

**האוניברסיטה העברית בירושלים**  
**THE HEBREW UNIVERSITY OF JERUSALEM**

---

**EVOLUTIONARY STABILITY IN  
GENERAL EXTENSIVE-FORM GAMES  
OF PERFECT INFORMATION**

**By**

**ZIBO XU**

**Discussion Paper # 631**

**Jan 2013**

**מרכז לחקר הרציונליות**

**CENTER FOR THE STUDY  
OF RATIONALITY**

---

**Feldman Building, Givat-Ram, 91904 Jerusalem, Israel**  
**PHONE: [972]-2-6584135      FAX: [972]-2-6513681**

**E-MAIL:                      [ratio@math.huji.ac.il](mailto:ratio@math.huji.ac.il)**

**URL:                      <http://www.ratio.huji.ac.il/>**

# Evolutionary stability in general extensive-form games of perfect information

Zibo Xu<sup>1</sup>

Center for the Study of Rationality, the Hebrew University of Jerusalem

1st December 2012

## **Abstract**

We consider a basic dynamic evolutionary model with rare mutation and a best-reply (or better-reply) selection mechanism. A state is evolutionarily stable if its long-term relative frequency of occurrence is bounded away from zero as the mutation rate decreases to zero. We prove that, for all finite extensive-form games of perfect information, only Nash equilibria are evolutionarily stable. We show that, in games where a player may play at more than one node along some path, even when the populations increase to infinity, there may be some evolutionarily stable states which are not part of the backward induction equilibrium component. We give a sufficient condition for evolutionary stability and show how much extra value is needed in the terminal payoffs to make an equilibrium evolutionarily stable.

---

<sup>1</sup>The author would like to acknowledge financial support by European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement No. 249159.

# 1 Introduction

Nash equilibrium is the most common solution concept in game theory. However, in a game with multiple equilibria, Nash equilibrium is silent on how to determine which equilibrium will be played. We would like to follow one evolutionary approach to answer this question. In an evolutionary dynamic process, a mutation can be viewed as a random perturbation; the selection condition requires that the current better/best strategies will be used more frequently in the future. In such a dynamic process with both mutation and selection, some Nash equilibria are more robust against persistent random perturbations (not just isolated ones), and are more likely to emerge in such ‘noisy’ environment in the long run. We say a state is evolutionarily stable if its long-term relative frequency of occurrence is bounded away from zero as the mutation rate decreases to zero. This notion is also called ‘stochastically stable’ by Foster and Young in [4], by Young in [14] and [15], ‘long-run equilibrium’ by Kandori et al. in [9], ‘in the support of the limit distribution’ by Samuelson in [10] and by Fudenberg and Levine in [3].

Such dynamic evolutionary models have been extensively studied under various selection mechanism for games in strategic form. (See [9], [14], [11] and many other literatures.) They show that certain Nash equilibria, e.g. the risk-dominant ones, are more stable than others. In general normal-form games, the evolutionary process brings inertia to the model: individuals’ observations are limited and imperfect, their understanding of the whole dynamic process is constrained, and replacing current strategies is costly. All these factors support the assumption that at each stage only a small fraction of individuals are adjusting their strategies simultaneously. The standard argument in games of normal form claims that the strategies that have proved to be successful up until today are likely to remain successful

for some time in the future.

For extensive-form games of perfect information, one needs to be careful when applying this argument. For individuals playing strategies at a disconnected node, when that node becomes connected by an individual from another population, the best/better strategies of the former population is entirely dependent on that single ‘connecting’ individual. If this ‘connecting’ individual leaves at the next stage, then the ones playing at the newly disconnected node will lose the incentive to adjust their strategies. Therefore, the set of better/best reply strategies can change at every stage. This is one basic proposition in the dynamics in extensive-form games.

Hart [8] proves that, in a basic and natural model, the backward induction equilibrium is the unique evolutionarily stable outcome, when the populations go to infinity, and the product of population size and mutation rate is bounded away from zero. This boundary condition can be interpreted as requiring that the expected number of mutation per generation does not go to zero. Gorodeisky [7] relaxes this constraint on the product of population size and mutation rate.

However, the papers listed above on best (or better) reply dynamics all concentrate on a specific type of perfect-information games: the population at each node is distinct. So, any player can never play at more than one node along any path of the game. We instead allow a player to play at more than one node along any path of the game. We use the basic dynamic model of population games proposed in Hart’s paper, which is essentially a most elementary process that provides for both selection and mutation. Our first conclusion is that, for all finite extensive-form games of perfect information, regardless of population size, only Nash equilibria can be evolutionarily stable. It is then a natural question whether the set of evolutionarily stable

equilibria contains non-backward-induction equilibria.

The starting point is that one should turn to the concept of *Nash equilibrium components*. Here, we adopt the conventional definition of a Nash equilibrium component: all equilibria in the same Nash equilibrium component give rise to identical outcome over terminal vertices, i.e., they differ only off the equilibrium path. Thus, a backward induction equilibrium component is a Nash equilibrium component whose final outcome is equivalent to the outcome of a backward induction equilibrium. In the case of large population, if a state is in a backward induction equilibrium component, then the distribution of a population at a disconnected node is dependent on how long the node is connected. When the population size is sufficiently large, we cannot conclude that all individuals in all populations converge to the (pure) backward induction equilibrium, even if the backward induction equilibrium component is the only evolutionarily stable result. (See Example 4 in [13] for more details.)

In [8], Hart also shows an example that equilibria other than the backward induction equilibrium may be evolutionarily stable, when the populations are fixed. However, he further proves that, if each player can play at only one node, then this no longer holds when the populations increase to infinity. Our result in Section 6 instead shows that if a player is allowed to play at multiple nodes, then the backward induction equilibrium component may not be the only evolutionarily stable outcome even when the populations go to infinity. Moreover, this example for large populations contains the minimum number of nodes for all finite extensive-form game of perfect information. For more details, please refer to the end of [13], where we discuss why such an example needs more than three nodes. The results hold for both the best reply and the better reply selection mechanisms.

The evolutionary approach can also explain the emergence of trust under bounded rationality modeled by a finite extensive-form game. Here we consider the case in which groups of agents play a game repeatedly where each group assumes the role of a player in the original extensive-form game. Our evolutionary model is compatible with the notion of ‘bounded rational behaviour’ introduced in [9]. That is, not all agents need to react instantaneously to their environment; when agents react, they react myopically; there is a small probability that agents change their strategies at random. Note that in our example, the evolutionarily stable Nash equilibrium component not containing the backward induction equilibrium is in fact Pareto-efficient. Thus, this evolutionary approach supports the argument that, under the condition of bounded rationality, players also select Pareto-efficient equilibria in the long run as well, even if they are not the backward induction equilibrium. Moreover, this Pareto-efficient result is achieved by trust built on a sufficiently long multiple-move game, since, as we have pointed out, every extensive-form game with no more than three nodes has the backward induction equilibrium component as its unique evolutionarily stable outcome. We shall discuss it in details in Section 6.

In Section 7, we generalise our example and provide a sufficient condition for the Nash equilibrium component different from the backward induction equilibrium component to be evolutionarily stable, regardless of population size. The theorem is in the context of two-player finite extensive-form generic games of perfect information, and it only needs a payoff condition and a structure condition roughly as follows. The payoff vector of the alternative Nash equilibrium component is sufficiently high, and the paths of the backward induction equilibrium and the alternative Nash equilibrium can be ‘properly’ separated. Further generalisation is also possible, as long as an individual

selects each better (or best) response strategy with a probability bounded from below. The proof needs several steps, some of which are quite subtle. So we may view the earlier result as a warm-up for this general theorem.

We can apply this sufficient condition for evolutionary stability to solve a problem of improving stability by increasing payoffs. It is straightforward to see that, when a payoff vector gives global best payoff to both players, it is the result of the backward induction equilibrium in the game, and this backward induction equilibrium component is evolutionarily stable for any population size. If we can only lift one payoff of a Nash equilibrium to be the global maximum and the other payoff to be a second best, under what circumstances will it still be enough to achieve evolutionary stability of the equilibrium component? We'll answer this question in Section 8.

Regarding the technical difficulties of the proofs, as Hart predicts in his paper, the complexity in multiple-move games arises since selection operates at the level of the players rather than the nodes. For instance, consider the case that the unique backward induction equilibrium path terminates somewhere in the middle of the game tree by player (population) I, and this backward induction equilibrium is evolutionarily stable. Then at some time in the induced Markov process very few individuals in population  $M(1)$  are playing below that backward induction equilibrium path. As the sum of individuals playing at each node is not fixed, a small change in the proportions of strategies in population  $M(1)$  triggered by mutation may bring a huge effect to the local proportions of strategies in population  $M(1)$  at some descendant node below that backward induction equilibrium path. That may further destabilise the backward induction equilibrium. While we regard the central issue in Hart's model as 'connectedness' of path, in this generalised model we are more concerned with the 'dynamic of proportions'.

## 2 The Model

We consider a generic finite  $N$ -player game  $\Gamma$  in extensive form with perfect information. In the generic assumption, no player obtains the same utility at any two terminal vertices. Thus, the backward induction equilibrium in  $\Gamma$  is unique. Without loss of generality, we assume that  $\Gamma$  does not contain any chance move. We denote the strategy set of player  $i$  to be  $A^i$ ,  $i = 1, \dots, N$ . Put  $A := \prod_{1 \leq i \leq N} A^i$ . For every  $a$  in  $A$  players receive payoffs determined by  $a$  in  $\Gamma$ . Without loss of generality, we consider *reduced-form strategies* in this paper. That is, we view two strategies of a player the same strategy if it is impossible to reach different payoff to this player from these two strategies. For instance, given the game in Figure 1, conventionally, for a strategy of player I where she plays the action  $a_1$  at the top node, that strategy must also specify the action she would play at the bottom node. We, however, do not specify it, as it is impossible to reach the bottom node in that case. Hence, in this one-player game, there are only three strategies  $a_1$ ,  $a_2$  and  $a_3$  in our definition.

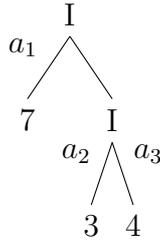


Figure 1: Example to the definition of strategies

For each player  $i$ , the set of mixed strategies of player  $i$  is a unit simplex on  $A^i$ , and we denote it by

$$X^i := \Delta(A^i) = \{x^i \in \mathbb{R}_+^{A^i} : \sum_{a^i \in A^i} x_{a^i}^i = 1\}.$$

We then denote the set of mixed strategy profiles by  $X := \prod_{1 \leq i \leq N} X^i$ . Let  $u^i : A \rightarrow \mathbb{R}$  be the payoff function of player  $i$ .

We consider the following *population game* associated to  $\Gamma$ . Assume that for each  $i$  with  $1 \leq i \leq N$  there is a non-empty population  $M(i)$  of individuals playing the game in the role of player  $i$ . For simplicity, we put  $|M(i)| = m$  for all  $i$ . We define a state  $w$  to be the collection of pure strategies of each individual in each population, i.e.,  $w = (w^i)_{1 \leq i \leq N}$  where  $w^i := (w^i(q))_{q \in M(i)}$  and  $w^i(q) \in A^i$  for each  $i$ . Let the state space  $W$  be  $\prod_{1 \leq i \leq N} (A^i)^{M(i)}$ . At a state  $w$ , for each  $i$ , let  $m_{a^i}(w)$  be the number of individuals playing strategy  $a^i \in A^i$ , and let  $x_{a^i}^i(w)$  be the proportion of population  $M(i)$  that plays the strategy  $a^i$ , i.e.,  $x_{a^i}^i(w) = m_{a^i}(w)/m$ . We put  $x^i(w) := (x_{a^i}^i(w))_{a^i \in A^i}$ , and thus we may view each state  $w$  as an  $N$ -tuple mixed strategy  $x(w) = (x^i(w))_i$  in  $X$ .

We define a discrete-time Markov chain  $(Z_t)_{t \geq 0}$  such that each  $Z_t$  is a random variable valued in  $W$ . Note that for a rigorous analysis, one needs to define first the sample space  $\Omega := W^{\mathbb{N}}$ , which is the necessity for the definition of this Markov chain and its transition probabilities. We shall not refer to it explicitly, when the context allows. Given the population size  $m$  and mutation rate  $\mu$ , the transition probabilities  $P_{\mu, m}$  of this Markov chain specify the probability that  $Z_{t+1}$  equals a state  $\tilde{w}$  given a history  $Z_1 = w_1, \dots, Z_t = w_t$ , i.e.,  $P_{\mu, m}(Z_{t+1} = \tilde{w} | Z_1 = w_1, \dots, Z_t = w_t)$ , for each  $t \geq 0$ . We further assume  $(Z_t)_{t \geq 0}$  to be a stationary Markov chain dictated by a one-step transition probability matrix  $Q_{\mu, m}$ , i.e.,

$$P_{\mu, m}(Z_{t+1} = \tilde{w} | Z_1 = w_1, \dots, Z_t = w_t) = Q_{\mu, m}(\tilde{w} | w_t),$$

for every  $w_1, \dots, w_t, \tilde{w}$  in  $W$  and  $t = 1, 2, \dots$ . If context allows, we drop the subscript  $\mu$  and  $m$  of  $P$  and  $Q$ ; for a strategy  $a$  of some player  $i$ , at stage  $t$ , we also abbreviate  $m_a(Z_t)$  by  $m_a(t)$  or simply  $A(t)$ , which is the

random number of individuals in population  $M(i)$  playing the strategy  $a$  in the Markov chain at stage  $t$ . We view  $Q$  as a transition rule concerning ‘selection’ and ‘mutation’ described in the following basic model with a *better-reply selection mechanism*, which was introduced by Hart in [8].

- Transitions are conditional independence over populations, i.e.,

$$Q(\tilde{w}|w) = \prod_{1 \leq i \leq N} Q(\tilde{w}^i|w).$$

- At each stage, for each  $i$  with  $1 \leq i \leq N$ , an individual  $q^i$  in  $M(i)$  is chosen with probability  $1/m$ .
- All individuals in  $M(i)$  except  $q^i$  do not change their strategies.
- The chosen  $q^i$  undergoes mutation, selection or no change, with probability  $\mu$ ,  $\sigma$  and  $1 - \mu - \sigma$ , respectively.
- Conditional on mutation, the individual  $q^i$  chooses a random strategy, i.e.,  $\tilde{w}^i(q^i) = a^i$  with probability  $1/|A^i|$  for each  $a^i$  in  $A^i$ .
- Conditional on selection, we define a set of ‘better strategies’, namely

$$BT^i(q^i, w) := \{a^i \in A^i : u^i(a^i, w^{-i}) > u^i(w^i(q^i), w^{-i})\}, \quad (2.1)$$

where  $w^{-i}$  indicates the collection of pure strategies of all individuals in all populations other than  $M(i)$ . If no ambiguity, we abbreviate the notation  $BT^i(q^i, w)$  by  $BT^i$ . If  $BT^i$  is not empty, then the new strategy  $\tilde{w}^i(q^i)$  of  $q^i$  is a randomly chosen better strategy, i.e.,  $\tilde{w}^i(q^i) = a^i$  with probability  $1/|BT^i|$  for each  $a^i \in BT^i$ . If  $BT^i$  is empty, then there is no change in the strategy of  $q^i$ .

- For the case of no change, the strategy of  $q^i$  does not change:  $\tilde{w}^i(q^i) = w^i(q^i)$ .

We can also adopt the *best-reply selection mechanism* in the model above. Consider the case that at state  $w$  an individual  $q^i$  in  $M(i)$  is chosen and selection takes effect. We define a set of ‘best strategies’, namely

$$BS^i(w) := \arg \max_{a^i \in A^i} u^i(a^i, w^{-i}). \quad (2.2)$$

If  $w^i(q^i) \in BS^i(w)$ , then there is no change in  $q^i$ 's strategy:  $\tilde{w}^i(q^i) = w^i(q^i)$ ; otherwise, the new strategy  $\tilde{w}^i(q^i)$  is a randomly chosen best strategy, i.e.,  $\tilde{w}^i(q^i) = a^i$  with probability  $1/|BS^i(w)|$  for each  $a^i \in BS^i(w)$ . The only difference between the models with best-reply and better-reply mechanism is in this selection procedure.

