population.[32] If, for instance, (2.7) is satisfied as an equality and, say, the best-reply action a^i is currently played by k individuals, then the probability that a chosen non-a^i-individual will switch to a^i by selection is $\beta k/|M(i)|$ (rather than $\geqslant \beta = \sigma/|A^i|$, as in the basic model). When k is small, this probability becomes low (and converges to 0 with k). Moreover, if a^i is not currently present in the population, then selection cannot introduce it.[33,34] Such dynamics are appropriate in particular in *imitation*-type models (where the “visibility” of an action depends on its prevalence in the population). For a simple example, suppose that the chosen individual samples a random individual in his population, and switches to his action (with probability $\beta > 0$) if it currently yields a higher payoff than his own. Note that in this case $x^i_{a^i}$, the proportion in the population of a better action a^i, increases by selection at a rate that is proportional to $x^i_{a^i}$.

The general model thus assumes that: (i) the probabilities that various individuals in the same population will be chosen are comparable;[35] (ii) the (relative[36]) effect of selection—towards better actions—is bounded away from zero (independently of μ); and (iii) the effect of mutation—whereby every action is possible—is of the order μ. The reader is referred to Sections 5.1 and 5.2 for further generalizations.

3. The results

3.1. Preliminary results

A general model with a one-step transition matrix Q satisfying (2.4)–(2.9) yields a Markov chain which is *irreducible*, since the probability of reaching any

---

[32] This condition was suggested by Ilan Eshel.

[33] A requirement suggested by Karl Schlag.

[34] To understand why this weakening of the selection mechanism does not affect our main result, the reader is referred to Footnote 53 and Lemma 4.6.

[35] I.e., the ratios $Q[q(i) = q \mid \omega]/Q[q(i) = q' \mid \omega]$ are uniformly bounded.

[36] I.e., the change in x^i relative to x^i.

state $\omega' \in \Omega$ from any other state $\omega \in \Omega$ is positive (as follows from (2.5) and (2.8), by using an appropriate sequence of mutations). Hence there exists a unique *invariant distribution* π on Ω; i.e., a unique $\pi \in \Delta(\Omega)$ satisfying $\pi = \pi Q$, or

$$\pi[\tilde{\omega}] = \sum_{\omega \in \Omega} \pi[\omega] Q[\tilde{\omega} \mid \omega]$$

for every $\tilde{\omega} \in \Omega$. The Markov chain is moreover *aperiodic*, since, for instance, the probability of staying at a Nash equilibrium state is positive. Therefore the long-run behavior of the process is well described by π, in the following two senses:

- In any long enough period of time, the relative frequency of visits at a state ω is approximately $\pi[\omega]$; i.e., for every $\omega \in \Omega$,
$$\lim_{T_2 - T_1 \to \infty} \frac{|\{t\colon T_1 < t \leqslant T_2,\ \omega_t = \omega\}|}{T_2 - T_1} = \pi[\omega].$$
- The probability that the state ω occurs at a period t is approximately $\pi[\omega]$ for large t; i.e., for every $\omega \in \Omega$,
$$\lim_{t \to \infty} P[\omega_t = \omega] = \pi[\omega].$$

The two properties hold regardless of the initial state; moreover, they hold not only for single states ω but also for any set of states $\Theta \subset \Omega$.

We are interested in the behavior of the process when the mutation rate is low, i.e., in the limit of the invariant distribution π as $\mu \to 0$ and all the other parameters (the game Γ, the population sizes $|M(i)|$ and the constants $\alpha_1, \alpha_2, \beta, \gamma_1, \gamma_2$) are fixed. We call a state $\omega \in \Omega$ *evolutionarily stable* if its invariant probability $\pi[\omega]$ is bounded away from zero as $\mu \to 0$, i.e., if[37] $\liminf_{\mu \to 0} \pi[\omega] > 0$. Recall that each state $\omega \in \Omega$ may be viewed as an N-tuple of mixed actions $x(\omega) = (x^i(\omega^i))_{i \in N} \in X$ (see (2.2)). The invariant distribution π on Ω therefore induces a probability distribution[38] $\hat{\pi} := \pi \circ (x)^{-1}$ over X; i.e., $\hat{\pi}[Y] := \pi[\{\omega \in \Omega\colon x(\omega) \in Y\}]$ for every (measurable) $Y \subset X$. We therefore call an N-tuple of mixed actions $x \in X$ *evolutionarily stable* if there are evolutionarily stable states $\omega \in \Omega$ with $x(\omega) = x$, i.e., if $\liminf_{\mu \to 0} \hat{\pi}[x] > 0$. The following result states that the backward induction equilibrium b is always evolutionarily stable.

---

[37] This is called "stochastic stability" (Foster and Young, 1990 and Young, 1993, 1998), "long-run equilibrium" (Kandori et al., 1993), "in the support of the limit distribution" (Samuelson, 1997 and Fudenberg and Levine, 1998).

[38] $(x)^{-1}: X \to \Omega$ denotes the inverse of the mapping $x : \Omega \to X$. We could have equivalently defined the dynamics directly on the space X of mixed action profiles (by identifying all states ω with the same $x(\omega)$ and taking the expected transition probabilities); we found however that our (more primitive) model is more transparent.

Theorem 3.1. *For each* $\mu > 0$, *let* π_μ *be the unique invariant distribution of a dynamic process given by a one-step transition matrix* $Q \equiv Q_\mu$ *satisfying* (2.4)–(2.9). *Then*

$$\liminf_{\mu \to 0} \hat{\pi}_\mu[b] > 0.$$

Proof. Assume without loss of generality that $Q_\mu \to Q_0$ and $\pi_\mu \to \pi_0$ as $\mu \to 0$ (take a convergent subsequence if needed; recall that the state space Ω is finite and fixed). The invariance property $\pi_\mu = \pi_\mu Q_\mu$ becomes $\pi_0 = \pi_0 Q_0$ in the limit; thus π_0 is an invariant distribution of Q_0 (but Q_0 is in general not irreducible, so its invariant distribution need not be unique). Now Q_0 allows no mutations (by (2.9) and $\mu \to 0$), so only selection applies.

First, we claim that

$$Q_0 \text{ is } \textit{acyclic}; \tag{3.1}$$

i.e., there are no $\omega_0, \omega_1, \ldots, \omega_t, \ldots, \omega_T \in \Omega$ satisfying $\omega_t \neq \omega_{t-1}$ and $Q_0[\omega_t \mid \omega_{t-1}] > 0$ for every $t = 1, \ldots, T$ and $\omega_T = \omega_0$. Indeed, at a final node $i \in N$ (i.e., when $N(i) = \emptyset$), selection can only increase the sum of the "local" payoffs (in $\Gamma(i)$) of the population $M(i)$, i.e.,[39] $\sum_{q \in M(i)} u^i_{\Gamma(i)}(\omega^i_q)$; therefore $\omega^i_T = \omega^i_0$ implies that $\omega^i_t = \omega^i_0$ for all t; i.e., the population at i never moves. The same applies at any node i for which there were no changes at all its descendant nodes $N(i)$; backward induction thus yields $\omega^i_t = \omega^i_0$ for all t and all $i \in N$, a contradiction.

Let $\Omega_0 := \{\omega \in \Omega\colon\ Q_0[\omega \mid \omega] = 1\}$ be the set of absorbing states of Q_0. Eq. (3.1) implies that, under Q_0, an absorbing state must always be reached:

$$\pi_0[\Omega_0] = 1. \tag{3.2}$$

In other words, only Q_0-absorbing states may be evolutionarily stable.

Let $\omega, \omega' \in \Omega_0$ be two absorbing states; if ω' can be reached from ω by one mutation step in one population followed by any number of selection steps, then we will say that ω' is *one-mutation-reachable* from ω. We claim that[40]

$$\begin{aligned} &\text{if} \quad \omega' \text{ is one-mutation-reachable from } \omega, \\ &\text{then} \quad \pi_0[\omega] > 0 \quad \text{implies} \quad \pi_0[\omega'] > 0. \end{aligned} \tag{3.3}$$

Indeed, the invariance property $\pi_\mu = \pi_\mu Q_\mu$ implies[41] $\pi_\mu = \pi_\mu Q^k_\mu$ for any integer $k \geqslant 1$, and thus

$$\pi_\mu[\omega'] \geqslant \pi_\mu[\omega'] Q^k_\mu[\omega' \mid \omega'] + \pi_\mu[\omega] Q_\mu[\omega'' \mid \omega] Q^{k-1}_\mu[\omega' \mid \omega''],$$

---

[39] We write $u^i_{\Gamma(i)}$ for the payoff function of i in the subgame $\Gamma(i)$.

[40] This is shown in a general setup by Samuelson (1997), Proposition 7.7(ii).

[41] Q^k, the kth power of the one-step transition probability matrix Q, gives precisely the k-steps transition probabilities.

where ω'' satisfies $Q_\mu[\omega'' \mid \omega] \geqslant c_1\mu$ for some $c_1 > 0$ (by (2.5) and (2.8); this is the mutation step) and $Q_0^{k-1}[\omega' \mid \omega''] = c_2 > 0$ (these are the selection steps; thus Q_0 rather than Q_μ). Also, since ω' is a Q_0-absorbing state, it can change only by mutations, so $Q_\mu^k[\omega' \mid \omega'] \geqslant 1 - c_3\mu$ for an appropriate constant $c_3 > 0$ (by (2.9)). Therefore $c_3\mu\pi_\mu[\omega'] \geqslant c_1\mu\pi_\mu[\omega]Q_\mu^{k-1}[\omega' \mid \omega'']$, which, after dividing by μ and then letting $\mu \to 0$, yields

$$\pi_0[\omega'] \geqslant \frac{c_1}{c_3}\pi_0[\omega]Q_0^{k-1}[\omega' \mid \omega''] = \frac{c_1c_2}{c_3}\pi_0[\omega] > 0,$$

and thus $\pi_0[\omega'] > 0$.

