Evolutionary Dynamics and Backward Induction*

Sergiu Hart†

November 14, 1999

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.

Keywords: games in extensive form, games of perfect information, backward induction equilibrium, subgame-perfect equilibrium, evolutionary dynamics, evolutionary stability, mutation, selection, population games.

Journal of Economic Literature Classification: C7, D7, C6.

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

*Previous versions: September 1999; May 1999. Research partially supported by a grant of the Israel Academy of Sciences and Humanities. The starting point for this research was a discussion during the Seventh Annual Conference of the Center for Rationality. The author thanks Bob Aumann, Dani Cohen, Ilan Eshel, 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, Jörgen Weibull, Benji Weiss, Asher Wolinsky, and Peyton Young for their useful comments.

†Center for Rationality and Interactive Decision Theory; Department of Economics; and Department of Mathematics; The Hebrew University of Jerusalem, 91904 Jerusalem, Israel.
E-mail: hart@math.huji.ac.il *URL:* http://www.ma.huji.ac.il/~hart

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 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 parallel notions in the two fields are "strategic equilibrium" (introduced by J. Nash in 1950) and "evolutionarily stable strategy" (introduced by J. Maynard Smith and G. R. Price in 1973). Roughly speaking, when a game is played by populations of individuals (with identical payoff functions), then a Nash equilibrium point essentially yields evolutionarily stable strategies. 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]).

Most of these models deal with games in strategic (or normal) form. Here we consider instead *games in extensive form*, where a most complete description of the game is given, specifying exactly 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 such 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 R. Selten in 1965, 1975.

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), individual (experimentation, stimulus response), social (learning, imitation), and so on. What matters is that the process is "adaptive" or "improving," in the sense that the proportion of better strategies increases. Evolutionary models have been extensively analyzed in various classes of games in strategic form (starting with Kandori, Mailath and Rob [1993] and Young [1993]; see also the books of Young [1998] and Fudenberg and Levine [1998]).

Since mutations are like small perturbations that make everything possible (i.e., every pure strategy has positive probability), and this yields in the limit (as the perturbations go to zero) 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 up to now has found the above claim to be false: Evolutionary models do not necessarily pick out the backward induction equilibria. Specifically, except in special classes of games, other equilibria besides the backward induction ones also turn out to be "evolutionarily stable" (see Nöldeke and Samuelson [1993]; Gale, Binmore and Samuelson [1995]; Cressman and Schlag [1997]; and the books of Samuelson [1997] and Fudenberg and Levine [1998]).

1.2. Examples

Even without specifying exactly how selection and mutation operate, one can get some intuition by considering a few examples. The first one is the 2-person game Γ_1 of Figure 1.1. 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

[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.

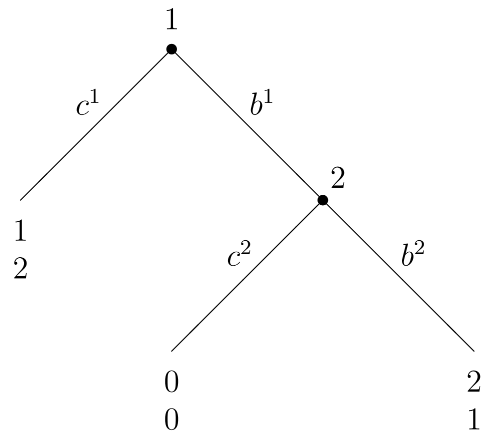

Figure 1.1: The game Γ_1

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: He gets 0 (since the individuals at 2 that he will meet play c^2) instead of 1. 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 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 (this is called "genetic drift"). Eventually, after a sufficiently long time,[3] more than half the population at 2 will consist of b^2-individuals. 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 back-

[3]The assumption is that mutations have positive—though small—probability at each period.

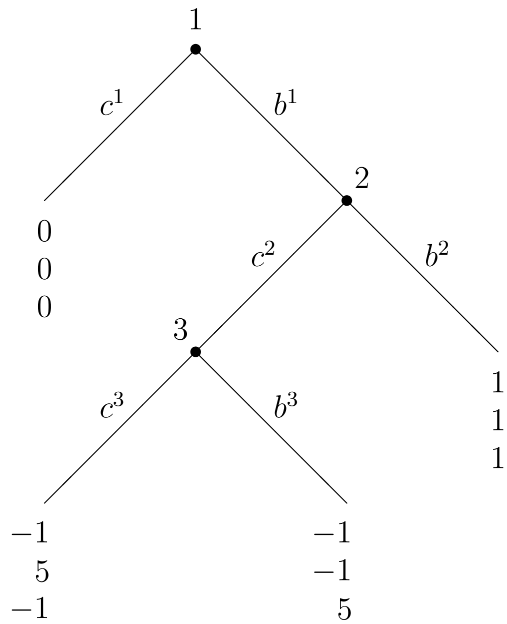

Figure 1.2: The game Γ_2

ward induction equilibrium; in other words, b is the evolutionarily stable equilibrium.

The next example is the 3-player game Γ_2 of Figure 1.2 (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 $c = (c^1, c^2, c^3)$. 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, 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—towards the best-reply 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 is a "reflecting barrier" below the 2/3 mark. All this happens fast enough that the

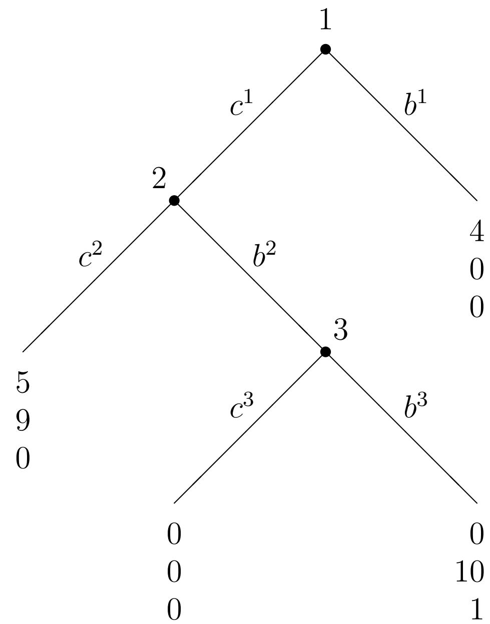

Figure 1.3: The game Γ_3

population at 1, which is playing b^1, can move only a little, if at all. Therefore we have essentially shown that the equilibrium component[4] 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 is in contrast to the result of Nöldeke and Samuelson [1993]: In their model,[5] the non-subgame-perfect equilibrium c also belongs to the stable set.