### 3 The Main Results

To give the formal definition of evolutionary stability, we study the behaviour of the steady-state analysis and limiting distribution of the Markov chain  $(Z_t)_{t \geq 0}$  with the one-step transition probability matrix  $Q_{\mu, m}$  generated from the transition process in Section 2. We first consider the selection mechanism of better reply applied in the model. It is straightforward to check that, for every finite  $m$  and every  $\mu > 0$ , the induced Markov chain is irreducible and aperiodic. Hence, for every pair of population size  $m$  and mutation rate  $\mu$ , there exists a unique invariant distribution  $\pi_{\mu, m} \in \Delta(W)$  on  $W$  such that  $\pi_{\mu, m} = \pi_{\mu, m} Q_{\mu, m}$ , or

$$\pi_{\mu, m}[\tilde{w}] = \sum_{w \in W} \pi_{\mu, m}[w] Q_{\mu, m}(\tilde{w}|w)$$

for every  $\tilde{w} \in W$ .

**Lemma 3.1** *Given an one-step transition probability matrix  $Q_{\mu, m}$  defined in Section 2, for a fixed population size  $m$ , the limit distribution  $\lim_{\mu \rightarrow 0} \pi_{\mu, m}$  exists.*

**Proof.** It immediately follows from Theorem 3.1 in [15].  $\square$

For a fixed population size, a state  $w$  in  $W$  is called *m-evolutionarily stable* if its invariant probability  $\pi_{\mu,m}[w]$  does not go to zero as  $\mu$  decreases to zero, i.e.,  $\lim_{\mu \rightarrow 0} \pi_{\mu,m}[w] > 0$ . We call a state  $w$  *evolutionarily stable for large populations* if  $\liminf_{m \rightarrow \infty} \lim_{\mu \rightarrow 0} \pi_{\mu,m}[w] > 0$ . Given a limiting process that  $m$  increases to infinity,  $\mu$  decreases to zero, and that  $\mu m$  is always in a set  $R \subseteq \mathbb{R}^+$ , we call a state  $w$  evolutionarily stable in this limiting process if  $\pi_{\mu,m}[w]$  is bounded away from zero throughout this process. All definitions above can be naturally extended to subsets included in the state space. So a Nash equilibrium component is evolutionarily stable if and only if a subset in it is evolutionarily stable. For every set of mixed strategy profiles  $Y \subseteq X$ , if no ambiguity, we abbreviate  $\pi_{\mu,m}[w \in W : x(w) \in Y]$  by  $\pi_{\mu,m}[Y]$ . Recall the game we consider is generic. Denote the unique backward induction equilibrium in  $\Gamma$  by  $BI = (s^i)_{1 \leq i \leq N}$ , and denote the  $\epsilon$ -neighbourhood of  $BI$  by

$$BI_\epsilon := \{x \in X : x_{s^i}^i \geq 1 - \epsilon \forall 1 \leq i \leq N\}.$$

We shall first prove in Section 4 that only Nash equilibria are evolutionarily stable regardless of population size.

The combined results from Hart [8] and Gorodeisky [7] are

$$\forall \epsilon > 0, \quad \lim_{\mu \rightarrow 0, m \rightarrow \infty} \pi_{\mu,m}[BI_\epsilon] = 1,$$

in the case that each player can only play at one node in a game. Thus, in the context of only two players, their conclusion can only be applied to a two-node game.

Recall that all equilibria in the same Nash equilibrium component give rise to identical outcome over terminal vertices, i.e., they differ only off the equilibrium path. We denote the backward induction equilibrium component

by  $BC$ . We then denote  $NC$  to be the set of Nash equilibria other than those included in the backward induction equilibrium component, and  $NE$  to be the set of Nash equilibria. It follows that  $NC = NE \setminus BC$ . We give a proof of the following theorem by Example 1 in Section 6.

**Theorem 3.2** *There exists a two-player game of perfect information and a positive number  $c$  such that for all finite  $m > 1$   $NC$  is  $m$ -evolutionarily stable and*

$$\lim_{\mu \rightarrow 0} \pi_{\mu, m}[NC] \geq c.$$

*Thus,  $NC$  is evolutionarily stable for large populations and*

$$\liminf_{m \rightarrow \infty} \lim_{\mu \rightarrow 0} \pi_{\mu, m}[NC] \geq c. \tag{3.1}$$

This theorem applies to both the best-reply selection mechanism and the better-reply selection mechanism.

## 4 Evolutionary Stability of Nash Equilibria

In this section, we prove that, under the best-reply or better-reply selection mechanism, only Nash equilibria are evolutionarily stable regardless of population size, after we introduce some notations and concepts.

### 4.1 Definitions, Notations and Operations

Suppose an extensive-form game  $\Lambda$  is given. As in the assumption in Section 2, without loss of generality, we assume no chance move in  $\Lambda$ . Recall that in such a game tree, each node is a nonterminal vertex, where one of the players takes a move. Given an extensive-form game  $\Lambda$ , we denote  $N(\Lambda)$  to be the

set of all nodes in  $\Lambda$ . For each node  $n$  in  $\Lambda$ , we denote  $d^\Lambda(n)$  and  $p^\Lambda(n)$  to be the set of all successor nodes and all predecessor nodes of  $n$ , respectively. Note that both  $d^\Lambda(n)$  and  $p^\Lambda(n)$  can be empty sets. We denote the set of immediate successor nodes and the set of immediate predecessor nodes of  $n$  by  $id^\Lambda(n)$  and  $ip^\Lambda(n)$ , respectively. We also denote  $\alpha^\Lambda(n)$  to be the player who moves at node  $n$ . If no ambiguity, we leave out the superscripts  $\Lambda$  in the above notations.

Consider a generic finite game  $\Gamma$  in extensive form of perfect information with the associated population game. Recall that a state describes the distribution of all individual's strategies in  $\Gamma$ . Given an individual  $x$ , we denote the population it belongs to by  $po(x)$ . For each node  $n$  in  $\Gamma$ , denote  $\beta(n)$  to be the corresponding population of  $\alpha^\Gamma(n)$ . Given a population  $i$ , we denote all nodes in  $\Gamma$  where the corresponding player of  $i$  moves by  $AN(i)$ . Given a state  $w$ , we say a node  $n$  is *connected* in  $\Gamma$  at  $w$  if at every  $n'$  in  $N(\Gamma)$  on the path from the root of  $\Gamma$  to  $n$  there is at least one individual  $x$  in  $\beta(n')$  whose move is towards  $n$ . When a node  $n$  in  $\Gamma$  is connected at state  $w$ , we write  $R_w(n) = 1$ ; otherwise  $R_w(n) = 0$ . Given a node  $n$  in  $\Gamma$ , we say an individual is *active* at node  $n$ , if at every predecessor node of  $n$  where that individual takes a move, he chooses the move towards the node  $n$ . That is, it moves towards  $n$  at all nodes in  $AN(po(x)) \cap p^\Gamma(n)$ . If player  $\alpha(n)$  does not play at any node in  $p(n)$ , then every individual in  $\beta(n)$  is active at  $n$ . A set of moves from an individual can be viewed as a (reduced) strategy if the set only specifies all the moves at the nodes at which the individual is active. For simplicity, we regard a move at a node as a directed edge in a game tree.

Given a game  $\Lambda$ , we denote the set of nodes which do not have any successor node in  $\Lambda$  by  $L(\Lambda) := \{n \in N(\Lambda) : d(n) = \emptyset\}$ . So, for each node  $n$  in  $L(\Lambda)$ , all children of  $n$  are terminal vertices in  $\Lambda$ . We introduce below two

operations on a node  $n$  in the set  $L(\Lambda)$  in a genetic extensive-form game  $\Lambda$  of perfect information.

- **Pruning  $\Lambda$  at node  $n$ .** [Elimination of dominated payoff vectors]

As  $\Lambda$  is a generic game, there is a terminal vertex  $max_n$  which gives player  $\alpha(n)$  the maximum payoff conditional on reaching node  $n$ . We cut all edges from  $n$  to its terminal vertices, and then transform node  $n$  to a terminal vertex with the payoff vector the same as of the original terminal vertex  $max_n$ . We denote such modified game by  $\Lambda^{\sim n}$ .

- **Cutting  $\Lambda$  at node  $n$ .** [Remove of node  $n$ ]

We delete the edge from  $ip(n)$  to  $n$ , and check whether  $ip(n)$  still has any child. If after deletion there is neither successor node nor terminal vertex associated to  $ip(n)$ , then we also delete the edge from  $ip(ip(n))$  to  $ip(n)$ , and check whether  $ip(ip(n))$  has any child (a successor node or terminal vertex). This process continues until we find a predecessor node  $\bar{n}$  of  $n$  with survival children after deletion. We then normalise the probability of the edges from this predecessor node  $\bar{n}$  to all its remaining children. We denote such modified game by  $\Lambda^{-n}$ .

## 4.2 Constraints on the Histories

Given a state  $w$  in  $W$ , if  $w$  is consistent with  $NE$ , we simply say  $w$  is in  $NE$  and write  $w \in NE$ . (The rigorous notation should be  $x(w) \in NE$ ; see the notation  $\pi_{\mu,m}[Y]$  introduced after the proof of Lemma 3.1.)

Given a finite generic extensive-form game  $\Gamma$  of perfect information and an associated Markov chain defined in Section 2, for any initial state  $w$ , we shall prove the existence of a finite transition history  $(w_0, w_1, \dots, w_k)$  with positive probability such that  $w_0 = w$  and  $w_k \in NE$ . This is the central step

to show that only  $NE$  can be evolutionarily stable in the stochastic process.

We shall construct such a finite history under the following constraints.

1. **No-mutation Constraint:** At each stage, the chosen individual in each population either moves to pick his best-reply strategy or keep using his best-reply strategy against the current distribution of other populations. In other words, we do not consider the case of mutation and the case of ‘no change’ which is assigned with probability  $1 - \mu - \sigma$  at each stage.
2. **Disconnected-moves-unchanged Constraint:** For all stages  $t \geq 0$ , when a chosen individual replaces his strategy with a best-reply strategy at stage  $t$ , for all nodes  $n$  with  $R_{w_{t+1}}(n) = R_{w_t}(n) = 0$  and the property that the individual is active at  $n$  in both stage  $t$  and  $t + 1$ , his move at  $n$  will be the same in these two stages. In other words, when an individual changes his current strategy to his best reply, if both two strategies consist of a move at a disconnected node  $n$ , then his move at  $n$  is unchanged.
3. **Fixed-move Constraint:** It is concerning about a node  $n$  in  $\Gamma$  and an edge  $e$  departing from  $n$ . This constraint applied at some stage  $t$  specifies the following requirement. From stage  $t$  on, whenever node  $n$  is disconnected and an individual turns active at node  $n$ , his move at  $n$  is always the edge  $e$ . This is possible, since, when an individual changes to his best-reply strategy, his moves on disconnected nodes do not affect his payoff.

### 4.3 Construction of a Finite History Ending at $NE$

Given any initial state  $w_0$ , we construct a finite history  $h = (w_0, w_1, \dots, w_t)$  in the stochastic process under the best-reply selection mechanism such that  $w_t \in NE$ .

To this end, we shall define below a sequence of extensive-form games  $(\Gamma_{n_1}, \dots, \Gamma_{n_r})$  where index  $n_i \in N(\Gamma)$  for all  $i$  with  $1 \leq i \leq r$  and  $N(\Gamma_{n_l}) \supset N(\Gamma_{n_{l+1}})$  for all  $l$  with  $1 \leq l < r$ . Moreover, we require that  $\Gamma_{n_1} = \Gamma$  and that  $\Gamma_{n_r}$  is a one-node game. Furthermore, for  $i \geq 1$ ,  $\Gamma_{n_{i+1}}$  is obtained from pruning or cutting at a node  $n_{i+1}$  in  $L(\Gamma_{n_i})$  in  $\Gamma_{n_i}$ . Each  $\Gamma_{n_i}$  will be associated with a finite sequence of states  $(w_i^0, w_i^1, \dots, w_i^{k(i)})$  under the best-reply selection mechanism. In doing so, we could fix a specific move at a node in turn in a Nash equilibrium, and from some stage on all individuals active at that node could always follow that move in the history  $h$ .

Let  $w_1^0$  be  $w_0$ . Given any  $i \geq 1$ , the game  $\Gamma_{n_i}$  and state  $w_i^0$ , to obtain  $\Gamma_{n_{i+1}}$ , we pick one node  $n_{i+1}$  in  $L(\Gamma_{n_i})$ . Under the best-reply selection mechanism, we then check whether there is a finite transition process  $y = (y_0, y_1, \dots, y_l)$  with the property  $y_0 = w_i^0$  and  $R_{y_l}(n_{i+1}) = 1$  which satisfies Constraint 1 and Constraint 2 as well as the Constraints 3 specified during the construction of  $(\Gamma_{n_1}, \dots, \Gamma_{n_i})$ . Note that if  $n_{i+1}$  is connected at  $w_i^0$ , then the answer is trivially yes.

**Case I:** No such finite process exists. We cut  $\Gamma_{n_i}$  at node  $n_{i+1}$  and obtain  $\Gamma_{n_{i+1}} := \Gamma_{n_i}^{-n_{i+1}}$ . Let  $w_{i+1}^0$  be  $w_i^0$ .

**Case II:** There exists one such finite process. We adopt one such process and index it as  $(w_i^0, w_i^1, \dots, w_i^l)$ . We denote the equivalent node (or terminal vertex) in  $\Gamma$  to the terminal vertex  $max_{n_{i+1}}$  in  $\Gamma_{n_i}$  by  $\bar{m}_{n_{i+1}}$ . (Recall that  $max_{n_{i+1}}$  gives player  $\alpha(n_{i+1})$  the maximum payoff in  $\Gamma_{n_i}$  conditional on reaching node  $n_{i+1}$ . Note that  $\bar{m}_{n_{i+1}}$  is a terminal vertex in  $\Gamma$ , or there

exists a game  $\Gamma_l$  such that  $l < i$  and  $\Gamma_l$  is pruned at node  $\bar{m}_{n_{i+1}}$ .) Suppose at state  $w_i^l$  there are  $q$  active individuals at node  $n_{i+1}$  whose move is not towards  $\bar{m}_{n_{i+1}}$ . In the next  $q$  stages, we still use the best-reply selection mechanism, and we always pick one such individual in population  $\beta(n_{i+1})$  and make it to adopt a best reply against the current distribution of all other populations. Meanwhile, we can make node  $n_{i+1}$  always connected in these  $q$  stages: when there is only one individual moving towards  $n_{i+1}$  at some node  $n' \in (p(n_{i+1}) \setminus AN(\beta(n_{i+1})))$  in the path from the root of  $\Gamma$  to  $n_{i+1}$ , we always avoid picking this individual in population  $\beta(n')$ . There is no further restriction on the transition process  $(w_i^{l+1}, w_i^{l+2}, \dots, w_i^{l+q})$  under the best-reply selection mechanism during these  $q$  stages.

Denote the stage when reaching the state  $w_i^{l+q}$  by  $t^{n_{i+1}}$ . For  $\Gamma$ , we apply Constraint 3 on node  $n_{i+1}$  and the edge from  $n_{i+1}$  to  $\bar{m}_{n_{i+1}}$  at stage  $t^{n_{i+1}}$ . We prune  $\Gamma_{n_i}$  at node  $n_{i+1}$  and obtain  $\Gamma_{n_{i+1}} := \Gamma_{n_i}^{\sim n_{i+1}}$ , and let  $w_{i+1}^0$  be  $w_i^{l+q}$ .

In either case, at state  $w_{i+1}^0$ , we pick one node  $n_{i+2}$  in  $L(\Gamma_{n_{i+1}})$  in  $\Gamma_{n_{i+1}}$  and check the existence of a qualified finite transition process as for game  $\Gamma_{n_i}$ . This process continues until we obtain an one-node game  $\Gamma_{n_r}$  and the associated finite transition process  $(w_r^0, \dots, w_r^{k(r)})$ , where  $\Gamma_{n_r}$  only contains the root of  $\Gamma$ .

We stick all finite transition history together in order together with the initial state  $w_0$ , and leave out  $w_i^0$  for all  $i$  with  $0 < i \leq r$ . We denote such transition history by  $h$ . Recall that in the construction process, in the sequence of induced extensive-form games  $(\Gamma_{n_1}, \dots, \Gamma_{n_r})$ ,  $N(\Gamma_{n_l}) \supset N(\Gamma_{n_{l+1}})$  for all  $l$  with  $1 \leq l < r$ . Recall also that the associated sequence  $(w_i^0, w_i^1, \dots, w_i^{k(i)})$  to each game  $\Gamma_{n_i}$  is finite. Because the population size  $m$  and  $|N(\Gamma)|$  are both

finite, we may infer that  $h$  is finite. The transition history  $h$  is in the form

$$h = \left( w_0, w_{l_1}^1, \dots, w_{l_1}^{k(l_1)}, w_{l_2}^1, \dots, w_{l_2}^{k(l_2)}, \dots, w_r^1, \dots, w_r^{k(r)} \right)$$

where, for any  $i > 0$  and any  $j$  with  $l_i < j < l_{i+1}$ ,  $w_j^1$  is not defined in the above process. That is, we cut  $\Gamma_{n_j}$  at node  $n_{j+1}$ , or we prune  $\Gamma_{n_j}$  at node  $n_{j+1}$  but the associated sequence of states  $(w_j^t)_{0 < t \leq k(j)}$  is empty.