Let $i \in N$ be a final node. We claim that there is an absorbing state $\omega \in \Omega_0$ with $\pi_0[\omega] > 0$, at which all the population at i plays the backward induction action b^i. If not, let $\omega \in \Omega_0$ be such that $\pi_0[\omega] > 0$ and the proportion $x^i_{b^i}(\omega)$ of b^i is maximal (among all such $\omega \in \Omega_0$ with $\pi_0[\omega] > 0$; recall that $\pi_0[\Omega_0] = 1$). Thus $x^i_{b^i}(\omega) < 1$. Consider a mutation of a non-b^i-individual into b^i (and no changes at all nodes $j \neq i$; recall that $\omega \in \Omega_0$), followed by any number of selection periods until a state $\omega' \in \Omega_0$ is (necessarily) reached. By (3.3), we have $\pi_0[\omega'] > 0$. But $x^i_{b^i}(\omega') > x^i_{b^i}(\omega)$, since the first mutation step increased this proportion, and the selection steps could not have decreased it; this contradicts our choice of ω. Therefore there are states $\omega \in \Omega_0$ with $\pi_0[\omega] > 0$ and $x^i_{b^i}(\omega) = 1$.

The same argument applies at any node $i \in N$ for which there are absorbing states $\omega \in \Omega_0$ with $\pi_0[\omega] > 0$ and $x^j_{b^j}(\omega) = 1$ for all $j \in N(i)$ (just choose from among these states one with maximal $x^i_{b^i}(\omega)$). Therefore, by backward induction, we get $\pi_0[\omega_b] > 0$ for that state $\omega_b \in \Omega_0$ with $x^i_{b^i}(\omega_b) = 1$ for all $i \in N$—thus $\hat{\pi}_0[b] > 0$. □

Remarks.

1. Evolutionarily stable states must be absorbing states for the dynamic without mutations Q_0 (cf. (3.2)). Clearly, every Nash equilibrium state is Q_0-absorbing; but in the general case, where selection cannot induce a switch to a better action unless it is currently present in the population, other states may also be absorbing: For instance, all states corresponding to pure action profiles (i.e., ω where $x(\omega)$ is pure).[42] In fact, profiles that are not Nash equilibria may well be evolutionarily stable. For an example, consider dynamics which satisfy (2.7) as an equality, applied to the following two-person game Γ_1': It is like the game Γ_1 of Fig. 1 in the Introduction, except that we invert the sign of all the payoffs of player 2 (and thus the backward induction equilibrium is now (c^1, c^2), with payoffs $(1, -2)$). A mixed profile $x = ((\xi^1, 1-\xi^1), (\xi^2, 1-\xi^2))$, where ξ^i denotes the proportion of c^i—we will write this as $x = (\xi^1, \xi^2)$ for short—is a Nash equilibrium whenever

---

[42] We thank an anonymous referee for pointing this out.

$\xi^1 = 1$ and $\xi^2 \geqslant 1/2$. In contrast, the set of Q_0-absorbing profiles contains all x with $\xi^1 = 1$ (i.e., for *any* ξ^2), as well as the pure $(0,1)$ and $(0,0)$ (i.e., (b^1, b^2) and (b^1, c^2)). Now Theorem 3.1 implies that $\hat{\pi}_0[b] = \hat{\pi}_0[(1,1)] > 0$. One mutation (in population 2) leads from $(1,1)$ to[43] $(1, 1-1/m)$, which is an absorbing state. Therefore, by (3.3), we get $\hat{\pi}_0[(1, 1-1/m)] > 0$. In the same manner it then follows from this that $\hat{\pi}_0[(1, 1-2/m)] > 0$, and so on, up to $\hat{\pi}_0[(1,0)] > 0$—but $(1,0)$ (i.e., (c^1, b^2)) is not a Nash equilibrium.[44] If one assumes that there exists a constant $\beta' > 0$ such that

$$Q\big[\tilde{\omega}^i_{q(i)} \in B^i \bigm| \omega\big] \geqslant \beta', \tag{3.4}$$

whenever $B^i \neq \emptyset$, then only Nash equilibria can be fixed points of the selection mechanism: $\omega \in \Omega_0$ implies[45] $x(\omega) \in EQ$. From (3.2) we thus have: If the dynamic satisfies (2.4)–(2.9) and (3.4), then

$$\lim_{\mu \to 0} \hat{\pi}_\mu[EQ] = 1.$$

Note that the basic dynamic satisfies (3.4).

2. It is not difficult to show that in the game Γ_1 of Fig. 1 in the Introduction one obtains $\lim_{\mu\to 0} \hat{\pi}_\mu[b] = \hat{\pi}_0[b] = 1$. However, in other games we may have $\limsup_{\mu \to 0} \hat{\pi}_\mu[b] < 1$; in fact, $\limsup_{\mu\to 0} \hat{\pi}_\mu[C(b)] < 1$, where $C(b)$ is the equilibrium component of b. For example, consider the basic dynamic in the game[46] Γ_3 of Fig. 3. All the equilibria in the component of b—i.e., all $x = (\xi^1, \xi^2, \xi^3)$ with $\xi^1 = 0$ and $\xi^2 \leqslant 4/5$, where ξ^i denotes the proportion of c^i—have positive $\hat{\pi}_0$, since they are reached by a chain of single mutations from b (this follows by repeatedly applying (3.3), similar to the arguments for the game Γ_1' given above). In particular,[47] $\hat{\pi}_0[(0, 4/5, 1)] > 0$. Now $(1, 1, 4/5)$ is one-mutation-reachable from $(0, 4/5, 1)$:

$$\begin{aligned}\left(0, \frac{4}{5}, 1\right) &\xrightarrow{MU(2)} \left(0, \frac{4}{5} + \frac{1}{m}, 1\right) \xrightarrow{SE(1,3)} \left(\frac{1}{m}, \frac{4}{5} + \frac{1}{m}, 1 - \frac{1}{m}\right)\\ &\xrightarrow{SE(1,2,3)} \left(\frac{2}{m}, \frac{4}{5} + \frac{2}{m}, 1 - \frac{2}{m}\right) \cdots \xrightarrow{SE(1,2,3)} \left(\frac{1}{5}, 1, \frac{4}{5}\right)\\ &\xrightarrow{SE(1)} \left(\frac{1}{5} + \frac{1}{m}, 1, \frac{4}{5}\right) \cdots \xrightarrow{SE(1)} \left(1, 1, \frac{4}{5}\right).\end{aligned}$$

---

[43] We write m for the size of population $M(2)$. To streamline the argument, we are ignoring the distinction between a state ω and its corresponding pair of mixed actions $x = x(\omega)$ (given by (2.2)).

[44] In fact, the same argument then yields $\hat{\pi}_0[(0,0)] > 0$ (start with a mutation in population 1), and then $\hat{\pi}_0[(0,1)] > 0$ (a mutation in population 2); thus $\hat{\pi}_0[x] > 0$ for all absorbing profiles x.

[45] If the converse also holds, then the selection is called *Nash-compatible* in Samuelson (1997), (7.8).

[46] The game Γ_2 of Fig. 2 is discussed in Section 5.3 (a).

[47] Assume for simplicity that all populations are of size m which is a multiple of 5.

Therefore the equilibrium $(1, 1, 4/5)$—which is in the component of c—has positive $\hat{\pi}_0$ (and thus so do all the other equilibria there).[48]

3.2. The main result

As the examples show, when the populations are fixed, equilibria other than the backward induction equilibrium b (including some that are very different from b) may be evolutionarily stable. We now consider the case where the populations increase, i.e., $|M(i)| \to \infty$ for $i \in N$. Put $\mathbf{m} = (|M(i)|)_{i\in N}$ for the vector of population sizes; we will refer to $\mathbf{m}$ as the *population profile*. As $\mathbf{m} \to \infty$, the state space changes and becomes infinite in the limit; we need therefore[49] to consider (arbitrary small) neighborhoods of $BI = \{b\}$ in the set of mixed actions X: For every $\varepsilon > 0$, put $BI_\varepsilon := \{x \in X\colon\ x^i_{b^i} \geqslant 1 - \varepsilon$ for all $i \in N\}$. That is, x belongs to the ε-neighborhood BI_ε of b if in all populations most of the individuals play their backward induction action; we emphasize that this holds for *all* $i \in N$—whether node i is reached or not.

Theorem 3.2 (Main Theorem). *For every mutation rate $\mu > 0$ and population profile $\mathbf{m} = (|M(i)|)_{i\in N}$, let $\pi_{\mu,\mathbf{m}}$ be the unique invariant distribution of a dynamic process given by a one-step transition matrix $Q \equiv Q_{\mu,\mathbf{m}}$ satisfying* (2.4)–(2.9). *Then, for every $\varepsilon > 0$ and $\delta > 0$,*

$$\lim_{\substack{\mu\to 0,\ m\to\infty \\ \mu m\geqslant\delta}} \hat{\pi}_{\mu,\mathbf{m}}[BI_\varepsilon] = 1,$$

where $m := \min_{i\in N} |M(i)|$. Moreover, there exists a constant c, depending on the game, on the dynamics parameters α_1, α_2, β, γ_1, γ_2, and on ε, δ, such that[50]

$$E_{\mu,\mathbf{m}}\big[x^i_{b^i}(\omega)\big] \geqslant 1 - c\mu \quad \text{for all } i \in N, \quad \text{and} \tag{3.5}$$

$$\pi_{\mu,\mathbf{m}}\big[x^i_{b^i}(\omega) \geqslant 1 - \varepsilon \text{ for all } i \in N\big] \geqslant 1 - c\mu, \tag{3.6}$$

for all $\mu > 0$ and all $\mathbf{m} = (|M(i)|)_{i\in N}$ with $|M(i)| \geqslant \delta/\mu$ for all $i \in N$.