Consider now another 3-player game: the game Γ_3 given in Figure 1.3. The backward induction equilibrium is b, at which both nodes 2 and 3 are unreached.

[4] Two (mixed) Nash equilibria belong to the same (*equilibrium*) *component* if their equilibrium paths coincide, and they differ only at unreached nodes.

[5] The main difference between their model and ours is that each one of their individuals, in addition to playing a pure action (as ours do), has conjectures about the composition of the populations at all nodes, whether reached or not. The dynamic then affects both actions and conjectures. In a biological world, it is hard to see how such conjectures are encoded into genes; our individuals are therefore characterized by their actions only.

The populations at 2 and 3 therefore move by mutations. Eventually, when the proportion of b^2 at node 2 goes 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 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 has 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.*

1.3. This Paper

As stated above, the games we study 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[7]). 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[8] 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

[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).

[7]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.

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

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—to which we refer 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").

Since such dynamics yield an ergodic system,[9] the 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 states which occur with positive frequency however low the mutation rate is.[10] Such states are called "evolutionarily stable." Two preliminary results are that only Nash equilibria can be evolutionarily stable, and that the backward induction equilibrium is always evolutionarily stable. The examples show that other Nash equilibria may however 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.*[11]

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

[9] Mutations make every state reachable from any other state.

[10] More precisely, states whose probability, according to the invariant distribution, is bounded away from zero as the probability of mutation goes to zero.

[11] With an additional proviso that the population sizes be large enough, depending on the mutation rate; see the precise statement in Subsection 3.2.

b itself.[12] The intuition for the role of the large population assumption will be provided in Subsection 4.1 below. Suffice it to say here that it has to do with a change of action (whether by mutation or selection) being 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 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]

To conclude the introduction, we note that two almost diametrically opposed approaches each 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 something which they know is not optimal), and moreover assumes 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 (this paper), in the realm of evolutionary dynamics, is essentially machinelike and requires no conscious optimization or rationality.[15] It is striking that such disparate models converge.[16]

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

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

[13] An exception is Nöldeke and Samuelson [1995].

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

[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).

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 all outgoing branches; or a move of one of the players, in which case the vertex will be 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 multi-linearly to *randomized* (or *mixed*) actions; thus $u^i : X \to \mathbb{R}$, where $X := \prod_{i \in N} X^i$ and $X^i := \Delta(A^i) = \left\{x^i \in \mathbb{R}_+^{A^i} : \sum_{a^i \in A^i} x^i_{a^i} = 1\right\}$, the unit simplex on A^i, is the set of probability distributions over A^i.

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] $u^i(x) \geq u^i(y^i, x^{-i})$ 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-

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

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

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 player i. The populations are assumed to be distinct (i.e., $M(i) \cap M(j) = \emptyset$ for $i \neq j$). Each individual $q \in M(i)$ plays a pure action in A^i, which we denote by $\omega^i_q \in A^i$;

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

put $\omega^i = \left(\omega^i_q\right)_{q \in M(i)}$ and $\omega = (\omega^i)_{i \in N}$. For each $a^i \in A^i$, let[20]

$$x^i_{a^i} \equiv x^i_{a^i}\left(\omega^i\right) := \frac{\left|\left\{q \in M(i) : \omega^i_q = a^i\right\}\right|}{|M(i)|} \tag{2.1}$$

be the proportion of population $M(i)$ that plays the action a^i; then $x^i \equiv x^i\left(\omega^i\right) := \left(x^i_{a^i}\left(\omega^i\right)\right)_{a^i \in A^i} \in X^i$ 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., $u^i(\omega^i_q, x^{-i})$; we shall slightly abuse notation by writing this as $u^i(\omega^i_q, \omega^{-i})$.

We refer to the above model as the *gene-normal form of* Γ (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., each player plays at most once in Γ). When however a player may play more than once (and thus has more than one agent), then the "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 should have the two nodes i and j in the same information set.[21,22] For a discussion of these issues and their import, the reader is referred to Subsection 5.1.

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 $\omega^i = \left(\omega^i_q\right)_{q \in M(i)}$ and $\omega^i_q \in A^i$ for each $i \in N$ and each $q \in M(i)$. Let $\Omega := \prod_{i \in N} \left(A^i\right)^{M(i)}$ be the state space. We consider discrete-time stochastic systems: Starting with an initial

[20] $|Z|$ denotes the number of elements of the 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 allowed here.

state[23] $\omega_1 \in \Omega$, a sequence of states $\omega_1, \omega_2, ..., \omega_t, ...$ in Ω is generated, according to certain probabilistic rules. These so-called "transition probabilities" specify, for each $t = 1, 2, ...$, the probabilities $P[\omega_{t+1} = \widetilde{\omega} \mid \omega_1, ..., \omega_t]$ that ω_{t+1} equals a state $\widetilde{\omega} \in \Omega$, given the history $\omega_1, \omega_2, ..., \omega_t$. 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 $\omega_1, ..., \omega_{t-1}$, nor on the "calendar time" t). That is, there is a stochastic matrix[24] $Q = (Q[\widetilde{\omega} \mid \omega])_{\widetilde{\omega}, \omega \in \Omega}$ such that $P[\omega_{t+1} = \widetilde{\omega} \mid \omega_1, ..., \omega_t] = Q[\widetilde{\omega} \mid \omega_t]$ for every $\omega, \widetilde{\omega} \in \Omega$ and $t = 1, 2, ...$ 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 $\mu + \sigma \leq 1$. The one-step transition probabilities $Q[\widetilde{\omega} \mid \omega]$ are given by the following process, performed independently for each $i \in N$:

- Choose a player $q(i) \in M(i)$ at random: All m players in $M(i)$ have the same probability $1/m$ of being chosen.