#### 4.4 Results and Proofs

In the game  $\Gamma$ , we call a node  $n$  an *active node* if it is pruned in the construction of finite history  $h$ . Thus, a node  $n$  is inactive if there is a predecessor node of  $n$  or  $n$  itself cut in the construction of  $h$ .

For the constructed finite history  $h = (w_i)_i$ , we firstly observe that each one-step transition from  $w_i$  to  $w_{i+1}$  is consistent with the best-reply selection mechanism. Thus, the finite  $h$  happens with positive probability. Note also that Constraint 2 and Constraint 3 are only concerning with disconnected nodes.

**Lemma 4.1** *Given the constructed finite history  $h$  under the best-reply selection mechanism, for each active node  $n_i$  in  $\Gamma$ , every active individual at node  $n_i$  takes a move towards some fixed node in  $id(n_i)$  from some stage on. Therefore, from that stage on, when  $n_i$  is connected, for an individual  $q$  in  $\beta(n_i)$ , if the best-reply strategy  $s$  of  $q$  consists a move on node  $n_i$ , then it must be towards that fixed node in  $id(n_i)$ .*

**Proof.** Recall the definition of  $\bar{m}_{n_i}$  and  $t^{n_i}$  in Case II in the above construction process when the active node  $n_i$  is considered in the game  $\Gamma_{n_{i-1}}$ . We show below by induction that every active individual at node  $n_i$  takes a move towards  $\bar{m}_{n_i}$  from stage  $t^{n_i}$  on.

For active nodes in  $L(\Gamma)$ , it is straightforward to see that the above statement holds under the Constraints 1, 2 and 3.

Now suppose that node  $n$  is labeled as  $n_i$  in the construction of  $h$  and  $id(n) \neq \emptyset$ . Given a node  $n$  in  $\Gamma$ , denote the subgame rooted at  $n$  by  $\Gamma(n)$ . The **induction hypothesis** is that for all active nodes  $n_r$  with  $r < i$ , in particular for all active nodes  $n_r$  in  $\Gamma(n_i)$ , every active individual at node  $n_r$  takes a move towards  $\bar{m}_{n_r}$  from stage  $t^{n_r}$  on. We first prove the following claim.

**Claim:** At stage  $t^{n_i}$  all active individuals at the active node  $n_i$  are taking a move towards  $\bar{m}_{n_i}$  at state  $w_{i-1}^{l+q}$ . (It is possible that no individual is active at node  $n_i$  at that time.)

**Proof of the claim:** Suppose the stage when it reaches the state  $w_{i-1}^l$  in the construction (see Case II) is  $\bar{t}$ , then  $\bar{t} > t^{n_j}$  for all active nodes  $n_j$  with  $j < i$ . Thus, by induction hypothesis, every active individual at any active node  $n_j$  in  $\Gamma(n_i)$  takes a move towards  $\bar{m}_{n_j}$  for any stage  $t \geq \bar{t}$ . Note that if a node  $n_r$  is inactive, then after any stage  $t^{n_j}$  where  $j > r$  and  $n_j$  is active, node  $n_r$  is always disconnected. Recall that  $\bar{m}_{n_j}$  is the backward-induction payoff vector if all inactive nodes in  $\Gamma(n_j)$  are excluded. It follows that, at all stages  $t \geq \bar{t}$  when  $n_i$  is connected, the local best-reply strategy of an individual in  $\beta(n_i)$  in  $\Gamma(n_i)$  consists of a move towards  $\bar{m}_{n_i}$ . Consider those  $q$  active individuals at node  $n_i$  whose move is not towards  $\bar{m}_{n_i}$ . We know from Case II in the construction of  $h$  that, when anyone of those  $q$  active individuals is chosen,  $n_i$  is connected. The chosen individual either selects a best-reply strategy which does not consist of a move at  $n_i$ , or it picks a strategy with moves in  $\Gamma(n_i)$ . That completes the proof of the above claim.

With the same argument above, for any stage  $t > t^{n_i}$ , when  $n_i$  is connected, an individual in  $\beta(n_i)$  has a best-reply strategy  $s$  with a move towards

$\bar{m}_{n_i}$ , if that  $s$  involves a move in  $\Gamma(n_i)$ . When  $n_i$  is disconnected, Constraint 2 excludes the case that an individual's move at  $n_i$  drifts from  $\bar{m}_{n_i}$  to another node in  $id(n_i)$ . When an individual's best reply consists of a move at  $n_i$  but with  $n_i$  disconnected, Constraint 3 requires that move towards  $\bar{m}_{n_i}$ . We have completed the proof of the lemma.  $\square$

**Lemma 4.2** *Suppose the constructed finite history  $h$  under the best-reply selection mechanism is  $(w_0, w_1, \dots, w_k)$ , then  $w_k$  is in  $NE$ .*

**Proof.** Consider the last active node, the root of  $\Gamma$  in the construction of  $h$ . At state  $w_k = w_r^{k(r)}$ ,  $\Gamma_{n_r}$  is pruned at the root. From Lemma 4.1, it follows that all individuals in all populations have no incentive to move and they are all playing their best-reply strategies at the last state  $w_k$  in  $h$ . So  $w_k \in NE$ .  $\square$

**Comment:** To deduce the result that the backward induction equilibrium is always evolutionarily stable, Hart proved in (3.4) in [8] that, for the case in  $\Gamma$  that each player only plays once,  $Q_0$  (the limit of one-step transition matrix when  $\mu$  goes to zero) is *acyclic*. That is, there are no states  $w_0, w_1, \dots, w_t, \dots, w_T$  in  $W$  with the property that  $w_t \neq w_{t-1}$  and  $Q_0(w_t|w_{t-1}) > 0$  for every  $t = 1, \dots, T$  and  $w_T = w_0$ . However, for the model we consider here,  $Q_0$  is not always acyclic. Consider the following two-player extensive-form game in Figure 2.

Suppose  $|M(1)| = |M(2)| = m = 2n + 1$  for some natural number  $n$ . A state  $w$  is described in the form  $(m_{a_2}, m_{a_3}; m_{b_2})$ . Now consider a finite sequence of states  $(w_t)_{0 \leq t \leq 4}$  satisfying the following process:  $(0, 1; n)$ ,  $(0, 1; n+1)$ ,  $(1, 1; n+1)$ ,  $(1, 1; n)$ ,  $(0, 1; n)$ . Then  $w_0 = w_4$ , and one can check that, for each  $t = 1, \dots, 4$ ,  $Q_0(w_t|w_{t-1}) > 0$ . Hence,  $Q_0$  is not acyclic. To some extent, the case of non-acyclic  $Q_0$  constitutes a substantial difficulty in the study of our models.

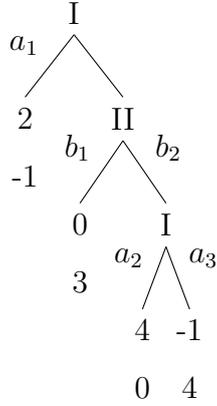


Figure 2: A game with non-acyclic  $Q_0$

**Lemma 4.3** *In a finite  $N$ -player game in extensive form of perfect information, we suppose that each player has no more than  $S$  pure strategies. Under the selection mechanism of best reply or better reply, if a state  $w$  can enter  $NE$  in finite stages with probability  $p$ , then it can also enter  $NE$  in no more than  $(S^m)^N$  stages with probability no less than  $p$ , where  $m$  is the population size.*

**Proof.** We firstly observe the size of the state space  $W$  is at most  $(S^m)^N$ . For any partial transition history  $(w_0, w_1, \dots, w_t)$  with positive probability  $p$ , if there exist two stages  $i$  and  $j$  such that  $i < j$  and  $w_i = w_j$ , then the shortened partial history  $(w_0, \dots, w_{i-1}, w_i, w_{j+1}, \dots, w_t)$  is with probability greater than  $p$ . This follows from the Markov property. In this way, we can reduce the original partial history to a finite sequence  $(w_{t_0}, w_{t_1}, \dots, w_{t_k})$  where  $w_{t_i} \neq w_{t_j}$  for all  $i \neq j$ . Having taken into account the size of the state space  $W$ , we complete the proof.  $\square$

**Theorem 4.4** *Under the best-reply selection mechanism, for a finite generic  $N$ -player game in extensive form with perfect information, it has*

$$\lim_{\mu \rightarrow 0} \pi_{\mu, m}[NE] = 1$$

for all  $m$  in  $\mathbb{N}$ , and hence

$$\lim_{m \rightarrow \infty} \lim_{\mu \rightarrow 0} \pi_{\mu, m}[NE] = 1.$$

**Proof.** We suppose that each player has no more than  $S$  pure strategies. From Lemma 4.2 and Lemma 4.3, it follows that under the best-reply selection mechanism and the Constraint 1 in Section 4.2, a state  $w$  not in  $NE$  can reach a state in  $NE$  in no more than  $(S^m)^N$  stages with positive probability when  $\mu = 0$ . For the general case, given a state  $w$  not in  $NE$ , we show below the existence of a natural number  $t \leq S^{mN}$ , a state  $w'$  in  $NE$  and a positive number  $c(w, w', t)$  with the property that

$$Q_{\mu, m}^t(w'|w) \geq c(w, w', t).$$

(Recall that  $Q$  is a one-step transition probability matrix.)

Without loss of generality, we suppose that there exists a number  $\bar{\sigma}$  with  $0 < \bar{\sigma} < 1$  such that  $\sigma \geq \bar{\sigma}$  in the limiting process of evolutionary dynamics. Recall that Constraint 1 requires that at each stage the chosen individual in each population either moves to pick his best-reply strategy or keep using his best-reply strategy against the current distribution of other populations. So at each stage, the transition probability to the next state where  $l$  populations change due to selection and the other  $N - l$  populations unchanged is no less than

$$m^{-N} \left( \frac{\bar{\sigma}}{S} \right)^l (1 - \mu)^{N-l}$$

for all  $1 \leq l \leq N$ . Therefore, when  $\mu < 1/2$ , we have

$$c(w, w', t) > \left( 2^{1-N} \left( \frac{\bar{\sigma}}{mS} \right)^N \right)^t, \quad (4.1)$$

and

$$\frac{t}{c(w, w', t)} \leq \frac{S^{mN}}{\left(2 \left(\frac{\bar{\sigma}}{2mS}\right)^N\right)^{S^{mN}}} =: f_{\bar{\sigma}}(m). \quad (4.2)$$

Denote the right hand side of the inequality above by  $f_{\bar{\sigma}}(m)$ . Thus  $f_{\bar{\sigma}}(m)$  is a uniform upper bound of the expected number of steps from  $W$  to  $NE$  when no mutation involved, where the uniformity means from any state  $w \in (W \setminus NE)$  to any reachable  $NE$ . From the invariance property, we may infer that

$$\pi_{\mu, m}[w'] \geq \pi_{\mu, m}[w'] Q_{\mu, m}^t(w'|w') + \pi_{\mu, m}[w] Q_{\mu, m}^t(w'|w).$$

Note that  $Q_{\mu, m}^t(w'|w') \geq 1 - Nt\mu$ , as  $w'$  is in  $NE$ . Hence,

$$\pi_{\mu, m}[w] \leq \frac{Nt}{c(w, w', t)} \pi_{\mu, m}[w'] \mu \leq N f_{\bar{\sigma}}(m) \mu.$$

It follows that

$$\sum_{w \in (W \setminus NE)} \pi_{\mu, m}[w] \leq N\mu \sum_{w \in (W \setminus NE)} f_{\bar{\sigma}}(m) < N\mu |W| f_{\bar{\sigma}}(m) = N\mu S^{mN} f_{\bar{\sigma}}(m),$$

when  $\mu < 1/2$ . Therefore,

$$\pi_{\mu, m}[NE] = 1 - \sum_{w \in (W \setminus NE)} \pi_{\mu, m}[w] \geq 1 - N\mu S^{mN} f_{\bar{\sigma}}(m).$$

We then arrive at, for all finite  $m$ ,  $\lim_{\mu \rightarrow 0} \pi_{\mu, m}[NE] = 1$ . It then follows that  $\lim_{m \rightarrow \infty} \lim_{\mu \rightarrow 0} \pi_{\mu, m}[NE] = 1$ .  $\square$

**Comment 1:** The result above also holds for the better-reply selection mechanism described in Section 2. To see this, note that, if player  $I$  is playing some strategy  $c$  and  $a$  is a best reply strategy yielding strictly higher payoff than strategy  $a$ , then  $a$  is also a better reply strategy than  $c$ .

**Comment 2:** The proof presented here is inspired by (3.3) in [8]. The conclusion may also follows from Theorem 3.1 in [15]. However, one needs to find all recurrent classes in the Markov chain with mutation rate being zero.

It is not straightforward to check whether there exists a recurrent class which is a set of states not in  $NE$ . We find the proof above is more informative, and we would like to use the definition of  $f_{\bar{\sigma}}(m)$  introduced there in later proofs.

**Comment 3:** Note that the above theorem holds for both the case  $\sigma = 1 - \mu$  and  $\sigma < 1 - \mu$ .  $\bar{\sigma}$  is introduced in the proof for the latter case.

## 5 From time-average asymptotic behaviour to evolutionary stability

We shall give a proof of Theorem 3.2 by a game shown in Section 6 in the following approach.

Recall that the long-run behaviour of a Markov chain  $(Z_t)$  is well described by its invariant distribution  $\pi$  in the following way. In any long enough period of time, the relative frequency of visits at a subset  $S$  included in  $W$  is approximately  $\pi[S]$ . That is,

$$\forall S \subseteq W, \lim_{T_2 - T_1 \rightarrow \infty} \frac{|\{t : T_1 < t \leq T_2, Z_t \in S\}|}{T_2 - T_1} = \pi[S]. \quad (5.1)$$

This property holds regardless of the initial state.

Our approach to the invariant distribution between  $BC$  and  $NC$  is to analyse the long-term relative proportion of visiting time in  $NC$ .

From Theorem 4.4, we may infer that, under the best-reply selection mechanism without mutation, from any initial state, the expected transition time to  $NE$  is bounded with respect to the population size  $m$ . See Corollary 10.1 for more details. Recall that we consider a model in which the mutation rate decreases to zero. When the mutation rate is much smaller than the inverse of the bound above, the expected duration of two sequential muta-

tions can be enough long such that the probability of moving to a state in  $NE$  before the next mutation is very big. That is, the transition time to  $NE$  triggered by a mutation is in expectation only a fraction of the interval between two mutations.

From another point of view, the probability that a mutation happens at a state not in  $NE$  is very low when mutation rate is small. So, after any mutation happens, we can concentrate on a transition process to  $NE$  in the best-reply dynamic without mutation. If an equilibrium component  $C_1$  is evolutionarily stable, and if a mutation from  $C_1$  can trigger a transition to another component  $C_2$  with positive probability  $p$  in best-reply dynamic without mutation, then  $C_2$  is also evolutionarily stable. To see this, note that in expectation every  $\lceil 1/p \rceil$  times of the mutation above can lead to at least one successful transition to  $C_2$ . Once at  $C_2$ , in expectation it will stay in  $C_2$  at the next  $1/(N\mu)$  stages before the next mutation, where  $N$  is the number of players. Since  $C_1$  is evolutionarily stable, the relative frequency of visits at  $C_1$  is positive. If the process has stayed at  $C_1$  for  $k/(N\mu)$  stages, then in expectation it has visited  $C_2$  for no less than  $kp/(N\mu)$  stages. (We ignore the period when the state is not in  $NE$  here.) From (5.1), we can conclude that  $C_2$  is also evolutionarily stable. Hence, to show  $NC$  being evolutionarily stable, we only need to show a transition triggered by some mutation at  $BC$  can reach  $NC$  with positive probability. The proof of Lemma 5.1 will follow this argument. To apply this lemma in the proof in Section 6, we need  $U(BC, NC)$ , which is the maximum expected transition time from a state in  $BC$  to any state in  $NC$ . The rigorous definition of  $U(BC, NC)$  is introduced in Section 10.1.