Thus, as the mutation rate is low and the populations are large, the proportion of each population i that does not play the backward induction action is small. Hence, in the long run, the dynamic system is most of the time in states where almost every individual plays his backward induction action.

---

[48] Computations (using MAPLE) for $m = 1, 2, \ldots, 8$ yield the following (approximate) values for $\hat{\pi}_0[C(b)]$, respectively: 0.89, 0.83, 0.821, 0.825, 0.8302, 0.8333, 0.8391, 0.8444.

[49] The probability of a single point may become 0 in the limit. For example, if $1/m$ is much smaller than μ (i.e., if $\mu m \to \infty$), then we may well get $\hat{\pi}[b] \to 0$. Indeed, consider the simplest case of a one-person game. The transition probability from the state ω_0 where everyone plays b, to any state ω_1 where all but one individual play b, is of the order of μ, whereas the transition from ω_1 to ω_0 has probability of the order of $1/m$.

[50] $E_{\mu,\mathbf{m}}$ denotes the expectation with respect to the probability distribution $\pi_{\mu,\mathbf{m}}$.

Remarks.

1. The only assumption made on the relative rates of convergence of μ and m is that $\mu m \geqslant \delta > 0$, i.e., that μm is bounded away from 0. It follows in particular that $\lim_{\mu \to 0} \lim_{m \to \infty} \hat{\pi}_{\mu,\mathbf{m}}[BI_\varepsilon] = 1$ (but this need not hold for the other iterative limit $\lim_{m \to \infty} \lim_{\mu \to 0}$). To interpret the condition on μm, define a "generation" to be that period of time in which each individual gets one opportunity to change his action; in our model, it is about m stages. Then the requirement is that the expected number of mutations per generation be bounded away from 0. (See also (e) in Section 5.2.)
2. No assumptions are made on the relative population sizes $|M(i)|$; one population may well be much larger than another.[51] However, the mutation rates in the different populations are assumed to be of the same order of magnitude—see (2.8) and (2.9).
3. The estimates we get in (3.5) and (3.6) involve the mutation rate μ but no higher powers of μ (as is the case in much of the existing literature, in particular in evolutionary dynamics for games in strategic form). This means that the effect of *simultaneous mutations* (whose probability—a power of μ—is relatively small) may indeed be ignored. Thus our result does not rely on the fact that, when μ is small, 100 simultaneous mutations are much more probable than 101 simultaneous mutations (both of these events are extremely improbable).[52] Our proof therefore does not use any of the techniques based on "counting mutations."

4. Proof of the Main Theorem

4.1. An informal argument

We begin by presenting informally the main ideas of the proof of our result; in particular, we explain the role of large populations. We do so for the simpler basic model; similar arguments apply to the general model.[53]

---

[51] However, notice that we have assumed that changes occur in the various populations with similar frequencies (e.g., in the basic model, one mutation per $1/\mu$ periods). If the populations are significantly different, one may want to modify this assumption: For example, if one population is 100 times larger than another, then one change in the latter corresponds to 100 changes in the former. As long as the populations are comparable (i.e., the ratios $|M(i)|/|M(j)|$ are bounded away from 0 and ∞), this modification will not affect the results.

[52] In a sense, the comparison here is between different coefficients of μ (i.e., of μ to the power 1), rather than between the first powers of μ with non-zero coefficient.

[53] The general model allows the effect of selection to be much weaker; for instance, if no individual currently plays the best-reply action, then the probability of switching to it may be only of the order of μ rather than σ. However, as we will show in Lemma 4.6, most of the time the proportion of

Clearly, if a node i is reached (i.e., if at every node along the path from the root to i there are individuals playing the action that corresponds to the path), then mutation $MU(i)$ at i has probability of the order of μ, which is much smaller[54] than the probability of selection $SE(i)$ at i. Therefore, at reached nodes most of the individuals play their best-reply actions. The problem is how to obtain a similar conclusion at the *unreached* nodes.

Consider first the three-player game Γ_2 of Fig. 2 in the Introduction. Assume that the dynamic system is in a state where all individuals at nodes 1 and 2 play b^1 and b^2, respectively; thus node 3 is not reached. Then selection $SE(3)$ does not affect the population at 3; only mutation $MU(3)$ does. Mutation by itself will lead in the long run to a distribution close to $(1/2, 1/2)$ (since each individual is eventually chosen, and then his action is replaced with equal probabilities by c^3 or b^3). However, there are also *mutations at node* 2 that yield a c^2 action, with a frequency of $\mu/2$. After such a mutation, the probability that the action of the mutant individual will revert to b^2 is at most $(1/m)\rho$ (since his probability of being chosen is $1/m$; here $\rho = \sigma + \mu/2$); thus, it will take on the average[55] m/ρ periods for it to happen. Therefore, over a long stretch of time, say T periods, the number of periods that there is a c^2 in population 2 is about[56] $(\mu/2)(m/\rho)T = \mu mT/(2\rho)$. These are periods at which 3 is reached and thus selection $SE(3)$—into b^3—is effective. At the same time, mutation $MU(3)$ occurs at 3 in roughly μT periods. Comparing the two implies that, when the population is large (i.e., as $m \to \infty$), selection has a much greater effect than mutation. Therefore, in the long run, we will get most of the population at node 3 playing b^3—even though 3 may be unreached most of the time.

Consider next the four-person game Γ_4 of Fig. 4. Assume again that everyone plays b^i at nodes 1 and 2, and thus both 3 and 4 are not reached. In the same way as in the previous example, we get the following: At node 4, mutations $MU(4)$ occur with a frequency of μ, whereas selection $SE(4)$ there—which requires mutations at *both* nodes 2 and 3 in order for 4 to be reached—occurs with a frequency of $(\mu/2)^2 m/(2\rho) = \mu^2 m/(8\rho)$ (indeed, the probability of a mutation at 2 into c^2 is $\mu/2$; the same goes for a mutation into c^3 at 3; and then it takes about $(m/\rho)/2 = m/(2\rho)$ periods until at least one of the mutants reverts). But we

---

individuals who play b^i is bounded away from 0, and thus the actual effect of selection towards b^i is in fact similar to that of the basic model. The difficulties in obtaining our result do *not* lie in the introduction of b^i and in keeping its proportion positive (mutations ensure that), but rather in making this proportion close to 1.

[54] We take the term "f is much smaller than g" to mean that the ratio f/g goes to 0 as $\mu \to 0$ and $m \to \infty$.

[55] An event whose probability is p every period will occur on average pT times during T periods, or once every $1/p$ periods. In our arguments we shall go back and forth between the two computations as needed.

[56] More precisely, $\min\{\mu mT/(2\rho), T\}$ (when $\mu/2 \geqslant \rho/m$, changes from b^2 into c^2 are more frequent than those from c^2 into b^2, and thus there will almost always be c^2-individuals).

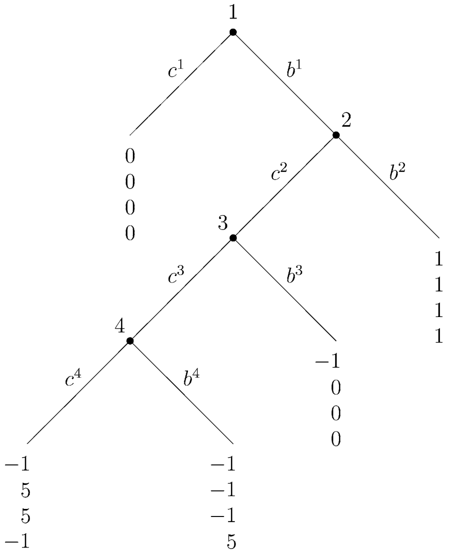

Fig. 4. The game Γ_4.

cannot say that $\mu^2 m$ is much larger than μ (we only assumed that μm is bounded away from zero), so we cannot conclude that, at node 4, selection "overpowers" mutation. Without this happening at 4, there is no reason for the populations at the higher nodes (like 3) to choose their backward induction action either. Moreover, when a node is even further away from the equilibrium path—say, k nodes away—the previous argument will work out only if $\mu^k m$ is much larger than μ.

A more careful analysis is thus called upon at this point.

Let us consider the first time that there is a c^3-individual in population 3; this happens (by mutation) on the average once every $2/\mu$ periods. If, at that point, there is a c^2 action in population 2, then 4 is reached and we are done. If not, then, as long as there is no c^2 in population 2, node 3 is not reached. Therefore, the only way that the c^3-individual can revert to b^3 is by mutation at 3; the probability of that happening is $(1/m)(\mu/2) = \mu/(2m)$ (since that *specific* individual must be chosen out of $M(3)$, and then undergo mutation). At the same time, the probability of getting a c^2-individual in population 2—by mutation—is $\mu/2$. Since $\mu/2$ is much larger than $\mu/(2m)$ for large m, in general the latter will happen much later than the former (and therefore can be essentially ignored). Thus altogether we have to wait at most on the order of $2/\mu$ periods for the mutation at 3, and then another $2/\mu$ periods for the mutation at node 2; in sum, $4/\mu$ periods until node 4 is reached (compare this to $4/\mu^2$ in the previous—unsuccessful—argument). Once 4 is reached, it takes on the order of $m/(2\rho)$ periods for either the c^2- or the c^3-individual to revert, so selection $SE(4)$ operates at node 4 with a frequency of approximately $(\mu/4)m/(2\rho) = \mu m/(8\rho)$. When $m \to \infty$, this is much larger than the mutation rate μ; so, again, selection "wins" at[57] 4. Once this has been established, it follows that most of the population at node 4 plays b^4 most of the time, and we are essentially[58] left with a three-player game (like Γ_2); the proof is completed by (backward) induction.[59]

The crux of the argument is that, after a mutation in population 3 has generated a c^3, this c^3 action is "stuck" there for a long time[60]—at least until a mutation generates a c^2 in population 2. Thus, what appeared to require *simultaneous mutations* (with a frequency of the order of μ^k for some $k \geqslant 2$), turns out instead to rely on *sequential mutations* (with a frequency of the order of μ).[61]

It should now be clear what role the large populations play: The smaller a group of individuals is, the (relatively) less probable it is that a change of action will occur in that group. This is particularly true when comparing a *specific* individual (like the c^3-mutant in the analysis of game Γ_4, or the c^2-mutant in Γ_2), to *any* individual in a whole population (population 3 in Γ_4, or population 2 in Γ_2).