- Put $\widetilde{\omega}^i_q := \omega^i_q$ 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

  $$B^i \equiv B^i(q(i), \omega) := \left\{a^i \in A^i : u^i(a^i, \omega^{-i}) > u^i(\omega^i_{q(i)}, \omega^{-i})\right\}; \qquad (2.2)$$

  this is the set of actions of player i that are strictly better in Γ, against the populations at the other nodes, than the action $\omega^i_{q(i)}$ of the chosen player $q(i)$. If B^i is not empty, then the new action $\widetilde{\omega}^i_{q(i)}$ of $q(i)$ is a randomly chosen better action: $\widetilde{\omega}^i_{q(i)} := a^i$ 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: $\widetilde{\omega}^i_{q(i)} := \omega^i_{q(i)}$.

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

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

- If mutation MU(i) was chosen, then $\widetilde{\omega}^i_{q(i)}$ is a random action in A^i; i.e., $\widetilde{\omega}^i_{q(i)} := a^i$ with probability $1/\left|A^i\right|$ for each $a^i \in A^i$.

- If no-change NC(i) was chosen, then the action of $q(i)$ does not change: $\widetilde{\omega}^i_{q(i)} := \omega^i_{q(i)}$.

For example, in the game Γ_1 of Subsection 1.2 (see Figure 1.1; 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 $\widetilde{\omega} = ((b^1, c^1, c^1), (b^2, b^2, b^2))$, then $Q[\widetilde{\omega} \mid \omega] = (1/3) \cdot (\mu/2 + \sigma) \cdot (1/3) \cdot (\mu/2)$. 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 $1/3$, 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 that 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 the Poisson distribution as an approximation to the binomial distribution). Our assumption may thus be viewed as essentially nothing more than a convenient rescaling of time; if one expects, say, k changes each period, then time should be rescaled by a factor of $1/k$. As we shall see below (in particular in Subsection 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

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

better." Of course, "selection" may stand for various processes of adaptation, imitation, learning, experimentation, and so on. Our model abstracts away from particular selection mechanisms, and merely assumes that better actions fare better. See also Subsection 5.3 for the case where the selection probability of a better action is proportional to its current proportion in the population.

3. The basic dynamic is determined by two parameters,[26] 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. The no-change periods may be viewed as "payoff accumulation" periods, or as "selection realization" periods (i.e., periods during which the actual selection occurs[27]).

4. The one-step transition probabilities were defined to be independent over the agents; 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 will 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 of different size at different nodes. The process is a stationary Markov chain, whose one-step transition probability matrix $Q = (Q[\widetilde{\omega} \mid \omega])_{\widetilde{\omega},\omega\in\Omega}$ satisfies:

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

[27] This may help 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. See also Subsection 5.3.

- Conditional independence over $i \in N$, i.e.,[28]

$$Q\left[\widetilde{\omega} \mid \omega\right] = \prod_{i \in N} Q\left[\widetilde{\omega}^i \mid \omega\right]. \tag{2.3}$$

- 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

$$\frac{\gamma_1}{|M(i)|} \leq Q\left[q(i) = q \mid \omega\right] \leq \frac{\gamma_2}{|M(i)|} \text{ for each } q \in M(i); \text{ and} \tag{2.4}$$

$$Q\left[\widetilde{\omega}_q^i = \omega_q^i \text{ for all } q \in M(i) \backslash \{q(i)\} \mid \omega\right] = 1. \tag{2.5}$$

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

$$Q\left[\widetilde{\omega}_{q(i)}^i = a^i \mid \omega\right] \geq \beta \text{ for each } a^i \in B^i, \tag{2.6}$$

where $B^i \equiv B^i(q(i), \omega)$ is the set of strictly better actions, as defined in (2.2).

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

$$Q\left[\widetilde{\omega}_{q(i)}^i = a^i \mid \omega\right] \geq \alpha_1 \mu \text{ for each } a^i \in A^i; \text{ and} \tag{2.7}$$

$$Q\left[\widetilde{\omega}_{q(i)}^i = a^i \mid \omega\right] \leq \alpha_2 \mu \text{ for each } a^i \notin B^i, a^i \neq \omega_{q(i)}^i. \tag{2.8}$$

Without loss of generality, all parameters $\alpha_1, \alpha_2, \beta, \gamma_1, \gamma_2$ 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 $\alpha_1 = \alpha_2 = 1/|A^i|$.

The general model thus assumes that: (i) the probabilities of various individuals in the same population being chosen are comparable;[29] (ii) the effect of selection—towards better actions—is bounded away from zero (independently of μ); and (iii) the effect of mutation—with every action being possible—is of the order μ. The reader is referred to Subsections 5.2 and 5.3 for generalizations.

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

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

3. The Results

3.1. Preliminary Results

A general model with a one-step transition matrix Q satisfying (2.3)–(2.8) yields a Markov chain which is *ergodic*, since the probability of reaching any state $\omega' \in \Omega$ from any other state $\omega \in \Omega$ is positive (this follows from (2.4) and (2.7), by using an appropriate sequence of mutations). Therefore there exists a unique *invariant distribution* π on Ω; i.e., a unique $\pi \in \Delta(\Omega)$ satisfying $\pi = \pi Q$, or

$$\pi\left[\widetilde{\omega}\right] = \sum_{\omega \in \Omega} \pi\left[\omega\right] Q\left[\widetilde{\omega} \mid \omega\right]$$

for every $\widetilde{\omega} \in \Omega$. The long-run behavior of the process is well described by π, in the following two senses:

- The relative frequency, in any long enough period of time, of the process visiting a state ω, is approximately $\pi\left[\omega\right]$; i.e., for every $\omega \in \Omega$:

$$\lim_{T_2 - T_1 \to \infty} \frac{\left|\left\{t : T_1 < t \leq T_2, \omega_t = \omega\right\}\right|}{T_2 - T_1} = \pi\left[\omega\right].$$

- The probability that the process is in state ω at some period t is approximately $\pi\left[\omega\right]$ for large t; i.e., for every $\omega \in \Omega$:

$$\lim_{t \to \infty} P\left[\omega_t = \omega\right] = \pi\left[\omega\right].$$

Note that the two properties hold regardless of the initial state; moreover, they hold 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 $\pi\left[\omega\right]$ is bounded away from zero as $\mu \to 0$; i.e., if $\liminf_{\mu \to 0} \pi\left[\omega\right] > 0$. Recall that each state $\omega \in \Omega$ may be viewed as an N-tuple of mixed actions $x\left(\omega\right) = \left(x^i\left(\omega^i\right)\right)_{i \in N} \in X$ (see (2.1)). The invariant distribution π on Ω therefore induces a probability distribution[30] $\widehat{\pi} := \pi \circ \left(x\right)^{-1}$ over X; i.e., $\widehat{\pi}\left[Y\right] :=$

[30] $\left(x\right)^{-1} : X \to \Omega$ denotes the inverse of the mapping $x : \Omega \to X$.