For Lemma 5.1, we present some related definitions of an event in the Markov process. Recall that an event is a subset of the sample space  $\Omega = W^{\mathbb{N}}$

which is specified in Section 2. An element  $\omega$  in the sample space  $\Omega$  is an infinite sequence of states, i.e.,  $\omega = (\omega_0, \omega_1, \dots)$  with  $\omega_i$  in  $W$  for all  $i \geq 0$ . Denote the truncated sequence of  $\omega$  with length  $n$  by  $\omega|(n-1) = (\omega_0, \omega_1, \dots, \omega_{n-1})$ . We can view an event as a set of sequences of states. We denote  $\{\omega \in \Omega : \omega \in F\}$  by simply  $F$ . In the proof of Theorem 3.2, we shall define an event  $F$  as an intersection of a finite sequence of events on the dynamic process such that, if  $F$  happens, then  $Z_t$  is in  $NC$  at some stage  $t > 0$ . Note that we can in fact start from any stage  $\bar{t} > 0$ , ignore the partial history  $(Z_0, \dots, Z_{\bar{t}-1})$ , and check whether it will reach  $NC$  after  $\bar{t}$  according to a ‘similar’ transition pattern as in  $F$ . That is, by Markov property, we can shift the original Markov chain  $Z_t$  to  $Z_{t-\bar{t}}$  and then consider the event  $F$ . For instance, when we count from stage  $\bar{t}$ ,  $F$  includes that the first mutant in population  $M(1)$  picks strategy  $a_2$  and the second mutation happens  $2\mu^{-1}$  stages later... Here the exact time index is not important, only transition order and time interval matter.

Note that  $F$  may only be a sufficient condition but not a necessary condition for  $Z_t \in NC$  at some  $t > 0$ . Denote the complement of  $F$  by  $F^c$ . For an initial state  $s$ , we define  $T_s(F)$  (or  $T_s(F^c)$ ) to be the expected time that we know  $F$  is true (or false), respectively. That is,

$$T_s(F) = E[\min\{t : \forall \omega \in \Omega \text{ with } \omega|t = (Z_0 = s, Z_1, \dots, Z_t), \omega \in F\} | F], \quad (5.2)$$

and

$$T_s(F^c) = E[\min\{t : \forall \omega \in \Omega \text{ with } \omega|t = (Z_0 = s, Z_1, \dots, Z_t), \omega \notin F\} | F^c].$$

Note that  $g_{\bar{\sigma}}(m)$  in the lemma below is introduced in Lemma 10.3. To understand and use this lemma, we can view the condition  $\mu < g_{\bar{\sigma}}(m)$  as  $\mu$  being enough small.

**Lemma 5.1** *Without loss of generality, we suppose that there exists a number  $\bar{\sigma}$  with  $0 < \bar{\sigma} < 1$  such that  $\sigma \geq \bar{\sigma}$  in the limiting process of evolutionary dynamics. Given the induced Markov chain with a finite population size  $m$ , suppose that, if  $F$  happens, then  $Z_t$  is in  $NC$  at some stage  $t > 0$ . We further suppose that there exist two positive numbers  $k$  and  $p$  such that given any  $g' \leq g_{\bar{\sigma}}(m)$  and for any mutation rate  $\mu < g'$  it has the property*

$$\min_{w \in BC} P(F|Z_0 = w) \geq p$$

and

$$\max_{w \in BC} T_w(F) + \max_{w \in BC} T_w(F^c) \leq k\mu^{-1}.$$

Then there exists a positive number  $c(p, k)$  such that

$$\forall \mu < g', \pi_{\mu, m}[NC] \geq c(p, k).$$

The proof is in Section 10.2.

## 6 A centipede-trust game

The backward induction equilibrium in the game in Figure 3  $\Gamma_1$  is  $e_b := (s_1, s_2)$ , and the equilibrium component  $BC$  is the set of Nash equilibria with the same result as in  $e_b$ . One alternative pure Nash equilibrium is  $e_n := (high, no)$ , and the corresponding equilibrium component is the set of Nash equilibria with the same result as in  $e_n$ . This is the only Nash equilibrium component different from  $BC$ . So we may also denote it as  $NC$ . For convenience, all backward induction moves are arrowed and the terminating moves in the alternative pure Nash equilibrium are double arrowed in Figure 3.

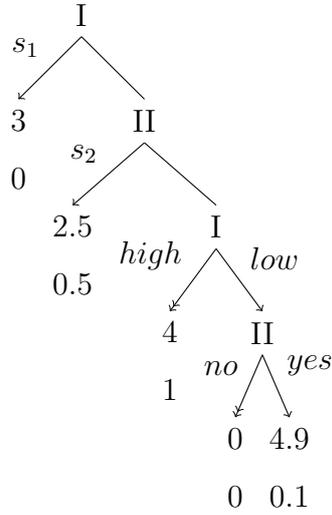


Figure 3: A centipede-trust game  $\Gamma_1$

We say an extensive-form game of perfect information is a *finite stopping game* if each node in it has at most one immediate succeeding node and all his other immediate children are terminal vertices.

We prove Theorem 3.2 by the finite stopping game  $\Gamma_1$  in Figure 3. Note that the paths of both *BC* and *NC* in  $\Gamma_1$  are terminated by player I. (We show the case that *BC* and *NC* are terminated by different players in [13].) We can see that  $\Gamma_1$  is not a generic game. However, we can easily modify the payoff vectors to make it satisfy the generic assumption, e.g., a game  $\Gamma'_1$  is the same as  $\Gamma_1$  except the payoff vector at the top node replaced by  $(2, 0.1)$ . We find the economics intuition is clearer in  $\Gamma_1$  than in  $\Gamma'_1$ . Moreover, the analysis and conclusion on evolutionary dynamics are the same for  $\Gamma_1$  and  $\Gamma'_1$ .

## 6.1 Overview of the game structure and dynamics in

### $\Gamma_1$

In game  $\Gamma_1$ , the subgame consisting of the last two nodes is an ultimatum minigame. (See [5] for experimental results and analysis of replicator-dynamic mechanism on ultimatum games.) In  $\Gamma_1$ , player II has limited bargaining power in the whole game. At stage 1, player I has a pie of size 3. He can either stop the game and eat the whole pie (action  $s_1$ ), or pass the pie to player II. If he chooses to continue the game, player II can then secure  $1/6$  of the pie (payoff 0.5) to himself by stopping the game, or pass the pie back to player I. If no one decides to stop the game in their first moves, then the pie grows to size 5, and player I is going to make a final offer to player II at stage 3. If player I gives a relative generous offer of size 1 to player II, we assume that player II always accepts it. The other option for player I at stage 3 is to demand almost the whole pie to himself. If so, player II can decide whether to accept or reject this tiny 0.1 payoff. If he rejects it, then both players leave with nothing.

If we just consider the subgame of the last two nodes in  $\Gamma_1$ , the result of  $NC$  in the subgame coincides with the result of  $NC$  in  $\Gamma_1$ . The conclusion from Hart and Gorodeisky says that the equilibrium component  $NC$  is not evolutionarily stable for large populations in this subgame. In contrast to that, I shall show that, in the multiple-move game  $\Gamma_1$ ,  $NC$  is evolutionarily stable for any finite populations, and even when the populations increase to infinity.

From another perspective, we may transform  $\Gamma_1$  into a game  $\Gamma_2$  where each player can only move at one node. In  $\Gamma_2$  below, player I and player III has common interest and the identical payoff, and the same case for player II and player IV. The difference of  $\Gamma_2$  from  $\Gamma_1$  is the strategies: in  $\Gamma_2$ ,

even when player I terminates the game by choosing strategy  $s_1$  at the top node, player III has to specify his (reduced) strategy. So there are always  $m$  individuals playing at the third node in the population game of  $\Gamma_2$ , while, in the population game of  $\Gamma_1$ , if all individuals in  $M(1)$  are playing  $s_1$ , then none is playing at the third node.

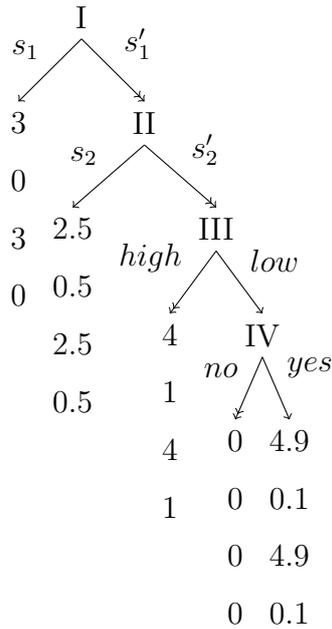


Figure 4:  $\Gamma_2$ : the case that each player plays at only one node

To compare the dynamics in these two population games, we firstly look at  $\Gamma_2$ , which fits Hart's model. Note that in  $\Gamma_2$  only one pure Nash equilibrium  $e_n := (s'_1, s'_2, high, no)$  is included in  $NC$ . One can show that the backward induction equilibrium is the only evolutionarily stable result when populations increase to infinity. To see this claim, from Theorem 4.4, we only need to consider the case that the initial state is a Nash equilibrium. Suppose that the initial state is  $e_b := (s_1, s_2, low, yes)$ , then any mutation in any population will give a worse payoff than the one of his backward induction strategy. So selection will wipe the mutation out, and the backward

induction is stable. The only way to transit from  $BC$  to  $NC$  is by consecutive mutations ( $km$  times for some constant  $k$ ), but that will happen very rarely when populations are large. If the initial state is  $e_n$ , then we consider a mutation in  $M(3)$  to strategy *low*. Once it happens, the bottom node is connected and  $M(4)$  is under the selection pressure to the backward induction strategy *yes*. Note that the probability is positive that the mutant at *low* is not selected in the next  $km$  (e.g.  $k=5$ ) stages and hence it stays there during this period. Thus the probability is also positive that a significant proportion (greater than  $40/49$ ) of  $M(4)$  has moved to *yes* in these  $km$  stages. That will further put  $M(3)$  under the selection pressure to *low* and the impact will finally push  $M(2)$  and  $M(1)$  to their backward induction equilibrium strategy. Therefore, in  $\Gamma_2$ , evolutionary dynamics lead necessarily to  $e_b$ , the backward induction equilibrium.

We now turn to  $\Gamma_1$ , and we present some intuition why it is possible to transit from  $BC$  to  $NC$  via a transition triggered by a single mutation. When a state is in  $BC$ , there is no individual playing at the third node. For a mutation from  $s_1$  to *high*, the mutant will stay at *high* for significantly long time with positive probability, similarly as in the case of  $\Gamma_2$  above. When it is at *high*,  $M(2)$  is under the selection pressure to a strategy *no* or *yes*, which terminates at the bottom node. Each strategy is chosen with equal probability. By the strong law of large numbers, we can show that with positive probability the distribution of individuals at the bottom node can always favours strategy *high* to *low* for  $M(1)$ . With the decrease of the proportion of  $M(2)$  playing at  $s_2$ ,  $M(1)$  will become under selection pressure to leave  $s_1$  at some stage, and that has to be *high* if the condition above is satisfied. Meanwhile,  $M(2)$  keeps moving to *no* and *yes* with equal probability. With positive probability, this transition will reach a state in  $NC$

before the next mutation. From the arguments in Section 5, we know that  $NC$  is evolutionarily stable if the probability is positive that a single mutation can trigger a transition from  $BC$  to  $NC$  under the best-reply dynamic.

## 6.2 Economics motivation

We discuss two points regarding the economics motivation of the game  $\Gamma_1$ . First, this evolutionary approach can support the argument that trust would be relatively easier built up in a game with more steps. (Recall that Hart shows only backward induction equilibria are evolutionarily stable for large populations in a game where each player can only move once.) In this example, if, at stage 1, player I resists the temptation to eat the whole pie immediately and chooses to continue the game with a promise of payoff 1 at stage 3 to player II, then player II would be likely to give a positive response to this decision of player I and continue this game as well. So player I's invitation to cooperation at stage 1 can be viewed as a 'carrot' to attract player II to be more patient. If player II tends to stop the game at stage 2, then player I gradually learns it in the long run and will finally simply stop the game at stage 1, i.e., the result of backward induction equilibrium with 0 payoff to player II. If player II indeed continues the game at stage 2, then the pie grows to size 5. Player I would be happy to reward such reciprocity from player II with  $1/5$  of the current pie and thus realize his promise. If player I is greedy at this stage and try to abuse his first-mover advantage, then it is reasonable to expect that player II might reject the offer to punish player I for cheating. He may simply do so to show his dignity or indifference between nothing and crumbs of the pie. The threat of saying no at the last stage can be viewed as a 'stick'. With both carrot and stick in the game, it is reasonable to predict that alternative Nash equilibrium component might

‘survive’ in the long run.

The trust scheme and the stimulating effects of both carrot and stick can be developed in a multiple-move game with more than 3 nodes. (For games with no more than 3 nodes, one can prove only backward induction equilibrium component is evolutionarily stable for large populations. See Section 9 for more details.)

The second point I would like to mention is that alternative Nash equilibrium component yields a Pareto-efficient equilibrium payoff (3,1). We know that the classical paper [9] of Kandori, Mailath and Rob considers whether Pareto-efficient equilibrium or risk-dominant equilibrium is selected in the long run in a coordination game as the mutation rate goes to 0. By this example, we show that Pareto-efficient equilibrium in an extensive-form game may also be a long-run equilibrium as the mutation rate decreases to 0. Moreover, it is still true even when the population size goes to infinity.

We show a class of such four-node games below in Figure 5 that the alternative evolutionarily stable Nash equilibrium component guarantees each player at least the second best payoff in the whole game, respectively, and it is a Pareto-optimal result.

### 6.3 The proposition and an easy extension

We prove Theorem 3.2 by Proposition 6.1 below, whose proof is included in Section 10.6.

**Proposition 6.1** *Given the game  $\Gamma_1$  in Figure 3, for the best-reply selection or better-reply selection mechanism described in Section 2, there exists a positive number  $c$  such that for all finite  $m$ ,*

$$\lim_{\mu \rightarrow 0} \pi_{\mu,m}[NC] > c.$$

It is not difficult to find a four-node game with an evolutionarily stable equilibrium component in addition to the backward induction equilibrium component. We consider a general four-node game in Figure 5 below with backward induction actions arrowed and the terminating moves in the alternative pure Nash equilibrium double arrowed.

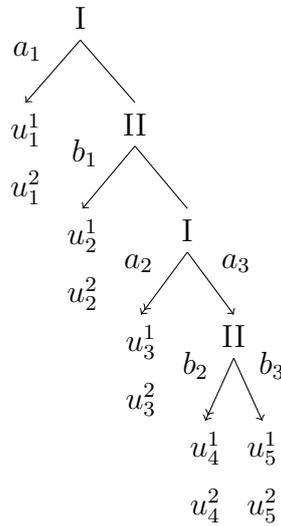


Figure 5: A general four-node game

By Lemma 10.6 and the analogous arguments as in the proof of Proposition 6.1, one can find the alternative Nash equilibrium component with payoff vector  $(u_3^1, u_3^2)$  evolutionarily stable regardless of population size if the following conditions are satisfied:

$$u_2^1 < u_1^1 < u_3^1 < u_5^1, \quad u_3^1 > \frac{u_4^1 + u_5^1}{2},$$

and

$$u_4^2 < u_5^2 < u_2^2 < u_3^2.$$

The transition from  $NC$  to  $BC$  may need two sequential mutations in population  $M(1)$  to strategy  $a_3$  and  $a_2$  in order. The mutant to  $a_3$  may lead to a state in  $BC$  such that the majority of population  $M(2)$  is playing its backward induction strategy  $b_1$ . When population  $M(1)$  is under the selection pressure and moving away from strategy  $a_1$  after the second mutation to strategy  $a_2$ , the relative proportion of  $M(2)$  playing  $b_2$  may be always high enough to make  $a_2$  as the best reply to  $M(1)$ . See the proof of Theorem 7.1 for a similar (but more general) treatment.

It is clear that payoff  $u_3^1$  is the second best payoff to player I, and  $u_3^2$  is at least the second best payoff to player II. We find this alternative Nash equilibrium component is Pareto-efficient. We shall further extend this result in Section 7.

## 7 A sufficient condition for $NC$ being evolutionarily stable

We show a sufficient condition that  $NC$  is evolutionarily stable in  $\Gamma$  for any population size, which is applied in the Theorem 8.2.