Finally, to understand the use of the condition that $\mu m \geqslant \delta > 0$, note that the above arguments show that the effect of selection is of the order of μm, whereas that of mutation is μ. The possibility that a non-negligible fraction of the population does not play the backward induction action, albeit an event of low probability for large m, cannot be ignored. A simple estimate[62] shows this probability to be of the order of at most $1/m$. When this event occurs at some descendant node, it may affect selection at the current node—away from the backward induction action. However, as long as $1/m$ is at most a constant times μ—which is the case when $\mu m \geqslant \delta > 0$—this effect, like that of random mutations, is again small relative to μm for large m.

---

[57] This argument clearly generalizes; for a node that is at distance k from the equilibrium path, the frequency of selection is of the order of $\mu m/(2k^2\rho)$, which, as $m \to \infty$, is much larger than the mutation rate μ.

[58] See the last paragraph in this subsection.

[59] Similar arguments apply to the three-player game Γ_3 of Fig. 3. When less than 90% of population 3 plays b^3, in order for 3 to be reached one needs mutants at 2 and at 1, and the computations are exactly as for node 4 in Γ_4. When the 90% proportion of b^3 is exceeded, then b^2 becomes a best reply of 2, so a mutant is needed only at 1—and the estimates are as for node 3 in Γ_2.

[60] It is a so-called "neutral mutation" that does not affect the payoffs.

[61] This kind of argument may also explain how matching mutations occur in interacting populations, i.e., mutations that yield no advantage to their own population, unless there are compatible mutations in the other populations. The computations above show that, in large populations, the frequency of such events may well be much higher than commonly thought: of the order of μ rather than a power of μ.

[62] Using Markov's inequality. More refined probabilistic computations may well lead to weaker conditions. See also (e) in Section 5.2.

4.2. *An outline of the proof*

We now provide an outline that may help the reader to follow the formal proof given in the next subsection. The proof proceeds by backward induction, starting from the final nodes and working back towards the root; see Proposition 4.5. The main claims that are proved for each node i are as follows:

1. The probability that b^i is not the local best reply of i is low (this happens only when a sizeable proportion of the population at some descendant node j does not choose b^j, which, by induction, has low probability); in fact, this probability is of the order of μ; see (4.4).
2. When b^i is the local best reply of i, the expected proportion of population i that does not play b^i when i is reached is small; again, it is of the order of μ (this holds since, with high probability, selection towards b^i has probability which is bounded away from 0, while mutation has probability μ); see Lemma 4.6 and Eq. (4.6).
3. When b^i is the local best reply of i, the ratio between the expected proportion of population i that does not play b^i when i is reached, and the same expected proportion when i is not reached, is of the order of μm; see (4.13). This is the central step in the proof, and its essence is the "sequential mutations" argument above; see Proposition 4.1. Together with the previous claim, it follows that the expected proportion of population i that does not play b^i when i is not reached is of the order of $1/m$; see (4.10).
4. Adding the above three estimates and noting that $1/m \leqslant (1/\delta)\mu$ implies that the expected proportion of population i that does not play b^i is of the order of μ, which yields (3.5) and (3.6).

4.3. *The proof*

We now prove the Main Theorem. Fix $\varepsilon, \delta > 0$; the mutation rate $\mu > 0$; the population profile $\mathbf{m} = (m^i)_{i \in N}$ with $\mu m \geqslant \delta$, where $m^i := |M(i)|$ and $m := \min_{i \in N} m^i$; and the transition probability matrix Q that satisfies (2.4)–(2.9). Let π be the resulting unique invariant distribution over the state space Ω. Take the state $\omega \in \Omega$ to be distributed according to π, and let $\tilde{\omega} \in \Omega$ be the next state, given by the one-step transition probabilities $Q[\tilde{\omega} \mid \omega]$; then $\tilde{\omega}$ is also distributed according to[63] π. From now on all probability statements and expectations will be according to this distribution.

---

[63] In other words: $(\omega, \tilde{\omega}) \in \Omega \times \Omega$ is distributed as two subsequent states of the Markov process, and ω is distributed according to the invariant distribution π (and thus so is $\tilde{\omega}$).

Before proceeding with the proof, we introduce a number of useful notations:

- For each node $i \in N$, put $Y^i := 1 - x^i_{b^i}(\omega)$ (the mapping $x : \Omega \to X$ is given by (2.2)); this is the proportion of population i that does *not* play the backward induction action in state ω. Similarly, put $\widetilde{Y}^i := 1 - x^i_{b^i}(\tilde{\omega})$ for the same proportion in the next-period state $\tilde{\omega}$. The random variables Y^i and $\widetilde{Y}^i$ are identically distributed (their distribution is $\pi \circ (1 - x^i_{b^i})^{-1}$); thus in particular $E[\widetilde{Y}^i] = E[Y^i]$.
- Given two nodes $i, j \in N$ such that i is a descendant of j (i.e., $i \in N(j)$), let $R^{j,i}$ be an indicator random variable, defined as 1 if node i is *reached* from node j in state ω, and 0 otherwise; i.e., $R^{j,i} = 1$ if and only if for every $k \in N$ on the path from j to i there is at least one individual $q \in M(k)$ whose choice ω^k_q is the action that leads towards i. When j is the root we will write R^i for the indicator that i is reached. Again, $\widetilde{R}^{j,i}$ is defined in the same way for $\tilde{\omega}$.
- When everyone plays the backward induction action—i.e., when $Y^j = 0$ for all $j \in N$—the unique local best reply for each $i \in N$ is b^i (recall that b is the unique backward induction equilibrium). Therefore there exists a $\lambda > 0$ (appropriately small) such that b^i is the unique local best reply of i for all $i \in N$ when $Y^j < \lambda$ for all $j \in N$ (i.e., when the proportion of the individuals at each node that do *not* play the backward induction action is less than λ). This λ depends on the game only, and will be fixed from now on.
- Let L^i be an indicator random variable, defined as 1 in state ω if $Y^j < \lambda$ for all $j \in N(i)$, and 0 otherwise. Thus, when $L^i = 1$ the backward induction action b^i is the unique *local* best reply of i in state ω, i.e., $u^i_{\Gamma(i)}(b^i, \omega^{N(i)}) > u^i_{\Gamma(i)}(a^i, \omega^{N(i)})$ for every $a^i \in A^i$, $a^i \neq b^i$. We denote by $\widetilde{L}^i$ the indicator that $\widetilde{Y}^j < \lambda$ for all $j \in N(i)$. When i is a final node (i.e., when $N(i) = \emptyset$) we have $L^i \equiv \widetilde{L}^i \equiv 1$.

We note that selection $SE(i)$ has an effect only when i is reached, i.e., when $R^i = 1$; if i is not reached, i.e., if $R^i = 0$, then all actions of i yield the same payoff in Γ and only mutation $MU(i)$ affects ω^i. If $R^i = 1$ and $L^i = 1$ then b^i is the global best reply of i, and thus certainly a "better action" for a "non-b^i-individual" (i.e., $b^i \in B^i(q(i), \omega)$ when $\omega^i_{q(i)} \neq b^i$). Since these arguments will be used repeatedly in the proof, for convenience we state the following implications of (2.4)–(2.9) here[64]

$$P\left[\tilde{\omega}^i_{q(i)} = a^i \;\middle|\; \omega\right] \geqslant \alpha_1 \mu \quad \text{for every } a^i \in A^i. \tag{4.1}$$

---

[64] We use the "big-O" notation: $f(x) = O(g(x))$ if there exists a constant $K < \infty$ such that $|f(x)| \leqslant K|g(x)|$ for all x in the relevant region. Thus $f(\mu, \mathbf{m}) = O(\mu)$ means that there exists K such that $|f(\mu, \mathbf{m})| \leqslant K\mu$ for all $0 < \mu < 1$ and all vectors $\mathbf{m} = (m^i)_{i \in N}$ with integer coordinates $m^i \geqslant \delta/\mu$ for all $i \in N$.

$$\text{If } R^i = 0 \text{ then } P\big[\tilde{\omega}^i_{q(i)} \neq \omega^i_{q(i)} \mid \omega\big] = O(\mu). \tag{4.2}$$

$$\text{If } L^i R^i = 1 \text{ and } \omega^i_{q(i)} \neq b^i \text{ then } P\big[\tilde{\omega}^i_{q(i)} = b^i \mid \omega\big] \geqslant \beta\big(1 - Y^i\big). \tag{4.3}$$

The crucial argument in our proof is the following proposition.