$\pi[\{\omega \in \Omega : x(\omega) \in Y\}]$ 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 $\liminf_{\mu \to 0} \widehat{\pi}[x] > 0$. The following result states that only Nash equilibria can be evolutionarily stable, and that the backward induction equilibrium b is always so.

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 (2.3)–(2.8). Then*

$$\begin{aligned} \lim_{\mu \to 0} \widehat{\pi}_\mu [EQ] &= 1, \text{ and} \\ \liminf_{\mu \to 0} \widehat{\pi}_\mu [b] &> 0. \end{aligned}$$

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 ergodic, so its invariant distribution is not unique). Now Q_0 allows no mutations (by (2.8) and $\mu \to 0$), so only selection applies. Therefore $\omega \in \Omega$ is a Nash equilibrium state (i.e., $x(\omega) \in EQ$) if and only if ω is an absorbing state of Q_0 (i.e., $Q_0[\omega \mid \omega] = 1$); denote by Ω_{EQ} the subset (of Ω) of all such states. To prove that an absorbing state must always be reached—and thus $\pi_0[\Omega_{EQ}] = 1$—we shall show that Q_0 allows no cycles. Indeed, assume that $\omega_0, \omega_1, ..., \omega_t, ..., \omega_T \in \Omega$ satisfy $\omega_t \neq \omega_{t-1}$ and $Q_0[\omega_t \mid \omega_{t-1}] > 0$ for every $t = 1, ..., T$, and $\omega_T = \omega_0$. At a final node $i \in N$ (i.e., with $N(i) = \emptyset$), selection can only increase the sum of the "local" payoffs (in $\Gamma(i)$) of the population $M(i)$, i.e.,[31] $\sum_{q \in M(i)} u^i_{\Gamma(i)}\left(\omega^i_q\right)$; therefore $\omega^i_T = \omega^i_0$ implies that $\omega^i_t = \omega^i_0$ for all t; thus 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.

Next, we shall show that $\widehat{\pi}_0[b] > 0$. First, we claim that, given two Nash equilibrium states $\omega, \omega' \in \Omega_{EQ}$, if $\pi_0[\omega] > 0$ and ω' can be reached from ω by one mutation step in one population followed by any number of selection steps, then

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

$\pi_0[\omega'] > 0$ too. Indeed, the invariance property $\pi_\mu = \pi_\mu Q_\mu$ implies[32] $\pi_\mu = \pi_\mu Q_\mu^k$ for any integer $k \geq 1$, and thus

$$\pi_\mu[\omega'] \geq \pi_\mu[\omega'] Q_\mu^k[\omega' \mid \omega'] + \pi_\mu[\omega] Q_\mu[\omega'' \mid \omega] Q_\mu^{k-1}[\omega' \mid \omega''],$$

where ω'' satisfies $Q_\mu[\omega'' \mid \omega] \geq c_1\mu$ for some $c_1 > 0$ (by (2.4) and (2.7); 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 Nash equilibrium state, it can change only by mutations, so $Q_\mu^k[\omega' \mid \omega'] \geq 1 - c_3\mu$ for an appropriate constant $c_3 > 0$ (by (2.8)). Therefore $c_3\mu\pi_\mu[\omega'] \geq c_1\mu\pi_\mu[\omega] Q_\mu^{k-1}[\omega' \mid \omega'']$, which, after dividing by μ and then letting $\mu \to 0$, yields

$$\pi_0[\omega'] \geq \frac{c_1}{c_3}\pi_0[\omega] Q_0^{k-1}[\omega' \mid \omega''] = \frac{c_1 c_2}{c_3}\pi_0[\omega] > 0,$$

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

Second, given a final node $i \in N$, we claim that there is a Nash equilibrium state $\omega \in \Omega_{EQ}$ with $\pi_0[\omega] > 0$, at which all the population at i plays the backward induction action b^i. If not, let $\omega \in \Omega_{EQ}$ be such that $\pi_0[\omega] > 0$ and the proportion $x^i_{b^i}(\omega)$ of b^i is maximal; 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 $\omega \in \Omega_{EQ}$), followed by any number of selection periods until a state $\omega' \in \Omega_{EQ}$ is reached. By the claim of the previous paragraph, 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 Nash equilibrium states $\omega \in \Omega_{EQ}$ 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 Nash equilibrium states $\omega \in \Omega_{EQ}$ with $\pi_0[\omega] > 0$ and $x^j_{b^j}(\omega) = 1$ for all $j \in N(i)$ (just choose 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 $\omega_b \in \Omega_{EQ}$ where $x^i_{b^i}(\omega_b) = 1$ for all $i \in N$. ■

We note that the above proof shows that the result of Theorem 3.1 holds also in the non-generic case where the backward induction equilibrium is not unique (the second statement then applies to each $b \in BI$).