We introduce some notations before we present the formal statement. Suppose that a two-player finite extensive-form generic game  $\Gamma$  of perfect information is given. Given a node  $N$  in  $\Gamma$ , in this section and Section 8 we denote the subgame of  $\Gamma$  rooted at  $N$  by  $\Gamma_N$ . Given a Nash equilibrium component  $EC \subseteq NE$ , we denote the path of  $EC$  by  $pa_{EC}$ , which is from the root of  $\Gamma$  to the terminal vertex associated to  $EC$ . Given a node or terminal vertex  $N$  in  $\Gamma$ , we denote the path from the root to  $N$  by  $pa(N)$ . Recall that the *depth* of a node is the length of the path from the root to that node. Given the backward induction equilibrium component  $BC$  and an

alternative Nash equilibrium component  $EC$ , we denote the deepest node on both  $pa_{BC}$  and  $pa_{EC}$  by  $\tilde{N}_{EC}$ . We can decide the identity of player I by the pair  $(BC, EC)$  in the way that we assign the player who moves at node  $\tilde{N}_{EC}$  as player I. (Thus the other player is denoted by player II.) Suppose that there are  $\lambda$  terminal vertices in the generic game  $\Gamma$ . We enumerate the attached values to each player in order, respectively, as  $(u_{(1)}^1, \dots, u_{(\lambda)}^1)$  and  $(u_{(1)}^2, \dots, u_{(\lambda)}^2)$  such that the superscripts correspond to the players and  $u_{(m)}^i > u_{(n)}^i$  for all  $m < n, i \in \{1, 2\}$ . For a Nash equilibrium component  $EC$ , we denote the payoff vector associated to the path  $pa_{EC}$  by  $u_{EC} = (u_{EC}^1, u_{EC}^2)$ . If  $EC$  ends at a terminal vertex  $A$ , we let  $u_A$  be  $u_{EC}$ .

We suppose that player I and player II has  $\varsigma_1$  and  $\varsigma_2$  reduced strategies in  $\Gamma$ , respectively.

**Condition 1:** suppose that a component  $SC$  ends at a terminal vertex denoted as  $V$ . We further suppose  $u_{SC}^2 = u_{(1)}^2$  and

$$u_{SC}^1 = u_{(2)}^1 > \frac{u_{(3)}^1}{\varsigma_2^2} + \frac{(\varsigma_2^2 - 1)u_{(1)}^1}{\varsigma_2^2}.$$

We denote the terminal vertex with the highest payoff  $u_{(1)}^1$  to player I by  $\bar{V}$ , and assume  $V \neq \bar{V}$ . Denote the deepest node on both path  $pa(V)$  and  $pa(\bar{V})$  by  $\bar{N}$ .

**Condition 2:** All nodes in subgame  $\Gamma_{\bar{N}}$  are not on  $pa_{BC}$ .

**Theorem 7.1** *If  $\Gamma$  satisfies both Condition 1 and 2, then there exists a positive number  $c$  such that for all finite  $m > 1$   $NC$  is  $m$ -evolutionarily stable and*

$$\lim_{\mu \rightarrow 0} \pi_{\mu, m}[NC] \geq c.$$

*Thus,  $NC$  is evolutionarily stable for large populations and*

$$\liminf_{m \rightarrow \infty} \lim_{\mu \rightarrow 0} \pi_{\mu, m}[NC] \geq c.$$

This theorem says that, in a two-player finite extensive-form generic game of perfect information, if the payoff vector of one pure Nash equilibrium  $SE$  is enough high and the  $BC$  and  $SE$  path are ‘separated properly’ by the path to  $u_{(1)}^1$ , then  $NC$  is evolutionarily stable regardless of population size.

**Comment:** It is possible to give a more generalised theorem on extensive games of more than two players in the same spirit, however the exact conditions would be more complicated. We can also generalise the condition that each chosen individual select a better (or a best strategy) with equal probability in Section 2. For instance, we can only require that the chosen individual chooses each strategy in  $BT$  with probability at least  $1/(\beta|BT|)$  (or for  $BS$  with probability  $1/(\beta|BS|)$ ) where  $\beta \geq 1$ . In this case, we only need to modify Condition 1 as follows, and Theorem 7.1 still holds.

Suppose the payoff vector of  $V$  is  $(u_{SC}^1, u_{SC}^2)$  with  $u_{SC}^2 = u_{(1)}^2$  and

$$u_{SC}^1 = u_{(2)}^1 > \frac{u_{(3)}^1}{(\beta\varsigma_2)^2} + \frac{((\beta\varsigma_2)^2 - 1) u_{(1)}^1}{(\beta\varsigma_2)^2}.$$

## 8 Extra value to make $NC$ evolutionarily stable

For simplicity, we apply the best-reply selection mechanism in this section. We also apply in the dynamic process the Disconnected-moves-unchanged Constraint, which is introduced in Section 4.2. This constraint seems reasonable from the view of biology.

If without further notice, in this section, we constrain ourselves within the case of two-player finite extensive-form generic games of perfect information with multiple equilibrium components. Given such a game  $\Gamma$ , we suppose that the backward induction equilibrium component is the only evolutionarily

stable result for large populations, i.e.

$$\lim_{m \rightarrow \infty} \lim_{\mu \rightarrow 0} \pi_{\mu, m}[BC] = 1.$$

Note that evolutionary stability can be viewed as a Nash refinement. A natural question is that, if we increase the payoff vector associated to a non-backward-induction equilibrium component, how much extra value do we need to make its equilibrium path an evolutionarily stable result for large populations. One may, of course, lift that payoff vector to be a complete dominant one in the whole game. That is, each payoff in the modified vector is the maximum of all possible payoffs in  $\Gamma$  to the player. That will give a new backward induction equilibrium component. (It is straightforward to see this new backward induction equilibrium component is evolutionarily stable for any population size.) If we have only limited power to raise the payoffs, in particular the case that we can only make the target vector consist of a best and a second best payoff to each player respectively in the modified  $\Gamma$ , is it enough to make its associated path an evolutionarily stable result for large populations? (For simplicity, when a component is evolutionarily stable, we also call its associated path evolutionarily stable.)

**Definition 8.1** *Given a two-player finite extensive-form generic game  $\Gamma$  of perfect information, we suppose that there is another Nash equilibrium component  $SC$  other than the backward induction equilibrium component  $BC$ , and we decide the player  $I$  from  $(BC, SC)$  as in the way above. We say that component  $SC$  is dominated by  $BC$  if and only if  $u_{BC}^1 = u_{(1)}^1$ .*

We first present a negative result by the following game  $\Lambda$  in Figure 6. It shows that, if one non-backward-induction equilibrium component  $SC$  is dominated by  $BC$ , then it is possibly not enough to make it evolutionarily

stable for large populations, by simply raising one payoff in the payoff vector attached to  $SC$  to be the global maximum and the other payoff in it to be the second best.

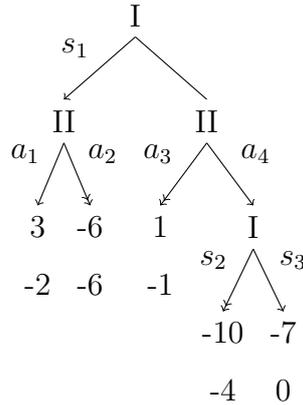


Figure 6: An example

In  $\Lambda$ , the backward induction equilibrium is  $(s_1, (a_1, a_4))$ , and we suppose its path ends at a terminal vertex denoted as  $A$ . (Note that a reduced strategy of player II in  $\Lambda$  needs two actions.) The path of the other equilibrium component, which we denote by  $SC$ , ends at a terminal vertex denoted as  $V$ , and the associated pure Nash equilibrium is  $(s_2, (a_2, a_3))$ . As before, the backward-induction moves are arrowed and the moves of  $SC$  are double arrowed. Note that player I is consistent with the pair  $(BC, SC)$ . We denote  $\bar{u}_{SC}$  to be the increased vector of  $u_{SC}$ . For convenience, we put  $\bar{u} = u$  for all  $u \neq u_{SC}$  and then denote the modified payoff vector sequence by  $(\bar{u})$ . We also denote the such modified game by  $\bar{\Lambda}$ .

One can see that only  $BC$  is evolutionarily stable for large populations in  $\Lambda$  by an adapted proof in [7] or [9]. We can further prove that to make  $SC$  evolutionarily stable with added values on the payoff vector of  $B$ , it needs  $\bar{u}_{SC} = \bar{u}_{(1)}^1$  and  $\bar{u}_{SC}^2 = \bar{u}_{(1)}^2$ .

[ Why is it not enough by putting  $(\bar{u}_{SC} = \bar{u}_{(1)}^1, \bar{u}_{SC}^2 = \bar{u}_{(2)}^2)$  or  $(\bar{u}_{SC} = \bar{u}_{(2)}^1, \bar{u}_{SC}^2 = \bar{u}_{(1)}^2)$ ? One can show two cases above, respectively.

(1)  $\bar{u}_{SC}^1 = \bar{u}_{(1)}^1$  and  $\bar{u}_{SC}^2 = \bar{u}_{(2)}^2$ . The path of  $SC$  in  $\bar{\Lambda}$  is the same as in  $\Lambda$ . This case is similar to the game shown in the Section ‘Further comments’ in [13].

(2)  $\bar{u}_{SC}^1 = \bar{u}_{(2)}^1$  and  $\bar{u}_{SC}^2 = \bar{u}_{(1)}^2$ . The path of  $SC$  in  $\bar{\Lambda}$  is the same as in  $\Lambda$  as well. Once the state is in  $BC$ , player I has no incentive to deviate from its  $pa_{BC}$  in the dynamic process.]

Under the circumstance that, in  $\Gamma$ , one non-backward-induction equilibrium component  $SC$  is not dominated, we have the following theorem.

**Theorem 8.2** *Given a two-player finite extensive-form generic game  $\Gamma$  of perfect information, we suppose that there is another Nash equilibrium component  $SC$  other than the backward induction equilibrium component  $BC$ , and we decide the player I from  $(BC, SC)$ . We further suppose that  $BC$  is the only evolutionarily stable result in  $\Gamma$  for large populations. If  $SC$  is not dominated by  $BC$  in  $\Gamma$ , then there exists a positive  $\epsilon(\Gamma)$  with the following proposition. If we raise  $u_{SC}$  to  $\bar{u}_{SC}$  such that  $\bar{u}_{SC}^2 = \bar{u}_{(1)}^2$  and  $\bar{u}_A^1 > \bar{u}_{(1)}^1 - \epsilon(\Gamma)$ , then the equilibrium component in  $\bar{\Gamma}$  consistent with  $pa(SC)$  becomes the evolutionarily stable result for any population size.*

**Comment:** One can also prove that, in a game  $\Gamma$  where every component  $SC \subset NC$  is dominated by  $BC$ , then only  $BC$  is evolutionarily stable for large populations. We omit the proof here.

**Proof.** Denote the terminal vertex of path  $SC$  to be  $A$ . Denote by  $\bar{V}$  the terminal vertex associated to the payoff vector  $u$  which contains  $u_{(1)}^1$  in  $\Gamma$ . Thus the path from the root to  $\bar{V}$  is denoted as  $pa(\bar{V})$ . We still denote the deepest node on both  $pa_{SC}$  and  $pa(\bar{V})$  by  $\bar{N}$ . Recall that the deepest

node on both  $pa_{BC}$  and  $pa_{SC}$  is  $\tilde{N}_{SC}$ . For simplicity, we re-denote it as node  $\tilde{N}$ . We prove the theorem in the following two cases.

1. **Case I:**  $\tilde{N} \in \Gamma_{\tilde{N}}$ . We show that, when  $\bar{u}_A^2 = \bar{u}_{(1)}^2$  and  $\bar{u}_A^1 = \bar{u}_{(2)}^1$ ,  $pa(A)$  becomes the path of the new backward induction equilibrium in  $\bar{\Gamma}$ . Firstly note that  $\bar{u}_A^1 > \bar{u}_p^1$  and  $\bar{u}_A^2 > \bar{u}_p^2$  for any path  $p$  which ends in a terminal vertex in subgame  $\bar{\Gamma}_{\tilde{N}}$ . So, the path of the backward induction equilibrium in  $\bar{\Gamma}_{\tilde{N}}$  leads to the payoff vector  $\bar{u}_A$ . If we turned the node  $\tilde{N}$  to a terminal vertex with payoff vector  $u_{BC}$  (the one of the original backward induction equilibrium in  $\Gamma$ ), then the backward induction path would proceed from the root to this new node. Now  $\bar{u}_A$  dominates  $u_{BC}$ , and thus the new backward induction equilibrium path proceeds towards  $A$ . It is straightforward to see this is an evolutionarily stable result for any population size.

2. **Case II:**  $\tilde{N} \notin \Gamma_{\tilde{N}}$ . We divide it into two sub-cases.

(a) **Sub-case I:** the backward induction equilibrium in subgame  $\Gamma_{\tilde{N}}$  terminates at  $\bar{V}$ . It follows that the path of the backward induction equilibrium remains the same after raising  $u_A$ .

As before, we suppose player I and player II has  $\varsigma_1$  and  $\varsigma_2$  reduced strategies in  $\Gamma$ , respectively. Define

$$\epsilon(\Gamma) := u_{(1)}^1 - \left( \frac{u_{(2)}^1}{\varsigma_2^2} + \frac{(\varsigma_2^2 - 1)u_{(1)}^1}{\varsigma_2^2} \right).$$

We apply Theorem 7.1, and the desired conclusion follows.

(b) **Sub-case II:** the backward induction equilibrium in subgame  $\Gamma_{\tilde{N}}$  does not terminate at  $\bar{V}$ . By the similar arguments as in Case I, we find that  $pa(A)$  becomes the path of the new backward induction equilibrium in  $\bar{\Gamma}$ , and the corresponding equilibrium component

is an evolutionarily stable result for any population size. (One mutation is enough to trigger a transition from any other equilibrium component to this new backward induction equilibrium component.)

□

## 9 Further research directions

There are many directions for further research. The basic model in Section 2 is a reasonable one in selection-mutation process when  $m$  is not too large. If  $m$  goes to infinity, one may find the condition that even very small difference of payoff, e.g.  $1/m$ , caused by one mutant will change the selection power completely. In Example 1, during the period of only one mutant in population  $M(1)$  playing strategy  $a_2$ , on average more than  $1/4$  of  $M(2)$  would be attracted to strategy  $b_1$ , away from its backward induction strategy  $b_3$ . That looks too strong for some real situations. (However, for some situations in epidemiology and epizology, one ‘bad’ mutation can indeed influence all related populations.) So one might try to smooth the selection process and make it continuous. For instance, Hart conjectures to use the payoff difference to a positive power less than 1. (See (5.1) in [8].)

We only consider the case that  $\mu$  is extremely small with respect to  $m$  in the limiting process. (Recall Lemma 10.3.) One can also study other conditions. For instance, the expected number of mutation per generation does not go to zero. That is,  $m\mu > c$  for some  $c > 0$ . The more specific case is  $\liminf_{\mu \rightarrow 0} \liminf_{m \rightarrow \infty} \pi_{\mu, m}[BC_\epsilon]$ .

One may also study under what tree structure in our models the  $BC$  is the unique evolutionarily stable outcome. An interesting result is that the

evolutionary stability is affected by background parameters under some tree structures. For instance, in the game  $\Gamma_1$  of Example 1, we can prove that, given any pair of  $\delta$  and  $\eta$  with  $0 < \delta < \eta$ ,

$$\lim_{\substack{\mu \rightarrow 0, m \rightarrow \infty \\ \delta \leq m\mu \leq \eta}} \pi_{\mu, m}[BC_\epsilon] = 1,$$

for all  $\epsilon > 0$ . Thus the Pareto-efficient equilibrium component with payoff vector (3,1) is not evolutionarily stable in this limiting process.

## 10 Appendix

### 10.1 Results on transition time in Markov chain

The following corollary to Theorem 4.4 estimates the expected transition time when no mutation involved. Given a Markov chain  $(Z_t)_{t \geq 0}$  generated in the selection-mutation process with population size  $m$  and mutation rate  $\mu$ , for a state  $w \in W$  and a subset  $S \subseteq W$ , we define the random variable  $T_{\mu, m}(w, S) := \min\{t \geq 0 : Z_t \in S, Z_0 = w\}$  to be the transition time from  $w$  to  $S$ , and define the expected transition time from  $w$  to  $S$  to be

$$U_{\mu, m}(w, S) := E_{\mu, m}[T_{\mu, m}(w, S)].$$

For every two subsets  $S_1, S_2 \subseteq W$ , we put

$$U_{\mu, m}(S_1, S_2) := \max_{w \in S_1} U_{\mu, m}(w, S_2).$$

**Corollary 10.1** *Given a Markov chain  $(Z_t)_{t \geq 0}$  with  $\mu = 0$  and the associated  $f_\sigma(m)$  defined in (4.2) for a  $\sigma > 0$ , we have  $U_{0, m}(W, NE) < f_\sigma(m)$ .*

**Proof.** The result follows from the definition of  $f_\sigma(m)$  in (4.2).  $\square$

We need the following lemma to show the expected transition time to Nash equilibria when  $\mu$  is small.