Proposition 4.1. *Consider the path from the root to a node $i \in N$; without loss of generality, assume that the nodes along this path are $1, 2, \ldots, i-1, i$ (in that order, with 1 the root). Let $\widetilde{Z}$ be a non-negative bounded random variable that depends on*[65] $(\tilde{\omega}^k)_{k \in \{i\} \cup N(i)}$. *Then*

$$\mu E\big[\widetilde{Z}\big(1 - R^i\big)\big] \leqslant O\left(\frac{1}{m}\right) E\big[\widetilde{Z} R^i\big] + O\left(\frac{\mu}{m}\right) + \sum_{j=1}^{i-1} O\big(E\big[\widetilde{Z}\big(1 - R^{j,i}\big)\big] - E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big)\big]\big).$$

Proof. For each $j = 1, 2, \ldots, i-1$, let $c^j \in A^j$ be the choice along the given path (i.e., towards $j+1$), and put $\theta^j(\omega) := |\{q \in M(j)\colon \omega^j_q = c^j\}|/m^j$, the proportion of individuals at node j that choose c^j; denote $V^j := \theta^j(\omega)$ and $\widetilde{V}^j := \theta^j(\tilde{\omega})$. We first prove three lemmata.

Lemma 4.2.

$$E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big) R^i\big] = O\left(\frac{1}{m}\right) E\big[\widetilde{Z} R^i\big].$$

Proof. For each ω with $R^i = 1$, to get $\widetilde{R}^{j,i} = 0$ there must be some node k with $j \leqslant k < i$ and $\widetilde{V}^k = 0$. But $V^k > 0$ (since $R^i = 1$), so in fact $V^k = 1/m^k$ (since $|\widetilde{V}^k - V^k|$ is 0 or $1/m^k$). Hence $P[\widetilde{V}^k = 0 \mid \omega] \leqslant \gamma_2/m^k = O(1/m)$ (by (2.5), since the single c^k-individual must be chosen). Therefore $P[\widetilde{R}^{j,i} = 0 \mid \omega] \leqslant \sum_{k=j}^{i-1} P[\widetilde{V}^k = 0 \mid \omega] = O(1/m)$, from which it follows that

$$E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big) \mid \omega\big] = P\big[\widetilde{R}^{j,i} = 0 \mid \omega\big] E\big[\widetilde{Z} \mid \omega\big] = O\left(\frac{1}{m}\right) E\big[\widetilde{Z} \mid \omega\big]$$

(we have used the conditional independence condition (2.4): $\widetilde{R}^{j,i}$ depends on nodes before i, whereas $\widetilde{Z}$ depends on i and nodes after i). Thus

$$E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big) \mid R^i = 1\big] = O\left(\frac{1}{m}\right) E\big[\widetilde{Z} \mid R^i = 1\big],$$

---

[65] We thank an anonymous referee for pointing out that an assumption on $\widetilde{Z}$ was missing here. It may be shown that the result also holds if $\widetilde{Z}$ is replaced by Z which depends on $\omega = (\omega^k)_{k \in N}$ (note: no restrictions here on which coordinates).

and

$$\begin{aligned} E\big[\widetilde{Z}\big(1-\widetilde{R}^{j,i}\big)R^i\big] &= E\big[\widetilde{Z}\big(1-\widetilde{R}^{j,i}\big) \bigm| R^i=1\big]P\big[R^i=1\big]+0P\big[R^i=0\big] \\ &= O\left(\frac{1}{m}\right)E\big[\widetilde{Z}\mid R^i=1\big]P\big[R^i=1\big] \\ &= O\left(\frac{1}{m}\right)E\big[\widetilde{Z}R^i\big]. \qquad \square \end{aligned}$$

Lemma 4.3.

$$E\big[\widetilde{Z}\big(1-\widetilde{R}^{j,i}\big)\big(1-R^j\big)R^{j,i}\big]=O\left(\frac{\mu}{m}\right).$$

Proof. For each $\omega \in \Omega$ with $R^j = 0$ and $R^{j,i} = 1$ (i.e., j is not reached, and i is reached from j), to get $\widetilde{R}^{j,i} = 0$ there must be some node k with $j \leqslant k < i$ and $\widetilde{V}^k = 0$. But $R^{j,i} = 1$ implies that $V^k > 0$, and thus $V^k = 1/m^k$. Therefore $P[\widetilde{V}^k = 0 \mid \omega] \leqslant (\gamma_2/m^k)O(\mu) = O(\mu/m)$ (by (2.5) and (4.2): the single c^k-individual must be chosen, and its action can change by mutation only, since j, and thus *a fortiori* k, is not reached). Hence $P[\widetilde{R}^{j,i} = 0 \mid \omega] \leqslant \sum_{k=j}^{i-1} P[\widetilde{V}^k = 0|\omega] = O(\mu/m)$, from which the result follows. □

Lemma 4.4.

$$\begin{aligned} &E\big[\widetilde{Z}\big(1-\widetilde{R}^{j,i}\big)\big(1-R^{j,j+1}\big)R^{j+1,i}\big] \\ &\quad \leqslant (1-\alpha_1\mu)E\big[\widetilde{Z}\big(1-R^{j,j+1}\big)R^{j+1,i}\big]+O\left(\frac{\mu}{m}\right). \end{aligned}$$

Proof. Take $\omega \in \Omega$ with $R^{j,j+1} = 0$ (i.e., $V^j = 0$), and $R^{j+1,i} = 1$. To get $\widetilde{R}^{j,i} = 0$ there must be some node k with $j \leqslant k < i$ and $\widetilde{V}^k = 0$. Now $P[\widetilde{V}^j > 0|\omega] = P[\tilde{\omega}^j_{q(j)} = c^j \mid \omega] \geqslant \alpha_1\mu$ (this follows from $V^j = 0$ and (4.1)), and $P[\widetilde{V}^k = 0 \mid \omega] \leqslant \gamma_2/m^k \leqslant \gamma_2/m$ for $k = j+1, \ldots, i-1$ (by (2.5), since $R^{j+1,i} = 1$ implies $V^k > 0$). Therefore, by (2.4),

$$\begin{aligned} P\big[\widetilde{R}^{j,i}=1 \bigm| \omega\big] &= P\big[\widetilde{V}^k>0 \text{ for all } j\leqslant k<i-1 \bigm| \omega\big] \\ &\geqslant (\alpha_1\mu)(1-\gamma_2/m)^{i-j-1} \end{aligned}$$

hence $P[\widetilde{R}^{j,i} = 0 \mid \omega] \leqslant (1-\alpha_1\mu) + O(\mu/m)$, and the result follows as in the proof of Lemma 4.2. □

The proof of Proposition 4.1 can now be completed.

Proof of Proposition 4.1 (continued). We have $(1 - R^j)R^{j,i} = R^{j,i} - R^i$ and $(1 - R^{j,j+1})R^{j+1,i} = R^{j+1,i} - R^{j,i}$. Adding the inequalities of Lemmata 4.2–4.4 together with

$$E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big)\big(1 - R^{j+1,i}\big)\big] \leqslant E\big[\widetilde{Z}\big(1 - R^{j+1,i}\big)\big]$$

yields

$$\begin{aligned} E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big)\big] \leqslant\ & O\left(\frac{1}{m}\right)E\big[\widetilde{Z}R^i\big] + (1 - \alpha_1\mu)E\big[\widetilde{Z}\big(R^{j+1,i} - R^{j,i}\big)\big] \\ & + E\big[\widetilde{Z}\big(1 - R^{j+1,i}\big)\big] + O\left(\frac{\mu}{m}\right). \end{aligned}$$

Rearranging terms gives

$$\begin{aligned} \alpha_1\mu E\big[\widetilde{Z}\big(R^{j+1,i} - R^{j,i}\big)\big] \leqslant\ & O\left(\frac{1}{m}\right)E\big[\widetilde{Z}R^i\big] + O\left(\frac{\mu}{m}\right) \\ & + \big(E\big[\widetilde{Z}\big(1 - R^{j,i}\big)\big] - E\big[\widetilde{Z}\big(1 - \widetilde{R}^{j,i}\big)\big]\big). \end{aligned}$$

Adding these inequalities for $j = 1, 2, \ldots, i - 1$ and noting that $R^{i,i} = 1$ and $R^{1,i} = R^i$ completes the proof. □

The next proposition proves the Main Theorem. The argument is divided into 8 steps.

Proposition 4.5. *For each node $i \in N$:*

$$P\big[L^i = 0\big] = O(\mu); \tag{4.4}$$

$$P\big[\widetilde{Y}^i < Y^i\big] = P\big[\widetilde{Y}^i > Y^i\big] = O(\mu); \tag{4.5}$$

$$E\big[Y^i L^i R^i\big] = O(\mu); \tag{4.6}$$

$$E\big[\widetilde{Y}^i \widetilde{L}^i R^i\big] = O(\mu); \tag{4.7}$$

$$E\big[\big|\widetilde{Y}^i \widetilde{L}^i - Y^i L^i\big|\big(1 - R^i\big)\big] = O\left(\frac{\mu}{m}\right); \tag{4.8}$$

$$E\big[\widetilde{Y}^i \widetilde{L}^i\big(1 - R^i\big)\big] = O\left(\frac{1}{m}\right); \tag{4.9}$$

$$E\big[Y^i L^i\big(1 - R^i\big)\big] = O\left(\frac{1}{m}\right); \tag{4.10}$$

$$E\big[Y^i\big] = O(\mu). \tag{4.11}$$

Proof. The proof is by backward induction on i. Assume that (4.4)–(4.11) hold for all[66] $j \in N(i)$; then each of the claims (4.4)–(4.11) for i will be proved in turn.

---

[66] The induction starts from final nodes i for which $N(i) = \emptyset$ (and thus there is no assumption).

Step 1: (4.4) *holds for* i. Indeed,[67] $L^i = 0$ implies that there is $j \in N(i)$ such that $Y^j \geqslant \lambda$, hence

$$P\big[L^i = 0\big] \leqslant \sum_{j\in N(i)} P\big[Y^j \geqslant \lambda\big] \leqslant \sum_{j\in N(i)} \frac{1}{\lambda} E\big[Y^j\big] = O(\mu),$$

where we have used Markov's inequality[68] and (4.11) for j (by the induction hypothesis).