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

3.2. The Main Result

As the examples show, other equilibria besides the backward induction equilibrium b (including some that are very different from b) may be evolutionarily stable when the populations are fixed. 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[33] to consider (small) neighborhoods of $BI = \{b\}$ in the set of mixed actions X: For every $\varepsilon > 0$, put $BI_\varepsilon := \left\{x \in X : x^i_{b^i} \geq 1 - \varepsilon \text{ for all } i \in N\right\}$. That is, x belongs to the ε-neighborhood BI_ε of b if most of the individuals, in all populations, 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 $\mathbf{m} = (|M(i)|)_{i \in N}$ with $m := \min_{i \in N} |M(i)|$, 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 (2.3)–(2.8). Then, for every $\varepsilon > 0$ and $\delta > 0$,*

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

Moreover, there exists a constant c, depending on the game, on the dynamics parameters $\alpha_1, \alpha_2, \beta, \gamma_1, \gamma_2$, and on ε, δ, such that[34]

$$E_{\mu,\mathbf{m}} \left[x^i_{b^i}(\omega)\right] \geq 1 - c\mu \text{ for all } i \in N, \text{ and} \tag{3.1}$$

$$\pi_{\mu,\mathbf{m}} \left[x^i_{b^i}(\omega) \geq 1 - \varepsilon \text{ for all } i \in N\right] \geq 1 - c\mu, \tag{3.2}$$

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

[33] The probability of a single point may become 0 in the limit. For instance, if $1/m$ is much smaller than μ (i.e., if $m\mu \to \infty$), then we may well get $\widehat{\pi}[b] \to 0$ (consider even the simplest 1-person game: The transition probability from the state ω_0 where everyone plays b, to the 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 $1/m$).

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

Thus, as the mutation rate is low and the populations are large enough, 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.

Remarks.

1. The only assumption made on the relative rates of convergence of μ and m is that μm is bounded away from 0. This implies that
$$\lim_{\mu \to 0} \lim_{\mathbf{m} \to \infty} \widehat{\pi}_{\mu,\mathbf{m}} [BI_\varepsilon] = 1 \tag{3.3}$$
(however this need not hold for $\lim_{\mathbf{m}\to\infty} \lim_{\mu\to 0}$).

2. No assumptions are made on the relative population sizes $|M(i)|$; one population may well be much larger than another. However, the mutation rates in the different populations are assumed to be of the same order of magnitude—see (2.7) and (2.8).

3. The estimates we get in (3.1) and (3.2) 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).[35] 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 will explain the role of large populations. We do so for the simpler basic model; the same arguments apply to the general model.

[35] 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.

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[36] than the probability of selection SE(i) at i. Therefore, at reached nodes most of the individuals play their best-reply actions. (This is essentially the argument of Theorem 3.1.) The problem is how to get a similar conclusion at the *unreached* nodes.

Consider first the 3-player game Γ_2 of Figure 1.2 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 and 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 back 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[37] m/ρ periods for it to happen. Therefore, during a long stretch of time, say T periods, the number of periods that there is a c^2 in population 2 is about $(\mu/2)\,(m/\rho)T = \mu m T/(2\rho)$. 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 is unreached most of the time.

Consider next the 4-person game Γ_4 of Figure 4.1. 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

[36] 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 $m \to \infty$.

[37] 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.

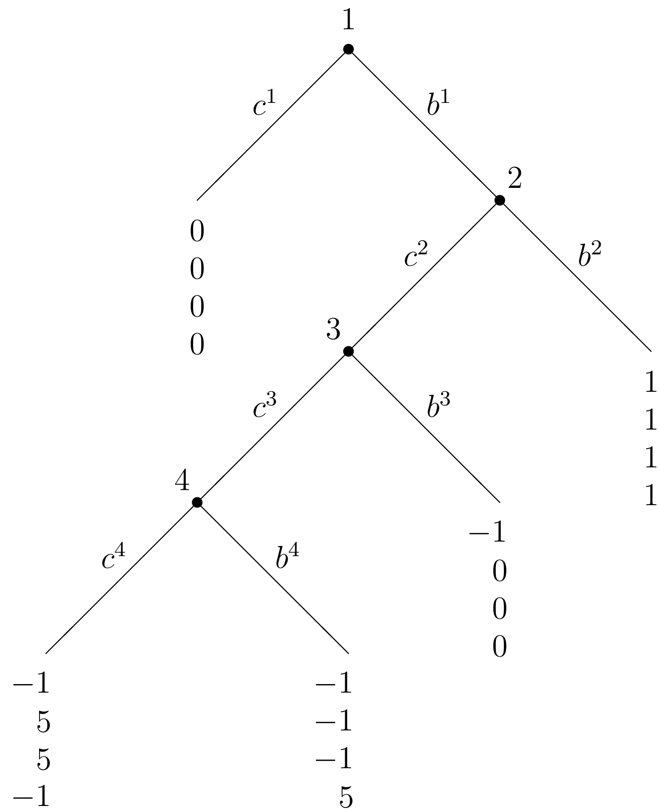

Figure 4.1: The game Γ_4

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 back). But we 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 back to b^3 is by mutation at 3; the probability of that happening is $(1/m)(\mu/2) = \mu/(2m)$ (since that 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 back, 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[38] 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[39] left with a 3-player game (like Γ_2); the proof is completed by (backward) induction.[40]

The crux of the argument is that, after a mutation in population 3 generated a c^3, this c^3 action is "stuck" there for a long time[41]—at least until a mutation generates a c^2 in population 2. Thus, what appeared to require *simultaneous*

[38]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.

[39]See the last paragraph in this subsection.

[40]Similar arguments apply to the 3-player game Γ_3 of Figure 1.3 of the Introduction. 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.

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

mutations (with a frequency of the order of μ^k for some $k \geq 2$), turns out instead to rely on *sequential mutations* (with a frequency of the order of μ).[42]

It should now be clear what role the large populations play: The smaller a group of individuals is, the (relatively) less probable it is for a change of action to 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 2 in Γ_4, or population 3 in Γ_2).

Finally, to understand the use of the condition that $\mu m \geq \delta > 0$, 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 sizeable fraction of the population does not play the backward induction action, albeit an event of low probability for large m, is not negligible. A simple estimate[43] 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 $\mu m \geq \delta > 0$—this effect (like that of random mutations) is again small relative to μm for large m.

4.2. An Outline of the Proof

We now provide an outline that may help the reader 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).

[42] This kind of argument may also explain how matching mutations occur in interacting populations; i.e., mutations that yield no advantage in their own population, unless there are fitting 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.

[43] Using Markov's inequality. More refined probabilistic computations may well lead to weaker conditions.

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 (this holds since selection has probability at least β, which is bounded away from 0, while mutation has probability μ); again, it is of the order of μ; see (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 (4.12). 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 $1/m \leq (1/\delta)\,\mu$ implies that the expected proportion of population i that does not play b^i is of the order of μ, which yields (3.1) and (3.2).