**Lemma 10.2** *Given a number  $p$  in  $(0, 1/2)$ , we consider a random variable  $X_p : \Omega \rightarrow \mathbb{N}$  in geometric distribution with*

$$P(X_p = k) = (1 - p)^{k-1}p \quad \forall k \in \mathbb{N}.$$

*Then, we have  $P(X_p > E[X_p]) > 1/8$ .*

**Proof.** It is straightforward to see that  $E[X_p] = 1/p$ . Then,

$$\begin{aligned} P(X_p > E[X_p]) &\geq \sum_{i > 1/p} P(X_p = i) \\ &\geq \sum_{i > \lceil 1/p \rceil} (1 - p)^{i-1}p \\ &= (1 - p)^{\lceil 1/p \rceil} \\ &> (1/2)(1 - p)^{1/p}, \end{aligned}$$

as  $p < 1/2$ . Note that the function  $y(n) = (1 - 1/n)^n$  is a strictly monotone increasing function when  $n > 1$ . We may then infer that

$$P(X_p > E[X_p]) > P(X_{1/2} \geq E[X_{1/2}]) > 1/8.$$

□

The next lemma shows that the expected transition time to Nash equilibria is bounded when mutation rate  $\mu$  is small.

**Lemma 10.3** *Consider a two-player extensive-form game of perfect information with the induced Markov chain. Without loss of generality, we suppose that there exists a number  $\bar{\sigma}$  with  $0 < \bar{\sigma} < 1$  such that  $\sigma \geq \bar{\sigma}$  in the limiting process of evolutionary dynamics. Denote  $c_{\bar{\sigma}}$  to be  $e^{-200/\bar{\sigma}}$ . Then, given  $f_{\bar{\sigma}}(m)$  defined in (4.2), for every  $\mu < c_{\bar{\sigma}}/(32f_{\bar{\sigma}}(m))$ , it satisfies  $U_{\mu,m}(W, NE) < c_{\bar{\sigma}}\mu^{-1}$ . Define  $g_{\bar{\sigma}}(m) := c_{\bar{\sigma}}/(32f_{\bar{\sigma}}(m))$ .*

**Proof.** From Corollary 10.1,  $U_{0,m}(W, NE) < f_{\bar{\sigma}}(m)$ , where  $f_{\bar{\sigma}}(m)$  is defined in (4.2). If no mutation is involved in the dynamic process, then any process starting from a state  $w$  not in  $NE$  would reach  $NE$  almost surely.

Given  $\bar{\sigma} > 0$ , note that  $1/(2\mu) > 16f_{\bar{\sigma}}(m)/c_{\bar{\sigma}} > 2f_{\bar{\sigma}}(m)$ . Denote the first stage that a mutation happens by  $t_{\mu} > 0$ . As it is for a two-player game, we apply  $p = 2\mu - \mu^2$  in Lemma 10.2, and obtain  $E[t_{\mu}] > 1/(2\mu)$ . It follows from Lemma 10.2 that

$$P(t_{\mu} > 1/(2\mu) > 2f_{\bar{\sigma}}(m)) > 1/8. \quad (10.1)$$

From Corollary 10.1 and Markov inequality, it follows that, for any initial state  $w \in W$ ,

$$P(T_{\mu,m}(w, NE) < 2f_{\bar{\sigma}}(m) | t_{\mu} > 2f_{\bar{\sigma}}(m)) > 1/2.$$

Therefore, by (10.1), the probability that it enters  $NE$  by stage  $2f_{\bar{\sigma}}(m)$  is bounded as follows.

$$\begin{aligned} & P(T_{\mu,m}(w, NE) < 2f_{\bar{\sigma}}(m)) \\ & \geq P(t_{\mu} > 2f_{\bar{\sigma}}(m)) P(T_{\mu,m}(w, NE) < 2f_{\bar{\sigma}}(m) | t_{\mu} > 2f_{\bar{\sigma}}(m)) \\ & > 1/16 \quad \forall w \in W. \end{aligned} \quad (10.2)$$

Denote  $\min_{w \in W} P(T_{\mu,m}(w, NE) < 2f_{\bar{\sigma}}(m))$  by  $\bar{P}$ . Then  $\bar{P} > 1/16$ . By Markov property, we may then infer that

$$U_{\mu,m}(W, NE) \leq \sum_{i \geq 1} (\bar{P}(1 - \bar{P})^{i-1} (2i)f_{\bar{\sigma}}(m)) = 2f_{\bar{\sigma}}(m)/\bar{P} < 32f_{\bar{\sigma}}(m).$$

As  $1/\mu > 32f_{\bar{\sigma}}(m)/c_{\bar{\sigma}}$ , the proof is completed.  $\square$

The following lemma shows in expectation the transition conditional on passing some subset takes longer than the one without such condition.

**Lemma 10.4**

$$\forall W_3 \subseteq W, U(W_3, NC) \leq U(W_3, BC) + U(BC, NC).$$

**Proof.** By Lemma A.1 in [7], for any  $w$  in  $W_3$ ,

$$U(w, NC) \leq U(w, BC) + U(BC, NC) \leq U(W_3, BC) + U(BC, NC),$$

which completes the proof.  $\square$

It is also important to know in expectation how long the Markov chain stays at the same state each time. The next lemma answers this question for an arbitrary state in  $NE$ . To present it formally, we first introduce some terminology. Given the Markov chain  $(Z_t)_{t \geq 0}$ , for a state  $w$  in  $W$ , suppose that  $Z_0 \neq w$  and mark stage 0 as  $t_0^e(w)$ . For each  $i \geq 0$ , we mark the stage  $\min\{t > t_i^e(w) : Z_t = w\}$  as  $t_{i+1}^s(w)$  and then the stage  $\min\{t > t_{i+1}^s(w) : Z_t \neq w\}$  as  $t_{i+1}^e(w)$  in turn. We define  $\bar{U}(w)$  to be the expected length of the period during which the dynamic process stays at  $w$ , i.e.,

$$\bar{U}(w) := E \left[ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{1 \leq i \leq N} (t_i^e(w) - t_i^s(w)) \right].$$

Note that the expected length of the period at  $w$  does not depend on the state at initial stage, so we can assume  $Z_0 \neq w$  in the condition above.

**Lemma 10.5** *Given a two-player extensive-form game of perfect information, for the induced Markov chain with the mutation rate  $\mu < g_a(m)$  where  $0 < a \leq 1$ , we have*

$$\bar{U}(w) \geq 1/(2\mu)$$

*for all  $w$  in  $NE$ .*

**Proof.** From stage 0, we enumerate the stages that at least one mutation happens as  $u_1, u_2, \dots$ . At every stage, the probability of a mutation is no more than  $2\mu$  from both populations. It follows that

$$E[u_{k+1} - u_k] \geq 1/(2\mu) \quad (10.3)$$

for all  $k$ . Note that once  $w$  is in  $NE$ , the state will keep unchanged at least until the next mutation. The result now follows.  $\square$

## 10.2 Proof of Lemma 5.1

From the definition of  $U(BC, NC)$  and Markov property, it follows that

$$\begin{aligned} & U(BC, NC) \\ &= \max_{w \in BC} U(w, NC) \\ &\leq \max_{w \in BC} (P(F|Z_0 = w) T_w(F) + P(F^c|Z_0 = w) (T_w(F^c) + U(W, NC))). \end{aligned}$$

By Lemma 10.4, we find  $U(W, NC) \leq U(W, NE) + U(BC, NC)$ . With the observation of

$$P(F^c|Z_0 = w) \leq 1 - p \quad \forall w \in BC,$$

it follows that

$$pU(BC, NC) \leq \max_{w \in BC} T_w(F) + \max_{w \in BC} T_w(F^c) + U(W, NE).$$

By Lemma 10.3 and initial assumptions, we find  $U(BC, NC) \leq p^{-1}(k + c_{\bar{\sigma}})\mu^{-1}$ . By Lemma 10.4 again, we arrive at

$$U(W, NC) \leq c_{\bar{\sigma}}\mu^{-1} + p^{-1}(k + c_{\bar{\sigma}})\mu^{-1}.$$

From (5.1), we observe that

$$\pi[NC] \geq \frac{\min_{w \in NC} \bar{U}(w)}{U(W, NC) + \min_{w \in NC} \bar{U}(w)}.$$

Lemma 10.5 then completes the proof.

### 10.3 A Combinatorial Lemma

Consider a finite sequence of i.i.d. random variables  $(Y_i)_{0 < i \leq n}$  with  $P(Y_i = 1) = P(Y_i = -1) = 1/2$  for all  $0 < i \leq n$ . For any finite number  $k$  with  $k > 1$ , the next lemma gives a lower bound of the probability that the number of  $Y_i = 1$  is always no less than the number of  $Y_i = -1$  divided by  $k$  during a sampling of this sequence  $(Y_i)_{0 < i \leq n}$ .

**Lemma 10.6** *Let  $(Y_i)_{i \in \mathbb{N}}$  be an infinite sequence of i.i.d. random variables with  $P(Y_i = 1) = P(Y_i = -1) = 1/2$  for all  $i \in \mathbb{N}$ . For each positive integer  $t$ , we let  $S_t$  denote the  $\sum_{i=1}^t Y_i$ . The sequence  $(S_t)_{t \in \mathbb{N}}$  is a standard random walk. Define  $B_t = \sum_{i=1}^t \mathbb{1}_{Y_i=1}$  and  $C_t = \sum_{i=1}^t \mathbb{1}_{Y_i=-1}$ . Given any  $k > 1$ , there exists a natural number  $s_k$  such that, for all  $n > 0$ , we have*

$$P(S_t/t \geq (1-k)/(k+1) \forall 0 < t \leq n) \geq 2^{-s_k},$$

and thus,

$$P(kB_t \geq C_t \forall 0 < t \leq n) \geq 2^{-s_k}.$$

**Proof.** Given  $k$  with  $k > 1$  and a natural number  $n$ , we firstly observe

$$P(kB_t \geq C_t \forall 0 < t \leq n) = P(S_t/t \geq (1-k)/(k+1) \forall 0 < t \leq n). \quad (10.4)$$

For each natural number  $l$ , we define a set

$$A_l := \bigcap_{t \geq l} \left\{ \frac{S_t}{t} > \frac{1-k}{2(k+1)} \right\}.$$

From Strong Law of Large Numbers, we know  $P(\lim_{t \rightarrow \infty} (S_t/t) = 0) = 1$ . It follows that  $P(\bigcup_{l \in \mathbb{N}} A_l) = 1$ . By monotone-convergence properties of probability measures, we arrive at  $\lim_{l \rightarrow \infty} P(A_l) = 1$ . Hence, there exists a

natural number  $t_k$  such that  $P(A_{t_k}) > 1/2$ . We also know that

$$P(A_{t_k}) = \sum_{l=-t_k}^{t_k} \left( P(S_{t_k} = l) P \left( \bigcap_{t \geq t_k} \left\{ \frac{S_t}{t} > \frac{1-k}{2(k+1)} \right\} \mid S_{t_k} = l \right) \right)$$

It then follows that

$$P \left( \bigcap_{t \geq t_k} \left\{ \frac{S_t}{t} > \frac{1-k}{2(k+1)} \right\} \mid S_{t_k} = t_k \right) > 1/2.$$

Therefore, for any  $n > t_k$ ,

$$\begin{aligned} & P \left( \frac{S_t}{t} \geq \frac{1-k}{k+1} \forall 0 < t \leq n \right) \\ & > P(S_{t_k} = t_k) P \left( \bigcap_{t \geq t_k} \left\{ \frac{S_t}{t} > \frac{1-k}{2(k+1)} \right\} \mid S_{t_k} = t_k \right) \\ & > 2^{-t_k-1}. \end{aligned}$$

For any  $n \leq t_k$ , it is straightforward to see

$$P \left( \frac{S_t}{t} \geq \frac{1-k}{k+1} \forall 0 < t \leq n \right) > 2^{-k}.$$

By (10.4), we complete the proof.  $\square$

## 10.4 Preliminary lemmas for the dynamics

**Lemma 10.7** *Given the induced Markov chain  $(Z_t)$  of a finite extensive-form game of perfect information with population size  $|m| > 1$ , suppose an individual  $q$  is chosen in population  $M(i) \ni q$  at some stage  $t$ , then for any positive number  $k$ , the probability that  $q$  is not chosen in  $M(i)$  for all stages between  $t$  and  $t + \lceil km \rceil$  is greater than  $2^{-2k-1}$ .*

**Proof.** Note that at every stage the probability that the individual  $q$  is chosen is  $1/m$ . We may then infer that the probability of  $q$  not chosen for all stages between stage  $t$  and  $t + km$  is at least  $(1 - 1/m)^{km} > 2^{-2k}$ .  $\square$

**Lemma 10.8** *Assume that at each stage the probability that selection takes effect at a chosen individual is no less than  $\bar{\sigma}$ . Given the induced Markov chain  $(Z_t)$  of a finite extensive-form game of perfect information, suppose that population  $M(i)$  is under selection pressure moving away strategy  $s_i$  for all stages  $t > \bar{t}$ . Then, for any pair of positive numbers  $k$  and  $\lambda$ , it follows that*

$$P(m_{s_i}(\bar{t} + \lceil km \rceil) > \lambda m | G_{\bar{t}, \bar{t} + \lceil km \rceil}) < \frac{1 - \lambda}{k\lambda\bar{\sigma}},$$

where  $G_{t_1, t_2}$  requires no mutation in all populations between stage  $t_1$  and  $t_2$ .

**Proof.** Without loss of generality, we assume  $m_{s_i}(t) = m$ . Denote by  $\bar{T}_\lambda$  the first time after  $t$  that the number of individuals in population  $M(i)$  which play strategy  $s_i$  is no more than  $\lambda m$ , i.e.,

$$\bar{T}_\lambda := \min\{t > \bar{t} : m_{s_i}(t) \leq \lambda m\}.$$

In evolutionary dynamics, when  $m_{s_i} > \lambda m$ , at any stage, the probability that an individual playing strategy  $s_i$  is chosen in population  $M(i)$  is greater than  $\lambda$ ; conditional on such an individual is chosen, the probability that selection takes effect at that stage is no less than  $\bar{\sigma}$ . Hence, for any stage  $t \geq \bar{t}$ ,

$$P(m_{s_i}(t+1) < m_{s_i}(t) > \lambda\bar{\sigma}).$$

It follows that

$$E[\bar{T}_\lambda - \bar{t}] < (1 - \lambda)m/(\lambda\bar{\sigma}).$$

By Markov inequality, we find

$$P(\bar{T}_\lambda - \bar{t} > km) < \frac{1 - \lambda}{k\lambda\bar{\sigma}},$$

which completes the proof.  $\square$

## 10.5 A sufficient condition that strategy *low* is not a best reply in $\Gamma_1$ in Figure 3

For the population game associated with  $\Gamma_1$ , for any two stages  $s_1$  and  $s_2$  with  $s_1 < s_2$ , denote the difference of number of individuals playing strategy *no* and *yes* by  $\Delta_{no}(s_1, s_2) := m_{no}(s_2) - m_{no}(s_1)$  and  $\Delta_{yes}(s_1, s_2) := m_{yes}(s_2) - m_{yes}(s_1)$ , respectively.

**Lemma 10.9** *Consider the Markov chain associated to the game  $\Gamma_1$  with the best-reply. Given two stages  $t_1$  and  $t_2$  with  $t_1 < t_2$ , we assume that no mutation happens between  $t_1$  and  $t_2$ , and that  $m_{high}(t_1) = 1$ ,  $m_{low}(t_1) = 0$ , and that population  $M(2)$  satisfies the following (10.5) at stage  $t_1$ .*

$$m_{s_2}(t_1) \geq 2m/3 \text{ and } 2.5m_{s_2}(t_1) + 4.9m_{yes}(t_1) \leq 3m. \quad (10.5)$$

If

$$1.01\Delta_{no}(t_1, t) \geq \Delta_{yes}(t_1, t) \geq 0 \quad \forall t \text{ with } t_1 < t \leq t_2, \quad (10.6)$$

then *low* is not the strict best reply of population  $M(1)$  at  $t_2$ , and hence no individual in population  $M(1)$  is playing *low* at stage  $t_2$ .