Step 2: (4.5) *holds for* i. We have $E[\widetilde{Y}^i] = E[Y^i]$ (recall that π is the invariant distribution), thus

$$0 = E\big[\widetilde{Y}^i - Y^i\big] = \left(\frac{1}{m^i}\right) P\big[\widetilde{Y}^i > Y^i\big] + \left(-\frac{1}{m^i}\right) P\big[\widetilde{Y}^i < Y^i\big],$$

since the only possible values of $\widetilde{Y}^i - Y^i$ are $0,\ 1/m^i$, and $-1/m^i$. Therefore

$$P\big[\widetilde{Y}^i > Y^i\big] = P\big[\widetilde{Y}^i < Y^i\big].$$

To get $\widetilde{Y}^i > Y^i$ we need a b^i-individual to become non-b^i (i.e., $\omega^i_{q(i)} = b^i$ and $\tilde{\omega}^i_{q(i)} \neq b^i$). This can happen either by selection—which requires b^i not to be a best reply of i (hence $L^i = 0$)—or by mutation—with probability equal to $O(\mu)$ (by (2.9)). Thus

$$P\big[\widetilde{Y}^i > Y^i\big] \leqslant P\big[L^i = 0\big] + O(\mu) = O(\mu),$$

by (4.4) for i, proving (4.5) for i.

Step 3: (4.6) *holds for* i. The case $\widetilde{Y}^i < Y^i$ occurs when a non-b^i-action is replaced by b^i; thus the chosen individual $q(i) \in M(i)$ is a non-b^i-individual (i.e., $\omega^i_{q(i)} \neq b^i$), which happens with probability $\geqslant \gamma_1 Y^i$ by (2.5). For every $\omega \in \Omega$ with $L^i R^i = 1$, the probability $P[\tilde{\omega}^i_{q(i)} = b^i | \omega]$ of changing the action to b^i is at least $\beta(1 - Y^i)$ by (2.7) or (4.3) (since the proportion of b^i in the population is $1 - Y^i$). Therefore

$$\begin{aligned} P\big[\widetilde{Y}^i < Y^i\big] &\geqslant E\big[\beta\gamma_1 Y^i\big(1 - Y^i\big) \bigm| L^i R^i = 1\big] P\big[L^i R^i = 1\big] \\ &\quad + 0\, P\big[L^i R^i \neq 1\big] \\ &= \beta\gamma_1 E\big[Y^i\big(1 - Y^i\big) L^i R^i\big], \end{aligned}$$

which, by (4.5) for i, implies

$$E\big[Y^i\big(1 - Y^i\big) L^i R^i\big] = O(\mu),$$

---

[67] We thank Michihiro Kandori for pointing out an error in this proof in the first version of the paper.

[68] Markov's inequality is: $P[Z \geqslant z] \leqslant (1/z)E[Z]$ for a non-negative random variable Z and $z > 0$.

and thus

$$E\big[Y^i L^i R^i \mathbf{1}_{Y^i \leqslant 1-\eta}\big] = O(\mu)/\eta = O(\mu) \tag{4.12}$$

for any $\eta > 0$, where $\mathbf{1}_{Y^i \leqslant 1-\eta}$ is the indicator that $Y^i \leqslant 1-\eta$.

To complete the proof of Step 3 we need the following lemma, showing that the probability that Y^i is in a neighborhood of 1 is low; that is, most of the time the proportion of b^i is bounded away from[69] 0.

Lemma 4.6. *There exist constants $\eta > 0$ and $c > 0$ such that*

$$P\big[Y^i > 1-\eta\big] = O\big(e^{-cm}\big).$$

Proof. The invariant distribution property implies[70]

$$\begin{aligned} &P\left[Y^i = \frac{k+1}{m^i}\right] P\left[\widetilde{Y}^i = \frac{k}{m^i} \,\middle|\, Y^i = \frac{k+1}{m^i}\right] \\ &\quad = P\left[Y^i = \frac{k}{m^i}\right] P\left[\widetilde{Y}^i = \frac{k+1}{m^i} \,\middle|\, Y^i = \frac{k}{m^i}\right] \end{aligned}$$

for every $k = 0, 1, \ldots, m^i - 1$. We have, first,

$$P\left[\widetilde{Y}^i = \frac{k}{m^i} \,\middle|\, Y^i = \frac{k+1}{m^i}\right] \geqslant \gamma_1 \frac{k+1}{m^i} \alpha_1 \mu$$

by (2.5) and (2.8); and second,

$$\begin{aligned} P\left[\widetilde{Y}^i = \frac{k+1}{m^i} \,\middle|\, Y^i = \frac{k}{m^i}\right] &\leqslant \gamma_2 \left(1 - \frac{k}{m^i}\right)\left(\alpha_2 \mu + P\big[L^i = 0\big]\right) \\ &\leqslant c_1 \left(1 - \frac{k}{m^i}\right)\mu \end{aligned}$$

for an appropriate constant $c_1 > 0$, where we have used (2.5) and (2.9) (selection can increase Y^i only when $L^i = 0$), and then $P[L^i = 0] = O(\mu)$ by (4.4) for i (proved in Step 1). Therefore

$$P\left[Y^i = \frac{k+1}{m^i}\right] \gamma_1 \frac{k+1}{m^i} \alpha_1 \mu \leqslant P\left[Y^i = \frac{k}{m^i}\right] c_1 \left(1 - \frac{k}{m^i}\right)\mu,$$

---

[69] This shows that in our model with large populations, (2.7) turns out to be effectively equivalent to the stronger (3.4).

[70] Let π be an invariant distribution of a Markov chain Q with state space Ω which is partitioned into two disjoint sets S and T, then it can be checked that $\sum_{s \in S} \pi[s] Q[T \mid s] = \sum_{t \in T} \pi[t] Q[S \mid t]$ (the "total flow"—i.e., the total invariant probability—from S to T equals the "total flow" from T to S). In our case, take $S = \{Y^i \geqslant (k+1)/m^i\}$ and $T = \{Y^i \leqslant k/m^i\}$.

or

$$P\left[Y^i = \frac{k+1}{m^i}\right] \leqslant c_2\left(\frac{m^i - k}{k+1}\right)P\left[Y^i = \frac{k}{m^i}\right],$$

where $c_2 := c_1/(\gamma_1\alpha_1)$. Let $\eta > 0$ be small enough so that $c_2(m^i - k)/(k+1) \leqslant 1/2$ for all[71] $k > k_0 := \lfloor(1-2\eta)m^i\rfloor$. Then we get

$$P\left[Y^i = \frac{k}{m^i}\right] \leqslant \left(\frac{1}{2}\right)^{k-k_0}$$

for all $k \geqslant k_0$ and thus

$$P\left[Y^i > 1-\eta\right] \leqslant \sum_{k>(1-\eta)m^i} \left(\frac{1}{2}\right)^{k-k_0} \leqslant \left(\frac{1}{2}\right)^{\eta m^i - 1},$$

as claimed. □

Proof of Step 3 (continued). Lemma 4.6 yields

$$E\left[Y^i L^i R^i \mathbf{1}_{Y^i > 1-\eta}\right] \leqslant P\left[Y^i > 1-\eta\right] = O\left(e^{-cm}\right),$$

which is at most $O(\mu)$ since $1/m \leqslant (1/\delta)\mu$. Adding this to the estimate of (4.12) gives (4.6) for i, thus completing Step 3.

Step 4: (4.7) *holds for* i. Write $E[\widetilde{Y}^i \widetilde{L}^i R^i] = E[\widetilde{Y}^i \widetilde{L}^i (1-L^i)R^i] + E[\widetilde{Y}^i \widetilde{L}^i L^i R^i]$. The first term is $O(\mu)$ by (4.4), and the second term is

$$\begin{aligned} E\left[\widetilde{Y}^i \widetilde{L}^i L^i R^i\right] &\leqslant E\left[\widetilde{Y}^i L^i R^i\right] = E\left[Y^i L^i R^i\right] + E\left[\left(\widetilde{Y}^i - Y^i\right)L^i R^i\right] \\ &\leqslant E\left[Y^i L^i R^i\right] + \left(\frac{1}{m^i}\right)P\left[\widetilde{Y}^i > Y^i\right] \end{aligned}$$

(since the only positive value of $\widetilde{Y}^i - Y^i$ is $1/m^i$). Applying (4.5) yields the desired inequality.

Step 5: (4.8) *holds for* i. We have

$$\begin{aligned} &E\left[\left|\widetilde{Y}^i \widetilde{L}^i - Y^i L^i\right|\left(1-R^i\right)\right] \\ &\quad\leqslant E\left[\left|\widetilde{Y}^i - Y^i\right|\widetilde{L}^i\left(1-R^i\right)\right] + E\left[Y^i\left|\widetilde{L}^i - L^i\right|\left(1-R^i\right)\right] \\ &\quad\leqslant E\left[\left|\widetilde{Y}^i - Y^i\right|\left(1-R^i\right)\right] + E\left[\left|\widetilde{L}^i - L^i\right|\left(1-R^i\right)\right]. \end{aligned}$$

The first term is bounded by $(1/m^i)P[\widetilde{Y}^i \neq Y^i,\ R^i = 0] = (1/m^i)\,O(\mu) = O(\mu/m)$ (see (4.2): $R^i = 0$ implies that the change from Y^i to $\widetilde{Y}^i$ is by mutation only). For the second term, note that $\widetilde{L}^i \neq L^i$ implies that there exists $j \in N(i)$ such that $Y^j \geqslant \lambda - 1/m^j$ (otherwise $Y^j < \lambda - 1/m^j$ and thus $\widetilde{Y}^j < \lambda$ for all

---

[71] $\lfloor x \rfloor$ denotes the largest integer that is $\leqslant x$.