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 $\mu m \geq \delta$, where $m := \min_{i \in N} m^i$) and the transition probability matrix Q that satisfies (2.3)–(2.8). 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 $\widetilde{\omega} \in \Omega$ be the next state, given by the one-step transition probabilities $Q[\widetilde{\omega} \mid \omega]$; then $\widetilde{\omega}$ is also distributed according to π. From now on all probability statements and expectations will be according to this distribution.

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 (2.1)); this is the proportion of population i that does *not* play the backward induction action in state ω. Similarly, put $\widetilde{Y}^i := 1 - x^i_{b^i}(\widetilde{\omega})$ for the same proportion in the next-period state $\widetilde{\omega}$. The random variables Y^i and $\widetilde{Y}^i$ are identically distributed (their distribution is $\pi \circ \left(1 - x^i_{b^i}\right)^{-1}$); thus in particular $E\left[\widetilde{Y}^i\right] = E\left[Y^i\right]$.

- 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 to be 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_q^k 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 $\widetilde{\omega}$.

- When everyone plays the backward induction action—i.e., when $Y^j = 0$ for all $j \in N$—the unique local best-reply of 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 to be 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., $N(i) = \emptyset$) 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 repeatedly used in the proof, for convenience we state the following implications of (2.3)–(2.8) here:[44]

$$P\left[\widetilde{\omega}^i_{q(i)} = a^i \mid \omega\right] \geq \alpha_1 \mu \text{ for every } a^i \in A^i. \tag{4.1}$$

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

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

$$\text{If } L^i R^i = 1 \text{ and } \omega^i_{q(i)} \neq b^i \text{ then } P\left[\widetilde{\omega}^i_{q(i)} = b^i \mid \omega\right] \geq \beta. \tag{4.3}$$

The crucial argument in our proof is the following:

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 $1, 2, ..., i-1, i$ (in that order, with 1 the root). Let Z be a non-negative random variable. Then*

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

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

Lemma 4.2.

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

Proof. For each ω with $R^i = 1$, to get $\widetilde{R}^{j,i} = 0$ there must be some node k with $j \leq k < i$ and $\widetilde{V}^k = 0$. But $V^k > 0$ (since $R^i = 1$), so in fact $V^k = 1/m^k$ (since $\left|\widetilde{V}^k - V^k\right|$ is 0 or $1/m^k$). Hence $P\left[\widetilde{V}^k = 0 \mid \omega\right] \leq \gamma_2/m^k = O\left(1/m\right)$ (by (2.4), since the single c^k-individual must be chosen). Therefore $P\left[\widetilde{R}^{j,i} = 0 \mid \omega\right] \leq \sum_{k=j}^{i-1} P\left[\widetilde{V}^k = 0 \mid \omega\right] = O\left(1/m\right)$, from which it follows that

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

■

Lemma 4.3.

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

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 $j \leq k < i$ and $\widetilde{V}^k = 0$. But $R^{j,i} = 1$ implies that $V^k > 0$, and thus $V^k = 1/m^k$. Therefore $P\left[\widetilde{V}^k = 0 \mid \omega\right] \leq \left(\gamma_2/m^k\right) O(\mu) = O(\mu/m)$ (by (2.4) 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 $P\left[\widetilde{R}^{j,i} = 0 \mid \omega\right] \leq \sum_{k=j}^{i-1} P\left[\widetilde{V}^k = 0 \mid \omega\right] = O(\mu/m)$, from which the result follows. ■

Lemma 4.4.

$$E\left[Z\left(1 - \widetilde{R}^{j,i}\right)\left(1 - R^{j,j+1}\right) R^{j+1,i}\right] \leq (1 - \alpha_1\mu)\, E\left[Z\left(1 - R^{j,j+1}\right) R^{j+1,i}\right] + O\left(\frac{\mu}{m}\right).$$

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 $j \leq k < i$ and $\widetilde{V}^k = 0$. Now $P\left[\widetilde{V}^j > 0 \mid \omega\right] = P\left[\widetilde{\omega}^j_{q(j)} = c^j \mid \omega\right] \geq \alpha_1\mu$ (this follows from $V^j = 0$ and (4.1)), and $P\left[\widetilde{V}^k = 0 \mid \omega\right] \leq \gamma_2/m^k \leq \gamma_2/m$ for $k = j+1, ..., i-1$ (by (2.4), since $R^{j+1,i} = 1$ implies $V^k > 0$). Therefore $P\left[\widetilde{R}^{j,i} = 1 \mid \omega\right] = P\left[\widetilde{V}^k > 0 \text{ for all } j \leq k < i-1 \mid \omega\right]$ $\geq (\alpha_1\mu)(1 - \gamma_2/m)^{i-j-1}$, hence $P\left[\widetilde{R}^{j,i} = 0 \mid \omega\right] \leq (1 - \alpha_1\mu) + O(\mu/m)$, 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, 4.3 and 4.4 together with

$$E\left[Z\left(1 - \widetilde{R}^{j,i}\right)\left(1 - R^{j+1,i}\right)\right] \leq E\left[Z\left(1 - R^{j+1,i}\right)\right]$$

yields:

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

Rearranging terms gives

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

Adding these inequalities for $j = 1, 2, ..., i-1$ 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 $i \in N$:*

$$\begin{aligned}
P\left[L^i = 0\right] &= O(\mu); && (4.4)\\
P\left[\widetilde{Y}^i < Y^i\right] = P\left[\widetilde{Y}^i > Y^i\right] &= O(\mu); && (4.5)\\
E\left[Y^i L^i R^i\right] &= O(\mu); && (4.6)\\
E\left[\widetilde{Y}^i \widetilde{L}^i R^i\right] &= O(\mu); && (4.7)\\
E\left[\left|\widetilde{Y}^i \widetilde{L}^i - Y^i L^i\right| \left(1 - R^i\right)\right] &= O\left(\frac{\mu}{m}\right); && (4.8)\\
E\left[\widetilde{Y}^i \widetilde{L}^i (1 - R^i)\right] &= O\left(\frac{1}{m}\right); && (4.9)\\
E\left[Y^i L^i (1 - R^i)\right] &= O\left(\frac{1}{m}\right); \text{ and} && (4.10)\\
E\left[Y^i\right] &= O(\mu). && (4.11)
\end{aligned}$$

Proof. The proof is by backward induction on i. Assume that (4.4)–(4.11) hold for all[45] $j \in N(i)$; then each of the claims (4.4)–(4.11) for i will be proved in turn.