**Proof.** We denote  $u_{low}(t)$  as the payoff to an individual in population  $M(1)$  who plays *low* at stage  $t$ . By (10.5), which is the constraint of population  $M(2)$  for *BC*, we observe  $u_{low}(t_1) \leq 3$ . Suppose the proportions of population  $M(2)$  playing the strategies  $s_2$ , *no* and *yes* at stage  $t_1$  are  $p_1 = x_{s_2}^2(Z_{t_1})$ ,  $p_2 = x_{no}^2(Z_{t_1})$  and  $p_3 = x_{yes}^2(Z_{t_1})$ , respectively. Then, the payoff of an individual in population  $M(1)$  playing strategy *low* at stage  $t_1$  is

$$u_{low}(t_1) = 2.5p_1 + 0p_2 + 4.9p_3.$$

Then, given a stage  $t$  with  $t_1 < t \leq t_2$ , with the possible move of population  $M(2)$  under the constraint (10.6), the payoff of an individual in popu-

lation  $M(1)$  playing strategy *low* at stage  $t$  is

$$\begin{aligned} u_{low}(t) &\leq (p_1 - x - x/1.01)2.5 + (p_2 + x/1.01)0 + (p_3 + x)4.9 \\ &\leq 2.5p_1 + 4.9p_3 - 0.075x \end{aligned}$$

where  $x \geq 0$  is the proportion difference of population  $M(2)$  playing *yes* between stage  $t_1$  and stage  $t$ . Thus  $u_{low}(t) \leq u_{low}(t_1) \leq 3$ , and hence strategy *low* is weakly dominated by strategy  $s_1$  at stage  $t$ . Recall  $m_{low}(t_1) = 0$  and the absence of mutation between stage  $t_1$  and  $t$ . We may then infer that there is still be no individual playing at strategy *low* at stage  $t$ .  $\square$

## 10.6 Proof of Proposition 6.1

As mentioned in Section 5, we are going to define  $F$  as an intersection of a finite sequence of events on the dynamic process such that, if  $F$  happens, then  $Z_t$  is in  $NC$  at some stage  $t > 0$ . In the proof below, event  $F_4$  itself leads to the entry into  $NC$ , however it is concerned with the states visited between the first and second mutation. We need events  $F_1$ ,  $F_2$  and  $F_3$  to deduce the probability of  $F = \bigcap_{i=1}^4 F_i$  in a dynamic process starting from stage 0.

**Proof of Proposition 6.1:** We consider the case of the best-reply selection mechanism, and the proof is essentially the same for the better-reply selection mechanism. Without loss of generality, we suppose that there exists a number  $\bar{\sigma}$  with  $0 < \bar{\sigma} < 1$  such that  $\sigma \geq \bar{\sigma}$  in the limiting process of evolutionary dynamics.

We take  $k = 1.01$  in Lemma 10.6, and in this proof denote  $s_{1.01}$  by  $\bar{s}$ . We further denote  $2^{-\bar{s}}g_{\bar{\sigma}}(m)$  by  $g'$ , and assume  $\mu < g'$ . (See Lemma 10.3 for the definition of  $g_{\bar{\sigma}}(m)$ .) As specified in the proof of Lemma 5.1, our main goal is to calculate  $U(BC, NC)$ . To this end, we are going to define a finite

sequence of events  $(F_1, F_2, F_3, F_4)$  with respect to the Markov chain  $(Z_t)_{t \geq 0}$ , and put event  $F$  to be  $\bigcap_{i=1}^4 F_i$ . For simplicity, we denote  $\bigcap_{i=1}^l F_i$  by  $D_l$  for all  $1 \leq l \leq 4$ . Thus  $D_l = D_{l-1} \cap F_l$ . To deduce  $U(BC, NC)$ , we assume  $Z_0 \in BC$  at stage 0. We shall see the argument below is independent to the exact initial state in  $BC$ . We enumerate the stages that at least one mutation happens as  $u_1, u_2, \dots$

**Event**  $F_1 := \{Z_{u_1}^1(q^1) = \text{high}, B_i(u_1) = B_i(0) \forall i = 1, 2, 3\}$ . [The first mutation is in population  $M(1)$  only and that yields a *high* strategy.]

**Claim:** When  $\mu < 1/7$ , for all  $w$  in  $BC$ ,

$$P(F_1|Z_0 = w) = (1 - \mu)/6 > 1/7. \quad (10.7)$$

Note that conditional on  $u_1$  is finite, the probability that the first mutation happens at population  $M(1)$  is  $1/2$ ; conditional on the first mutation affects population  $M(1)$ , the probability that the mutant's new strategy is *high* and no mutation happens in population  $M(2)$  is  $(1 - \mu)/3$ .

**Event**  $F_2 := \{\mu^{-1} < u_2 - u_1 < 100/\mu\}$ . [The number of stages between the first and the second mutation is between  $\mu^{-1}$  and  $100\mu^{-1}$ .]

**Claim:** When  $\mu < 1/4$ , for all  $w$  in  $BC$ ,

$$P(F_2|Z_0 = w) > 1/16 - 1/200. \quad (10.8)$$

It is straightforward to see

$$P(u_2 - u_1 > \mu^{-1}) \geq (1 - 2\mu)^{\mu^{-1}} = \left( (1 - 2\mu)^{1/2\mu} \right)^2 > 4^{-2} = 1/16, \quad (10.9)$$

when  $\mu < 1/4$ . (Recall that  $y(n) = (1 - 1/n)^n$ ,  $n \geq 2$ , is a strictly monotone increasing function with limit  $e^{-1}$  as  $n \rightarrow \infty$ .) From  $E[u_2 - u_1] = 1/(2\mu)$ , it follows that  $P(u_2 - u_1 > 100/\mu) < 1/200$  by Markov inequality.

We denote  $t_1$  to be the first stage after  $u_1$  when the state is in  $NE$ , i.e.,  $t_1 = \min\{t > u_1 : Z_t \in NE\}$ .

**Event**  $F_3 := \{t_1 - u_1 < 1/(2\mu)\}$ . [From stage  $u_1$ , it takes less than  $1/(2\mu)$  stages back into NE.]

**Claim:** Denote  $2 + (8/\bar{\sigma})$  by  $\bar{c}_{\bar{\sigma}}$ . For all  $w$  in  $BC$ ,

$$P(F_3|D_2, Z_0 = w) \geq 1 - e^{-\bar{c}_{\bar{\sigma}}2^{-\bar{s}}}. \quad (10.10)$$

If  $t_1 < u_2$ , then throughout the period between  $t_1$  and  $u_2$ , no change happens on either  $X^1$  or  $X^2$ , and it is at a state in  $NE$ . Recall the definition of  $f_{\bar{\sigma}}(m)$  in (4.2). From Corollary 10.1 and Markov inequality,

$$P(U_{0,m}(Z_{u_1}, NE) > e^{\bar{c}_{\bar{\sigma}}2^{\bar{s}}}f_{\bar{\sigma}}(m)) < e^{-\bar{c}_{\bar{\sigma}}2^{-\bar{s}}}. \quad (10.11)$$

We may then infer

$$P(t_1 - u_1 > e^{\bar{c}_{\bar{\sigma}}2^{\bar{s}}}f_{\bar{\sigma}}(m)|u_2 - u_1 > 1/\mu) < e^{-\bar{c}_{\bar{\sigma}}2^{-\bar{s}}},$$

from the fact of no mutation between stage  $u_1$  and  $u_2$  and the assumption  $\mu < g' \leq e^{-\bar{c}_{\bar{\sigma}}}/(2^{\bar{s}}32f_{\bar{\sigma}}(m))$ . Therefore, for all  $w$  in  $BC$ ,

$$P(t_1 - u_1 > 1/(2\mu)|D_2, Z_0 = w) < e^{-\bar{c}_{\bar{\sigma}}2^{-\bar{s}}}.$$

**Event**  $F_4 := \{\exists t : u_1 < t < u_2, Z_t \in NC\}$ . [The process enters into  $NC$  at some stage between the first and the second mutation.]

**Claim:**

$$P(F_3 \cap F_4|D_2, Z_0 = w) \geq 2^{-\bar{s}}(2^{-\bar{c}_{\bar{\sigma}}} - e^{-\bar{c}_{\bar{\sigma}}}) \forall w \in BC. \quad (10.12)$$

Note that, at any stage  $t > u_1$ , if  $m_{high}(t) > m_{low}(t) = 0$ , then population  $M(2)$  is under the selection pressure to strategies *no* and *yes*. For any two stages  $\tau_1$  and  $\tau_2$  with  $\tau_1 < \tau_2$ , denote the difference of number of individuals playing strategy *no* and *yes* by  $\Delta_{no}(\tau_1, \tau_2) := m_{no}(\tau_2) - m_{no}(\tau_1)$  and  $\Delta_{yes}(\tau_1, \tau_2) := m_{yes}(\tau_2) - m_{yes}(\tau_1)$ , respectively. Given two stages  $\tau_1 < \tau_2$ ,

denote the event of no mutation in both populations between stage  $\tau_1$  and  $\tau_2$  by  $G_{\tau_1, \tau_2}$ , and the event of  $1.01\Delta_{no}(\tau_1, t) \geq \Delta_{yes}(\tau_1, t)$  for all stages  $t$  with  $\tau_1 < t < \tau_2$  by  $H_{\tau_1, \tau_2}$ . We take  $m_{s_2}(u_1)$  as number  $n$  in Lemma 10.6, and find that, given any  $t > u_1$ , (for example, take  $t$  as  $\min\{t' > u_1 : m_{s_2}(t') = 0\}$ )

$$P(H_{u_1, t} | G_{u_1, t}, m_{high}(t') \geq m_{low}(t') = 0 \forall t' \text{ with } u_1 < t' < t) \geq 2^{-\bar{s}}. \quad (10.13)$$

Conditional on that *low* is never the best reply of population  $M(1)$ , with the decrease of  $x_{s_2}^2$ , *high* will become the best reply of population  $M(1)$  when  $x_{s_2}^2 < 1/3$ . Supported by Lemma 10.9, we show below that the following event happens with positive probability conditional on  $1.01\Delta_2(t) \geq \Delta_3(t)$  and no mutation involved for all stages  $t > u_1$ : the whole population  $M(1)$  moves to *high* and the whole population  $M(2)$  moves away from  $s_2$ , i.e., it enters a state in NC.

For the original Markov chain  $(Z_t)_t$ , we take  $k = 4/\bar{\sigma}$  in Lemma 10.7 and conclude that the probability is greater than  $2^{-1-2k}$  that the mutant will keep playing strategy *high* between stage  $u_1$  and  $u_1 + \lceil km \rceil$ . (\*) Taking the same  $k$  and  $\lambda = 1/3$  in Lemma 10.8, we find

$$P\left(m_{s_2}(u_1 + \lceil km \rceil) \leq \frac{m}{3} \mid F_{h>l}, G_{u_1, u_1 + \lceil km \rceil}\right) \geq \frac{1}{2}, \quad (10.14)$$

where  $F_{h>l} := \{m_{high}(t) > m_{low}(t) = 0 \forall u_1 < t \leq u_1 + \lceil km \rceil\}$ . If the events in (\*), (10.13) and (10.14) are all true, then  $(z_t)_{u_1 \leq t < u_2}$  is in a trajectory to NC. Combining all three results above as well as (10.10), we have proved the claim.

Finally, from (10.7), (10.8) and (10.12), it follows that

$$\min_{w \in BC} P(F | Z_0 = w) > (1/7)(1/16 - 1/200)2^{-\bar{s}}(2^{-\bar{c}\bar{\sigma}} - e^{-\bar{c}\bar{\sigma}}),$$

when  $\mu < g'$ . Now we calculate the expected transition time from BC to NC conditional on F:

$$\max_{w \in BC} T_w(F) = \max_{w \in BC} (E[u_1|F, Z_0 = w] + E[t_1 - u_1|F, Z_0 = w]).$$

By the definition of  $F_3$ ,  $E[t_1 - u_1|F, Z_0 = w] < 1/(2\mu)$ , for all  $w$  in  $BC$ . So,  $\max_{w \in BC} T_w(F) < 1/(2\mu) + 1/(2\mu) = \mu$ . ( $T_w(F)$  is defined in (5.2).) We may also infer from the definition of  $F_2$  that, for all  $w$  in  $BC$ ,

$$T_w(F^c) \leq E[u_1|F, Z_0 = w] + 100/\mu < 1/(2\mu) + 100/\mu.$$

We complete the proof by Lemma 5.1.  $\square$ .

**Comment:** We do not specify the exact value of  $\sigma \geq \bar{\sigma}$  in the above proof. One can see that the proof can be applied to any fixed  $\sigma$  with  $0 < \sigma < 1$  or  $\sigma = 1 - \mu$  in the limiting process. In the above proof, all (10.11), (10.13) and (10.14) need  $\sigma$  to be taken into account. However, the proof is the same, as long as all best-reply strategies are assigned with the same probability for each selection step.

## 10.7 Preliminary results to prove Theorem 7.1

In  $\Gamma$ , we denote the set of all nodes where player I and II moves by  $\mathcal{N}_1$  and  $\mathcal{N}_2$ , respectively.

**Lemma 10.10** *Condition 1 implies that player I moves at node  $\bar{N}$ .*

**Proof.** Denote the immediate descendant node of  $\bar{N}$  on path  $pa(V)$  by  $\hat{N} := id(\bar{N}) \cap pa(V)$ . (Recall the definition of  $id(N)$  in Section 4.1.) It is straightforward to see the path from  $\hat{N}$  to  $V$  is the path of the local backward induction equilibrium in subgame  $\Gamma_{\hat{N}}$ . If player II plays at node  $\hat{N}$ , then the local backward induction strategy requires player II to move towards  $\hat{N}$  at

$\bar{N}$ . Thus the path of the global backward induction equilibrium is from the root to  $V$ , since no divergence on path  $pa(\bar{N})$  can give any player a higher payoff. This contradicts with the assumption that  $pa_{SC} \neq pa_{BC}$ .  $\square$

**Lemma 10.11** *Denote by  $NB_2$  the set of nodes where player II takes a move in his backward induction strategy. If  $\Gamma$  satisfies both Condition 1 and Condition 2, then  $\Gamma_{\bar{N}}$  contains no node in  $NB_2$ .*

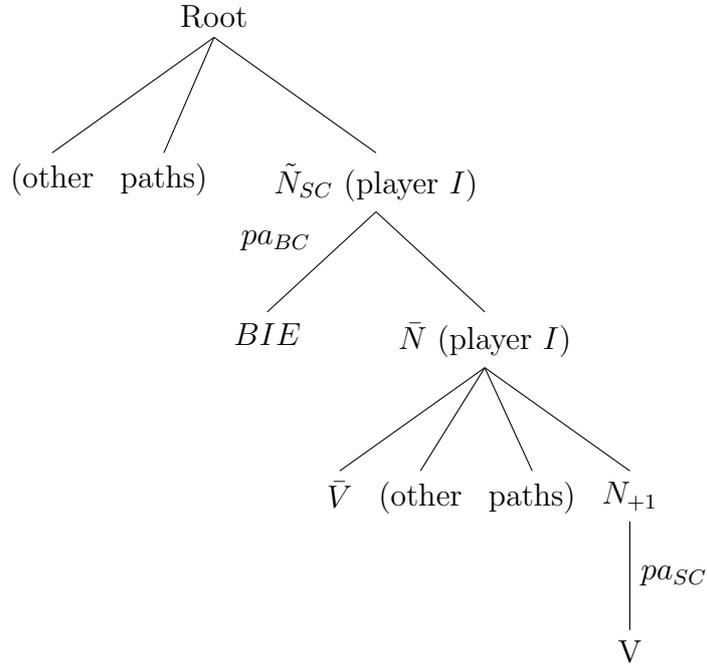


Figure 7: Tree structure of  $\Gamma$

**Proof.** We prove it by contradiction. See Figure 7 for the tree structure of  $\Gamma$ .

**Step 1:** Suppose there exists a node in  $\Gamma_{\bar{N}}$  where player II moves in his backward induction strategy, then player II moves towards  $\bar{N}$  in his backward induction strategy at all nodes in  $pa(\bar{N}) \cap \mathcal{N}_2$ .

**Step 2:** Denote the node  $id(\bar{N}) \cap pa(V)$  by  $N_{+1}$  if it exists. (If not, the analysis in this step can be ignored.) Because the payoff vector  $(u_{SC}^1, u_{SC}^2)$  gives the maximum payoff for each player in  $\Gamma_{N_{+1}}$ , the path of the local backward induction equilibrium starts from  $N_{+1}$  and ends at  $V$ . Hence, at each node in  $d(\bar{N}) \cap pa(V)$ , the player who moves chooses the direction towards  $V$ .