$j \in N(i)$, in which case $L^i = \widetilde{L}^i = 1$). Choose $j \in N(i)$ to be a last such node; thus $Y^r < \lambda - 1/m^r$ for all $r \in N(j)$, hence $L^j = 1$. Now $\widetilde{L}^i \neq L^i$ implies that $\tilde{\omega}^k \neq \omega^k$ for some $k \in N(i) \cup \{i\}$, which, by (4.2), has conditional probability equal to $O(\mu)$ for every ω with $R^i = 0$ and thus *a fortiori* $R^k = 0$. Therefore

$$\begin{aligned}
&E\big[\big|\widetilde{L}^i - L^i\big|\big(1 - R^i\big)\big] \\
&\quad\leqslant \sum_{j \in N(i)} P\left[\widetilde{L}^i \neq L^i,\ Y^j \geqslant \lambda - \frac{1}{m^j},\ L^j = 1,\ R^i = 0\right] \\
&\quad\leqslant O(\mu) \sum_{j \in N(i)} P\left[Y^j \geqslant \lambda - \frac{1}{m^j},\ L^j = 1,\ R^i = 0\right] \\
&\quad\leqslant O(\mu) \sum_{j \in N(i)} \frac{1}{\lambda - \frac{1}{m^j}} E\big[Y^j L^j \big(1 - R^j\big)\big],
\end{aligned}$$

where we have used $1 - R^i \leqslant 1 - R^j$ together with Markov's inequality. Applying (4.10) for each $j \in N(i)$ completes the proof of (4.8) for i.

Step 6: (4.9) *holds for* i. Take $\widetilde{Z} = \widetilde{Y}^i \widetilde{L}^i$ in Proposition 4.1. We have

$$\begin{aligned}
&\big|E\big[\widetilde{Y}^i \widetilde{L}^i \big(1 - R^{j,i}\big)\big] - E\big[\widetilde{Y}^i \widetilde{L}^i \big(1 - \widetilde{R}^{j,i}\big)\big]\big| \\
&\quad= \big|E\big[\widetilde{Y}^i \widetilde{L}^i \big(1 - R^{j,i}\big)\big] - E\big[Y^i L^i \big(1 - R^{j,i}\big)\big]\big| \\
&\quad\leqslant E\big[\big|\widetilde{Y}^i \widetilde{L}^i - Y^i L^i\big|\big(1 - R^{j,i}\big)\big] \\
&\quad\leqslant E\big[\big|\widetilde{Y}^i \widetilde{L}^i - Y^i L^i\big|\big(1 - R^i\big)\big] = O\left(\frac{\mu}{m}\right),
\end{aligned}$$

where we have used the fact that π is the invariant distribution, the inequality $1 - R^{j,i} \leqslant 1 - R^i$ (since $R^{j,i} = 0$ implies $R^i = 0$), and finally (4.8) for i. Thus each one of the right-most terms in the inequality obtained from Proposition 4.1 is $O(\mu/m)$, and therefore

$$\mu E\big[\widetilde{Y}^i \widetilde{L}^i \big(1 - R^i\big)\big] \leqslant O\left(\frac{1}{m}\right) E\big[\widetilde{Y}^i \widetilde{L}^i R^i\big] + O\left(\frac{\mu}{m}\right). \tag{4.13}$$

Applying (4.7) for i completes the proof.

Step 7: (4.10) *holds for* i. It follows immediately from (4.8) and (4.9) for i.

Step 8: (4.11) *holds for* i. Adding (4.6) and (4.10) for i yields

$$E\big[Y^i L^i\big] = O(\mu) + O\left(\frac{1}{m}\right) = O(\mu),$$

since $1/m \leqslant (1/\delta)\mu$. Together with

$$E\big[Y^i \big(1 - L^i\big)\big] \leqslant P\big[L^i = 0\big] = O(\mu)$$

by (4.4) for i, the proof is completed. □

Proof of Theorem 3.2 (Main Theorem). Inequality (4.11) of Proposition 4.5 is precisely (3.5); applying Markov's inequality then yields (3.6). □

5. Discussion

It should be clear at this point that in order to single out the backward induction equilibrium, the mutation rate must not be too low or too high relative to the selection rate: If it is too low then the system can get "stuck" at other equilibria, and if it is too high the system may be a completely random system. Studying other dynamical models will hopefully help delineate in which cases the backward induction outcome obtains, and in which cases it does not.

5.1. *Other selection dynamics*

Various generalizations and modifications of our dynamics may be considered. Some will not affect the main result, while others may well invalidate it. We list here several:

(a) Replace (2.7) with

$$Q\big[\tilde{\omega}^i_{q(i)} = a^i \bigm| \omega\big] \geqslant f\big(x^i_{a^i}(\omega)\big),$$

where $f : [0,1] \to [0,1]$ is a monotonic decreasing function with $f(\xi) > 0$ for all $\xi > 0$. For instance, in an "imitation" model, the chosen individual may sample s other individuals in his own population (rather than just one)—in which case we can take $f(\xi) = \beta\xi^s$. It may be shown that our result continues to hold: The only change is in Step 3 of the proof of Proposition 4.5, where we now have $P[\widetilde{Y}^i < Y^i] \geqslant \gamma_1 E[Y^i f(1 - Y^i) L^i R^i]$, which implies

$$E\big[Y^i L^i R^i \mathbf{1}_{Y^i \leqslant 1-\eta}\big] = O(\mu)/f(\eta) = O(\mu);$$

thus (4.12) holds.

(b) Replace (2.7) with (3.4), i.e., *some* better action has positive probability of being chosen (rather than *each*). It seems that our result is still correct.[72]

(c) Assume that in each period one individual—from *all* populations altogether—may change its action. For example, in each period choose one individual at random from $\bigcup_{i\in N} M(i)$; or, when the populations are of different sizes, first choose $i \in N$ at random, and then choose $q(i)$ in $M(i)$ at random (only this

---

[72] Consider for simplicity a final node i; one shows first that the proportion of the worst action at i must be small, after which the same is proven for the second-worst, and so on. Note that the proportion of the best reply—which can change only by mutation—is bounded away from zero.

chosen individual may change its action in the current period). We conjecture that this modification will not affect the result.[73]

(d) Make the probability of selection depend on the *payoff differences*. For example,

$$Q\big[\tilde{\omega}^i_{q(i)} = a^i \mid \omega\big] = \beta x^i_{a^i}(\omega)\big[u^i\big(a^i, \omega^{-i}\big) - u^i\big(\omega^i_{q(i)}, \omega^{-i}\big)\big], \tag{5.1}$$

for every $a^i \in B^i$. That is, the probability that selection induces a switch to a better action is proportional to how much better that action is, as measured by the payoffs against the current populations at the other nodes. In contrast, (2.7) only takes into account the *sign* of this difference (which is positive if and only if $a^i \in B^i$). Equation (5.1) yields a dynamic in the well-known class of "replicator dynamics."[74] Again, there are many variants here; we conjecture that our result need not hold under[75] (5.1) (though it might hold if we instead use a positive power, strictly less than 1, of the payoff differences).

(e) Modify selection so that the proportion of the better actions will increase with a probability that is close to 1 (i.e., of the order of $1 - \alpha\mu$). Note that in our model if, say, action a^i is currently better than action c^i and $2/3$ of the population play a^i, then the probability that the proportion of a^i will increase (to $2/3 + 1/m$) is of the order of $1/3$ only ($=$ the probability that a c^i-individual will be chosen). Some preliminary work indicates that our result holds here too, at least in the case where there are exactly two actions at each node.

(f) Change the probability that each individual in $M(i)$ will be chosen, in such a way that it still goes to zero as the population size $|M(i)|$ goes to infinity, though without necessarily being of the order of $1/|M(i)|$ as in[76] (2.5).

(g) Specify various rules as to how the individuals are to be matched in each period to play the game (recall Remark 3 in Section 2.3); for instance, when all populations are of size m, we may arrange all individuals into m random matchings (rather than everyone playing against everyone else). Such extensions might overcome the difficulty that an individual needs to play against everyone in the other populations in a time period that converges to zero as the populations increase.

---

[73] The conditional independence condition (2.4), which is no longer true, has been used only in the proofs of Lemmata 4.2 and 4.4.

[74] Our selection may thus be called an "ordinal replicator dynamic." Note that (5.1) is continuous (with respect to changes in the other populations), whereas (2.7) is not.

[75] Large populations decrease the effect of selection, since the difference in payoffs due to the action of one individual at a previous node is small.

[76] One then needs to work with expressions like $E[g(Y^i)]$ instead of $E[Y^i]$, where $g(Y^i)$ is the probability that a non-b^i-individual is chosen when the proportion of such individuals in the population is Y^i.

5.2. Extensions

A number of directions for further study are:

(a) *Non-unique backward induction equilibrium*. Analyze the non-generic case where there is more than one subgame-perfect equilibrium—for instance, when some of the payoffs at the final nodes are equal. It seems that a subset of BI, at times a strict subset, is obtained.
(b) *Multiple agents and non-distinct populations*. Study games where players may move at more than one node, and thus the populations at different nodes are no longer distinct (as in the gene-normal form). Note that difficulties may arise since selection operates at the level of the player (rather than that of the agent/gene).
(c) *Games with imperfect information*. Allow for general games in extensive form (not necessarily of perfect information). Here we conjecture that all evolutionarily stable equilibria will be subgame-perfect, but the converse will no longer be true; evolutionary dynamics may well pick out certain refinements rather than others.
(d) *Other selection dynamics*. This is discussed in Section 5.1.
(e) $\mu m \to 0$ *and the order of limits*. Study the case where the number of mutations per generation may go to zero (see Remark 1 in Section 3.2). While our proof makes use of the condition $\mu m \geqslant \delta > 0$, we do not know whether it is always necessary.[77] It is thus possible that, for a certain class of dynamics (like the basic dynamic, or those that satisfy (3.4)), one has in fact $\lim_{m\to\infty,\ \mu\to 0} \hat{\pi}[BI_\varepsilon] = 1$ for the double limit (and thus also for the iterative limit $\lim_m \lim_\mu$).