Step 1: (4.4) holds for i. Indeed,[46] $L^i = 0$ implies that there is $j \in N(i)$ such that $Y^j \geq \lambda$, hence

$$P\left[L^i = 0\right] \leq \sum_{j \in N(i)} P\left[Y^j \geq \lambda\right] \leq \sum_{j \in N(i)} \frac{1}{\lambda} E\left[Y^j\right] = O(\mu),$$

where we have used Markov's inequality[47] and (4.11) for j (by the induction hypothesis).

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

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

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

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

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

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

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

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 $\widetilde{\omega}^i_{q(i)} \neq b^i$). 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 (2.8)). Thus

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

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 $\geq \gamma_1 Y^i$ by (2.4). For every $\omega \in \Omega$ with $L^i R^i = 1$, the probability $P\left[\widetilde{\omega}^i_{q(i)} = b^i \mid \omega\right]$ of changing the action to b^i is at least β (see (4.3)), and we get

$$\begin{aligned} P\left[\widetilde{Y}^i < Y^i\right] &\geq E\left[\beta\gamma_1 Y^i \mid L^i R^i = 1\right] P\left[L^i R^i = 1\right] + 0 \; P\left[L^i R^i \neq 1\right] \\ &= \beta\gamma_1 E\left[Y^i L^i R^i\right]. \end{aligned}$$

Using (4.5) for i thus yields (4.6) for i.

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

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

(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

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

The first term is bounded by $(1/m^i)\, P\left[\widetilde{Y}^i \neq Y^i, R^i = 0\right] = (1/m^i)\, O(\mu) = O(\mu/m)$ (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 $Y^j \geq \lambda - 1/m^j$ (otherwise $Y^j < \lambda - 1/m^j$ and thus $\widetilde{Y}^j < \lambda$ for all $j \in N(i)$, in which case $L^i = \widetilde{L}^i = 1$). Choose $j \in N(i)$ to be a last such node; thus $Y^l < \lambda - 1/m^l$ for all $l \in N(j)$, hence $L^j = 1$. Now $\widetilde{L}^i \neq L^i$ implies that $\widetilde{\omega}^k \neq \omega^k$ 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

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

where we have used $1 - R^i \leq 1 - R^j$ 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 $Z = \widetilde{Y}^i \widetilde{L}^i$ in Proposition 4.1. We have

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

where we have used the fact that π is the invariant distribution, the inequality $1 - R^{j,i} \leq 1 - R^i$ (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

$$\mu E\left[\widetilde{Y}^i \widetilde{L}^i \left(1 - R^i\right)\right] \leq O\left(\frac{1}{m}\right) E\left[\widetilde{Y}^i \widetilde{L}^i R^i\right] + O\left(\frac{\mu}{m}\right). \tag{4.12}$$

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

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

since[48] $1/m \leq (1/\delta)\,\mu$. Together with

$$E\left[Y^i(1-L^i)\right] \leq P\left[L^i = 0\right] = O(\mu)$$

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 (3.1); applying Markov's inequality then yields (3.2). ■

5. Discussion

5.1. Multiple Agents

The gene-normal form, which was defined in Subsection 2.2, assumes that the populations at the different nodes are distinct. We shall now consider the implications of this assumption in the case where a player in the original game Γ consists of more than one agent; that is, that player moves at more than one node. For concreteness, assume that two nodes i and j belong to the same player. There are three main implications:

1. The actions at i and at j are determined by different "characteristics." This poses no problem, since different decision nodes in the tree correspond to different situations, and each one is controlled by a different "gene." As we have already argued in Subsection 2.2, if instead the same gene controls both decisions, it means that the player acts identically at both. The correct description of such an instance is then to have the two nodes belong to the same "information set"—i.e., a game with imperfect information (and possibly imperfect recall too).

[48] This is the only place in the proof where $\mu m \geq \delta$ is used.

2. The payoffs of all participants are computed based on independent draws from the populations; i.e., $u^k(\omega^k_q, \omega^{-k}) = u^k(\omega^k_q, x^{-k})$, where x^{-k} is the product of x^l for all $l \neq k$. In particular, the proportion of a pair of actions (a^i, a^j) is taken to be $x^i_{a^i} x^j_{a^j}$. If i and j were two gene loci of the same organism, this would be in general incorrect. However, an examination of our proofs suggests that a weaker assumption could be used: Namely, if both a^i and a^j are present (i.e., if $x^i_{a^i} > 0$ and $x^j_{a^j} > 0$), then the combination (a^i, a^j) is also present (i.e., its proportion is positive).

3. At each period, the dynamic process affects the two characteristics independently. As far as mutations go, this is natural. However, selection operates at the level of the organism rather than the single gene. When the proportion of the combination (a^i, a^j) is given by the product $x^i_{a^i} x^j_{a^j}$, it makes no difference. But difficulties arise once this independence no longer holds. Still, note that $x^i_{b^i} > 1 - \varepsilon$ and $x^j_{b^j} > 1 - \varepsilon$ imply that the proportion of the pair (b^i, b^j) must be at least $1 - 2\varepsilon$, regardless of the degree of interdependence between i and j. This argument suggests that our results may well hold in this more general setup.

5.2. Other Selection Dynamics

One straightforward generalization of our dynamics assumes that selection satisfies $Q\left[\widetilde{\omega}^i_{q(i)} \in B^i \mid \omega\right] \geq \beta$ instead of (2.6); i.e., at least one better action has positive probability of being chosen (rather than each one of them). It seems that the result will remain unchanged.[49] Another possible generalization, which also appears not to affect the result, assumes that the probability of each individual in $M(i)$ being chosen goes to zero as the population size $|M(i)|$ goes to infinity (without it necessarily being of the order of $1/|M(i)|$ as in (2.4)).[50] Other dynamics to consider—like the "replicator dynamics"—may have the selection take

[49] 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.

[50] One then needs to work with expressions like $E\left[p\left(Y^i\right)\right]$ instead of $E\left[Y^i\right]$, where $p\left(Y^i\right)$ is the probability that a non-b^i-individual is chosen when their proportion in the population is Y^i.

into account the actual payoff differences, rather than just their sign.[51] Further variants should also be analyzed; hopefully, a precise characterization of the class of dynamics that yield the backward induction outcome will be obtained.

One modification[52] makes selection choose better actions with probabilities that are proportional to their current proportions in the population. Thus, for instance, if the best-reply action c^i is currently played by k individuals, then the probability that a chosen non-c^i individual will switch to c^i by selection is $\sigma k/m^i$ (rather than $\sigma / |B^i|$). When k is small, this probability becomes low; in particular, if c^i is not currently present in the population, then selection cannot introduce it.[53] Such dynamics may be more appropriate in imitation-type models (where the "visibility" of an action depends on its prevalence in the population). Note also the property that $x^i_{a^i}$, the proportion in the population of a better action a^i, increases by selection at a rate that is proportional to $x^i_{a^i}$.

Formally, we weaken (2.6) to[54]

$$Q\left[\widetilde{\omega}^i_{q(i)} = a^i \mid \omega\right] \geq \beta x^i_{a^i}(\omega) \text{ for each } a^i \in B^i. \tag{5.1}$$

It turns out that our result continues to hold.

Theorem 5.1. *The results of the Main Theorem 3.2 hold also for dynamic processes satisfying (2.3), (2.4), (2.5), (5.1), (2.7) and (2.8).*

Proof. It can easily be checked that selection—i.e., (2.6) or (4.3)—is used in the proofs of Subsection 4.3 only in Step 3 of the Proof of Proposition 4.5. Replacing (2.6) by (5.1) implies that $P\left[\widetilde{\omega}^i_{q(i)} = b^i \mid \omega\right] \geq \beta(1 - Y^i)$ for every $\omega \in \Omega$ with $L^i R^i = 1$ (since the proportion of b^i in the population is $1 - Y^i$). Therefore the argument of Step 3 yields

$$P\left[\widetilde{Y}^i < Y^i\right] \geq \beta\gamma_1 E\left[Y^i\left(1 - Y^i\right) L^i R^i\right],$$

[51] Large populations may then decrease the effect of selection, since the difference in payoff due to one individual is small.

[52] Suggested by Ilan Eshel.

[53] A requirement suggested by Karl Schlag.

[54] Notice that (5.1) is more general than the description in the previous paragraph (since we only require $\geq$ rather than $=$).

which, by (4.5) for i, implies

$$E\left[Y^i\left(1-Y^i\right)L^iR^i\right]=O\left(\mu\right). \tag{5.2}$$

Next, we claim that the probability that Y^i is in a neighborhood of 1 is low.

Lemma 5.2. *There exist constants $\eta>0$ and $c>0$ such that*

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

Proof. We have

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

by (2.4) and (2.7). Next,

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

for an appropriate constant $c_1>0$, where we have used (2.4) and (2.8) (selection can increase Y^i only when $L^i=0$), and then $P\left[L^i=0\right]=O\left(\mu\right)$ by (4.4) for i (proved in Step 1).

Using the invariant distribution property yields

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

or

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

(where $c_2:=c_1/\left(\gamma_1\alpha_1\right)$). Let $\eta>0$ be small enough so that $c_2\left(m^i-k\right)/\left(k+1\right)\leq 1/2$ for all[55] $k>k_0:=\lfloor\left(1-2\eta\right)m^i\rfloor$. Then we get

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

[55] $\lfloor x\rfloor$ denotes the largest integer that is $\leq x$.

for all $k \geq k_0$ and thus

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

as claimed. ■

Proof of Theorem 5.1 (continued). (5.2) implies

$$\eta E\left[Y^i L^i R^i \mathbf{1}_{Y^i \leq 1-\eta}\right] = O\left(\mu\right),$$

where $\mathbf{1}_{Y^i \leq 1-\eta}$ is the indicator that $Y^i \leq 1-\eta$. Lemma 5.2 yields

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

which is at most $O\left(\mu\right)$ since $1/m \leq (1/\delta)\,\mu$. Adding the two estimates gives (4.6) for i, thus completing Step 3. The rest of the Proof of Proposition 4.5 is unchanged. ■

5.3. Extensions

This work is a first attempt to analyze basic evolutionary models in extensive form games. A number of directions for further study suggest themselves:

1. *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 are equal. It seems that a subset of BI, at times a strict subset, is obtained.

2. *Games with imperfect 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.

3. *Multiple agents and non-distinct populations*: This is discussed in Subsection 5.1, which suggests a number of relevant generalizations.

4. *Other selection dynamics*: These are discussed in Subsection 5.2.

References

Aumann, R. J. [1995], "Backward Induction and Common Knowledge of Rationality," *Games and Economic Behavior* 8, 6–19.

Aumann, R. J. [1998], "On the State of the Art in Game Theory," *Games and Economic Behavior* 24, 181–210.

Cressman, R. and K. H. Schlag [1997], "The Dynamic (In)Stability of Backwards Induction," mimeo.

Fudenberg, D. and D. K. Levine [1998], *The Theory of Learning in Games*, MIT Press.

Gale, J., K. Binmore and L. Samuelson [1995], "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior* 8, 56–90.

Hammerstein, P. and R. Selten [1994], "Game Theory and Evolutionary Biology," in *Handbook of Game Theory, Vol. II*, R. J. Aumann and S. Hart (eds.), North-Holland, 929–993.

Hofbauer, J. and K. Sigmund [1998], *Evolutionary Dynamics and Population Dynamics*, Cambridge University Press.

Kandori, M., G. Mailath and R. Rob [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 *Contributions to the Theory of Games II*, H. W. Kuhn and A. W. Tucker (eds.), *Annals of Mathematics Studies* 28, Princeton University Press, 193–216.

Nöldeke, G. and L. Samuelson [1993], "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior* 5, 425–454.

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

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.