5.3. Related literature

This subsection discusses relations to other work on dynamic models in games of extensive form (for games in strategic form, see the books referred to in Section 1.1).

(a) *Nöldeke and Samuelson*. The closest paper to ours is Nöldeke and Samuelson (1993) (see also Samuelson, 1997, Chap. 8). An important difference between their model and ours is that each one of their individuals is characterized not only by an action (like ours) *but also by conjectures* about the composition of the populations at all nodes, whether reached or not. The dynamic then affects actions *and* conjectures. Our model is thus more basic; it is “operational” or

---

[77] We note that the condition $\mu m \geqslant \delta$ is used *only* in Steps 3 and 8 in the proof of Proposition 4.5, in order to bound the probability that the backward induction action is not played at descendant nodes; it is *not* used to show that nodes are reached with positive probability (recall also Section 4.1, particularly the last paragraph there).

"behavioristic" (only actual actions and payoffs matter), whereas theirs requires in addition more sophisticated "cognitive" considerations (beliefs or conjectures).

Nöldeke and Samuelson show that, in general, equilibria other than the backward induction are stable; for example, the equilibrium c in the game Γ_2 of Fig. 2. To see why this is so, notice, first, that the absorbing states in their model are the "self-confirming equilibria," which include the Nash equilibria as well as additional points (which allow for arbitrary beliefs at unreached nodes). Thus every $x = (\xi^1, \xi^2, \xi^3)$ with $\xi^1 = \xi^2 = 0$, where ξ^i denotes the proportion of c^i in population i, constitutes a self-confirming equilibrium, whereas it is a Nash equilibrium only if $\xi^3 \leqslant 1/3$. Since all these absorbing states are connected by a chain of single mutations from the backward induction equilibrium b (see Theorem 3.1 and (3.3) in its proof), it follows that all of them—including $(0, 0, 1)$—have positive $\hat{\pi}_0$-probability. But now starting at $(0, 0, 1)$, after one mutation in population 2, selection there will increase ξ^2; if $\xi^2 > 1/2$ occurs before $\xi^3 < 1/3$, then selection will start affecting population 1 as well. There is therefore a positive probability that a $(1, \xi^2, \xi^3)$ equilibrium in the component of $c = (1, 1, 1)$ will eventually be reached by selection. This implies that $\hat{\pi}_0[x] > 0$ for all x there, in particular $\hat{\pi}_0[c] > 0$.

In our basic model, once the system leaves the equilibrium component of $b = (0, 0, 0)$ and[78] $(0, 0, 1/3 + 1/m)$ is reached (by a sequence of single mutations), selection starts affecting the population at 2, moving it towards c^2. This makes node 3 reachable and then selection at 3 decreases ξ^3. When m is large, the probability is extremely low that ξ^2 will increase from 0 to $1/2 + 1/m$—which is the crucial proportion for selection to start operating at node 1—before ξ^3 decreases from $1/3 + 1/m$ to $1/3$ (indeed, the former requires $m/2$ "steps," the latter only 1 step). Therefore, the system is led back to the b-component with very high probability, and so, for large populations, $\hat{\pi}_0$ gives very small weight to c and its component. In our general model, if $(0, 0, 1/3 + 1/m)$ turns out, as in Nöldeke and Samuelson, to be absorbing for the selection dynamic Q_0 (recall Remark 1 in Section 3.1), then the same argument applies to $(0, 1/m, 1/3 + 1/m)$; to go from there to $(0, 0, 1/3 + 1/m)$ however requires an additional argument (see for instance (4.13)).

We note that for some variants of our dynamics—as in (e) in Section 5.1—it can in fact be shown that $\hat{\pi}_0[x] = 0$ for all x in the equilibrium component of c, for *every* population size $m \geqslant 2$. However, large populations would still be needed to obtain the backward induction equilibrium in general—for instance, in game Γ_3 of Fig. 3.

---

[78] Assume for simplicity that m is divisible by 6.

Finally, it is possible[79] that the Nöldeke and Samuelson model will also yield the backward induction equilibrium for large populations; they did not consider this setup.

(b) *Population dynamics*. The most relevant paper here is Cressman and Schlag (1998), who consider continuous-time replicator dynamics (like (5.1) in Section 5.1) for continuum populations (which would correspond to our large populations limit). They study (interior) asymptotic stability (rather than evolutionary or stochastic stability), and they show that only in simple games[80] does the process select the backward induction equilibria. As we said in (d) of Section 5.1, we conjecture that our result need not hold for such selection mechanisms. Another related work is Swinkels (1993), which provides sufficient conditions for myopic adjustment dynamics—a class that includes the replicator dynamics—to lead to (Kohlberg–Mertens) stable sets.

(c) *Fictitious play dynamics*. There are a number of papers that provide adaptive learning models, of the “fictitious play” variety, leading to the backward induction equilibria: see Canning (1992), Theorem 4; Hendon et al. (1996), Theorem 2; and Groes et al. (1999), Theorem 2. While these are not population games (there is a single player in each role), their dynamics can nevertheless be interpreted as population dynamics—see Hofbauer and Sigmund (1998), Section 8.4. By comparison to the present paper, notice that in these models the dynamic is a best-reply dynamic, and it is *assumed* that all nodes are reached (the “beliefs” of the players—the counterpart of our population frequencies—are completely mixed); in our model, this constitutes a substantial difficulty (see Section 4.1).

Acknowledgments

This research was partially supported by a grant of the Israel Academy of Sciences and Humanities. The starting point was a discussion during the Seventh Annual Conference of the Center for Rationality. The author thanks Bob Aumann, Dani Cohen, Stefano Demichelis, Ilan Eshel, Drew Fudenberg, Michihiro Kandori, Andreu Mas-Colell, Jean-François Mertens, Uzi Motro, Abraham Neyman, Bezalel Peleg, Motty Perry, Phil Reny, Larry Samuelson, Karl Schlag, Reinhard Selten, Avi Shmida, Jeroen Swinkels, Jörgen Weibull, Benji Weiss, Asher Wolinsky, Peyton Young, the anonymous referees, and the editors for their most useful comments.

---

[79] As conjectured by one referee.

[80] Like our examples Γ_1 and Γ_2 (but not Γ_3) in the Introduction—see their Theorem 4 and Remark 5.

References

Aumann, R.J., 1995. Backward induction and common knowledge of rationality. Games Econ. Behav. 8, 6–19.

Aumann, R.J., 1998. On the state of the art in game theory. Games Econ. Behav. 24, 181–210.

Canning, D., 1992. Learning the subgame perfect equilibrium. Discussion paper No. 608, Columbia University.

Cressman, R., Schlag, K.H., 1998. The dynamic (in)stability of backwards induction. J. Econ. Theory 83, 260–285.

Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Rev. Econ. Stud. 67, 17–45.

Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics. Theor. Pop. Biol. 38, 219–232.

Fudenberg, D., Levine, D.K., 1998. The Theory of Learning in Games. MIT Press.

Gale, J., Binmore, K., Samuelson, L., 1995. Learning to be imperfect: The ultimatum game. Games Econ. Behav. 8, 56–90.

Groes, E., Jacobsen, H.J., Sloth, B., 1999. Adaptive learning in extensive form games and sequential equilibrium. Econ. Theory 13, 125–142.

Hammerstein, P., Selten, R., 1994. Game theory and evolutionary biology. In: Aumann, R.J., Hart, S. (Eds.), Handbook of Game Theory, Vol. II. North-Holland, pp. 929–993.

Hendon, E., Jacobsen, H.J., Sloth, B., 1996. Fictitious play in extensive form games. Games Econ. Behav. 15, 177–202.

Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Population Dynamics. Cambridge University Press.

Kandori, M., Mailath, G., Rob, R., 1993. Learning, mutation, and long-run equilibrium in games. Econometrica 61, 29–56.

Kuhn, H.W., 1953. Extensive games and the problem of information. In: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games II. Annals of Mathematics Studies, Vol. 28. Princeton University Press, pp. 193–216.

Maynard Smith, J., Price, G.R., 1973. The logic of animal conflict. Nature 246, 15–18.

Nash, J., 1950. Equilibrium points in n-person games. Proc. Nat. Acad. Sci. USA 36, 48–49.

Nöldeke, G., Samuelson, L., 1993. An evolutionary analysis of backward and forward induction. Games Econ. Behav. 5, 425–454.

Samuelson, L., 1997. Evolutionary Games and Equilibrium Selection. MIT Press.

Selten, R., 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragetragheit. Z. Ges. Staatswissen. 12, 301–324.

Selten, R., 1975. Re-examination of the perfectness concept for equilibrium points in extensive games. Int. J. Game Theory 4, 25–55.

Swinkels, J.M., 1993. Adjustment dynamics and rational play in games. Games Econ. Behav. 5, 455–484.

Vega-Redondo, F., 1997. Evolution, Games, and Economic Behavior. Oxford University Press.

Weibull, J.W., 1995. Evolutionary Game Theory. MIT Press.

Young, H.P., 1993. The evolution of conventions. Econometrica 61, 57–84.

Young, H.P., 1998. Individual Strategy and Social Structure. Princeton University Press.