**Step 3:** From Lemma 10.10, we know that player I moves at node  $\bar{N}$ . If he moves towards  $N_{+1}$ , then with the backward induction path in  $\Gamma_{N_{+1}}$  taken into account, player I secures payoff  $u_{SC}^1$ , conditional on that player II plays his backward induction strategy.

**Step 4:** Since we have shown that player II moves along the whole  $pa(V)$  in his backward induction strategy, player I's best reply against it, i.e., his backward induction strategy, should guarantee a payoff at least  $u_{SC}^1$ . However, by Condition 2,  $pa_{BC}$  does not go through  $\bar{N}$ , and hence player I receives a worse payoff in  $BC$ . Contradiction.  $\square$

**Lemma 10.12** *If  $\Gamma$  satisfies both Condition 1 and Condition 2, then in subgame  $\Gamma_{\bar{N}}$ , the path of the local backward induction equilibrium is from  $\bar{N}$  to  $\bar{V}$ .*

**Proof.** We prove it by contradiction. Suppose that the path of the local backward induction equilibrium in  $\Gamma_{\bar{N}}$  is not from  $\bar{N}$  to  $\bar{V}$ . Because the only terminal vertex not dominated by  $V$  is  $\bar{V}$ , the path of the local backward induction equilibrium in  $\Gamma_{\bar{N}}$  must be from  $\bar{N}$  to  $V$ , and the path  $pa_{BC}$  in  $\Gamma$  must be from the root to  $V$ . This contradicts Condition 2.

$\square$

Suppose player I moves at a node  $N$ , we define the set of immediate next successor nodes controlled by player I by  $id^1(N)$ . To put it formally, we

define  $id^1(N)$  to be a set with the following two propositions. (Recall that  $p(N)$  is the set of all predecessor nodes of  $N$ .)

1.  $id^1(N) \subset d(N)$
2.  $\forall N' \in id^1(N), \mathcal{N}_1 \cap (p(N') \setminus (p(N) \cup \{N\})) = \emptyset$

**Lemma 10.13** *Consider a reduced strategy  $s_1$  of player I which satisfies the following propositions.*

1. *At every node in  $p(\bar{N}) \cap \mathcal{N}_1$ , player I moves towards node  $\bar{N}$ .*
2. *For every node  $N$  in  $p(\bar{N}) \cap \mathcal{N}_1$ , player I plays a local backward induction strategy in all subtrees  $\Gamma_{N'}$  where  $N' \in id^1(N)$ .*

*The best response  $s_2$  of player II against  $s_1$  includes no move in  $\Gamma_{\bar{N}}$ .*

**Proof.** From Lemma 10.12, it follows that player I's induced strategy of  $s_1$  in  $\Gamma_{\bar{N}}$  is consistent with the path of local backward induction equilibrium in  $\Gamma_{\bar{N}}$ , which is from  $\bar{N}$  to  $\bar{V}$ .

We prove the statement by contradiction. Suppose that player II's best response  $s_2$  consists of some move at a node in  $\Gamma_{\bar{N}}$ . It follows that the induced strategy of  $s_2$  in  $\Gamma_{\bar{N}}$  is the local backward induction strategy in  $\Gamma_{\bar{N}}$ . We may also infer that player II moves towards  $\bar{V}$  at every node in  $pa(\bar{N}) \cap \mathcal{N}_2$  in strategy  $s_2$ , since  $s_2$  includes a move in  $\Gamma_{\bar{N}}$ . So the path generated by  $(s_1, s_2)$  starts from the root and goes through node  $\bar{N}$ .

We claim that  $(s_1, s_2)$  is the backward induction strategy profile in  $\Gamma$ . To see this, firstly recall that  $s_1$  and  $s_2$  are consistent with the path of backward induction equilibrium in  $\Gamma_{\bar{N}}$ . Suppose a node  $N$  in  $\{\bar{N}\} \cup p(\bar{N})$  is such that  $s_1$  and  $s_2$  are consistent with the path of the local backward induction equilibrium in subgame  $\Gamma_N$ . If the node  $ip(N)$  is played by player I, then the

move towards  $\bar{V}$  gives him the highest possible payoff. So,  $s_1$  is the backward induction strategy of player I in  $\Gamma_{ip(N)}$ . If  $ip(N)$  is played by player II, we know from the paragraph above that the best response in  $\Gamma_{ip(N)}$  against  $s_1$  conditional on reaching node  $ip(N)$  is moving towards  $\bar{V}$ . In this way, we conclude that  $(s_1, s_2)$  is the backward induction strategy profile in  $\Gamma$ . Thus, path  $pa_{BC}$  goes through node  $\bar{N}$ , which contradicts Condition 2.  $\square$

**Lemma 10.14** *Suppose that  $\Gamma$  satisfies Condition 1 and 2. In subgame  $\Gamma_{\bar{N}}$ , there are at least two reduced strategies of player II. In the whole game  $\Gamma$ ,  $s_2 > 2$ .*

**Proof.** If there is only one reduced strategy in  $\Gamma_{\bar{N}}$ , then  $pa_{SC}$  cannot terminate at  $V$ . Since  $pa_{BC}$  does go through  $\bar{N}$ , there is at least one more reduced strategy of player II beyond  $\Gamma_{\bar{N}}$ .  $\square$

We present below a generalised version of Lemma 10.6.

**Corollary 10.15** *Suppose a natural number  $r$  is given. Let  $(Y_i)_{i \in \mathbb{N}}$  be an infinite sequence of i.i.d. random variables with  $P(Y_i = l) = 1/r$  for all  $1 \leq l \leq r$  and for all  $i \in \mathbb{N}$ . For each natural number  $t$  and  $l$  with  $1 \leq l \leq r$ , define  $N_t^l = \sum_{i=1}^t \mathbb{1}_{Y_i=l}$ . Then, given any  $k > 1$ , there exists a natural number  $s_k$  such that, for all  $n > 0$ , we have for any  $l$  with  $1 \leq l \leq r$*

$$P(kN_t^l > \frac{t}{r} \forall 0 < t \leq n) \geq r^{-s_k}.$$

## 10.8 Sketch proof of Theorem 7.1

### 10.8.1 The transition from $BC$ to $SC$

Take a pure strategy of player I in  $SC$  as  $s_3$ . Denote the payoff vector of  $BC$  by  $(u_{BC}^1, u_{BC}^2)$ . For the equilibrium component  $BC$ , we enumerate

all pure Nash equilibria in it as  $((s_1^1, s_1^2), (s_2^1, s_2^2), \dots, (s_i^1, s_i^2))$ . Denote  $S^1 := (s_1^1, s_2^1, \dots, s_i^1)$  and  $S^2 := (s_1^2, s_2^2, \dots, s_i^2)$ . We define  $\bar{u}$  to be the best payoff to player I when he uses the strategy  $s_3$  against a strategy in  $S^2$  of player II, i.e.,  $\bar{u} := \max_{s \in S^2} u^1(s_3, s)$ . Note that  $\bar{u} < u_{BC}^1 < u_{(1)}^1$ . So there exists a positive number  $\rho < 1$  with

$$(1 - \frac{\rho}{\varsigma_2 - 2})\bar{u} + \frac{\rho}{\varsigma_2 - 2}u_{(1)}^1 < u_{BC}^1. \quad (10.15)$$

(Recall Lemma 10.14 that  $\varsigma_2 > 2$ .)

We only need to prove that there exists a positive number  $c$  such that for all finite  $m$ ,

$$\lim_{\mu \rightarrow 0} \pi_{\mu, m}[NC] > c.$$

As before, we consider the case of the best-reply selection mechanism, and the proof is essentially the same for the better-reply selection mechanism. Without loss of generality, we suppose that there exists a number  $\bar{\sigma}$  with  $0 < \bar{\sigma} < 1$  such that  $\sigma \geq \bar{\sigma}$  in the limiting process of evolutionary dynamics.

We reset  $c_{\bar{\sigma}}$  to be  $\exp \frac{-200(\varsigma_2 - \rho)}{\rho \bar{\sigma}}$  in Lemma 10.3. We take  $k = 1.5$  in Corollary 10.15, and in this proof denote  $s_{1.5}$  by  $\bar{s}$ . We further denote  $\varsigma_2^{-\bar{s}} g_{\bar{\sigma}}(m)$  by  $g'$ , and assume  $\mu < g'$ . (See Lemma 10.3 for the definition of  $g_{\bar{\sigma}}(m)$ .) As in the proof of Proposition 6.1, our main goal is to calculate  $U(BC, NC)$ . To this end, we are going to define a finite sequence of events  $(F_1, \dots, F_{10})$  with respect to the Markov chain  $(Z_t)_{t \geq 0}$ . For simplicity, we denote  $\bigcap_{i=1}^l F_i$  by  $D_l$  for all  $l$  with  $1 \leq l \leq 10$ . Thus  $D_l = D_{l-1} \cap F_l$ . To deduce  $U(BC, NC)$ , we assume  $Z_0 \in BC$  at stage 0. We shall see the argument below is independent to the exact initial state in  $BC$ . We enumerate the stages that at least one mutation happens as  $u_1, u_2, \dots$

Denote  $\mathcal{N}_1(\bar{N})$  and  $\mathcal{N}_2(\bar{N})$  to be the set of all strategies of player I and player II which requires at least one move in subgame  $\Gamma_{\bar{N}}$ , respectively.

**First period:** events  $(F_1, \dots, F_6)$  between stage  $u_1$  and  $u_2$

[Event  $D_6 = \bigcap_{i=1}^6 F_i$  will push the majority of  $M(2)$  to a strategy in  $\mathcal{N}_2 \setminus \mathcal{N}_2(\bar{N})$ . If the mutant at  $u_2$  happens to be an individual in  $M(1)$  playing a pure strategy in  $SC$ , then the distribution of the converted  $M(2)$  on  $\mathcal{N}_2(\bar{N})$  will possibly be approximately uniform all the time from  $u_2$ . Thus  $M(1)$  will favour a pure strategy in  $SC$  from some time after  $u_2$ . See Comment 1 at the end of this section for more details.]

Recall the strategy  $s_1$  defined in Lemma 10.13.

**Event**  $F_1 := \{Z_{u_1}^1(q^1) = s_1, Z_{u_1}^2(q) = Z_0^2(q) \forall q \in M(2)\}$ . [The first mutation is in population  $M(1)$  only and that yields a  $s_1$  strategy.]

With the similar analysis as in (10.7), we conclude that, when  $\mu < 1/2$ , for all  $w$  in  $BC$ ,

$$P(F_1|Z_0 = w) = (1 - \mu)/(2\zeta_1) > 1/(4\zeta_1). \quad (10.16)$$

**Event**  $F_2 := \{\mu^{-1} < u_2 - u_1 < 100/\mu\}$ . [The number of stages between the first and the second mutation is between  $\mu^{-1}$  and  $100\mu^{-1}$ .]

With the similar analysis as in (10.8), we conclude that, when  $\mu < 1/4$ , for all  $w$  in  $BC$ ,

$$P(F_2|Z_0 = w) > 1/16 - 1/200. \quad (10.17)$$

**Event**  $F_3 := \{Z_{u_2-1} \in BC\}$ . [After state transitions triggered by the mutant, it converts back to  $BC$ .]

If  $Z_{u_2-1} \notin BC$ , then we do not need to consider the events  $F_4, \dots, F_9$  and  $F_{10}$  below. See Comment 2 at the end of this section for more details about  $F_3$ .

We denote  $t_1$  to be the first stage after  $u_1$  when the mutant  $q^1$  changes his strategy, i.e.,  $t_1 = \min\{t > u_1 : Z_t^1(q^1) \neq Z_{u_1}^1(q^1)\}$ .

**Event**  $F_4 := \{t_1 - u_1 < 1/(2\mu)\}$ . [From stage  $u_1$ , it takes less than  $1/(2\mu)$  stages back into  $NE$ .]

With the similar analysis as in (10.10), we conclude that, for all  $w$  in  $BC$ ,

$$P(F_4|D_3, Z_0 = w) \geq 1 - \exp \frac{-200(\varsigma_2 - \rho)}{\rho\bar{\sigma}}. \quad (10.18)$$

**Event**  $F_5 := \{t_1 - u_1 > \lceil \frac{2(\varsigma_2 - \rho)m}{\rho\bar{\sigma}} \rceil\}$  [From stage  $u_1$ , it takes more than  $\lceil \frac{2(\varsigma_2 - \rho)m}{\rho\bar{\sigma}} \rceil$  stages for the mutant to change his strategy from  $s_1$ .]

With the similar analysis as in  $F_4$  in Example 1 in [13], we conclude that, when  $\mu < g'$ , for all  $w$  in  $BC$ ,

$$P(F_5|D_3, Z_0 = w) > 2^{-\frac{4(\varsigma_2 - \rho)}{\rho\bar{\sigma}} - 1} > \exp\left(-\frac{4(\varsigma_2 - \rho)}{\rho\bar{\sigma}} - 1\right).$$

Thus, for all  $w$  in  $BC$ ,

$$\begin{aligned} & P(D_5, Z_0 = w) \\ & \geq P(D_3, Z_0 = w) \left( \exp\left(-\frac{4(\varsigma_2 - \rho)}{\rho\bar{\sigma}} - 1\right) - \exp\left(-\frac{200(\varsigma_2 - \rho)}{\rho\bar{\sigma}}\right) \right). \end{aligned} \quad (10.19)$$

We denote  $c_5(\bar{\sigma}) := \exp\left(-\frac{4(\varsigma_2 - \rho)}{\rho\bar{\sigma}} - 1\right) - \exp\left(-\frac{200(\varsigma_2 - \rho)}{\rho\bar{\sigma}}\right)$  for simplicity.

**Event**  $F_6 := \{\sum_{s \in \mathcal{N}_2(\bar{N})} m_s(u_2) \leq \rho m / \varsigma_2\}$ . [At the stage of the second mutation, the proportion of population  $M(2)$  playing a move in  $\Gamma_{\bar{N}}$  is no greater than  $\rho m / \varsigma_2$ .]

With the similar analysis as in  $F_5$  in Example 1 in [13] (and some insight), we conclude that, for all  $w$  in  $BC$ ,

$$P(D_6, Z_0 = w) \geq P(D_5, Z_0 = w)/2. \quad (10.20)$$

**Second period:** events  $(F_7, \dots, F_{10})$  between stage  $u_2$  and  $u_3$

[Transition to SC.]

Recall that  $s_3$  is a pure strategy of player I in  $SC$ .

**Event**  $F_7 := \{Z_{u_2}^1(q^1) = s_3, Z_{u_2}^2(q) = Z_{u_2-1}^2(q), \forall q \in M(2)\}$ . [The second mutation is in population  $M(1)$  only and that yields an  $s_3$  strategy.]

With the similar analysis as in (10.7), we conclude that, when  $\mu < 1/2$ , for all  $w$  in  $BC$ ,

$$P(F_7|D_6, Z_0 = w) = (1 - \mu)/(2\zeta_1) > 1/(4\zeta_1). \quad (10.21)$$

**Event**  $F_8 := \{\mu^{-1} < u_3 - u_2 < 100/\mu\}$ . [The number of stages between the second and the third mutation is between  $\mu^{-1}$  and  $100\mu^{-1}$ .]

With the similar analysis as in (10.8), we conclude that, when  $\mu < 1/4$ , for all  $w$  in  $BC$ ,

$$P(F_8|D_7, Z_0 = w) > 1/16 - 1/200. \quad (10.22)$$

We denote  $t_2$  to be the first stage after  $u_2$  when the state is in  $NE$ , i.e.,  $t_2 = \min\{t > u_2 : Z_t \in NE\}$ . Denote  $2 + \frac{4\rho}{\bar{\sigma}(\zeta_2 - 2 - \rho)}$  by  $\bar{c}_{\bar{\sigma}}$ .

**Event**  $F_9 := \{t_2 - u_2 < 1/(2\mu)\}$ . [From stage  $u_2$ , it takes less than  $1/(2\mu)$  stages back into  $NE$ .]

With the similar analysis as in (10.10), we conclude that, for all  $w$  in  $BC$ ,

$$P(F_9|D_8, Z_0 = w) \geq 1 - e^{-\bar{c}_{\bar{\sigma}}\zeta_2^{-\bar{s}}}. \quad (10.23)$$

**Event**  $F_{10} := \{\exists t : u_2 < t < u_3, Z_t \in NC\}$ . [The process enters into  $NC$  at some stage between the second and the second mutation.]

**Claim:**

$$P(F_9 \cap F_{10}|D_8, Z_0 = w) \geq \zeta_2^{-\bar{s}}(2^{-\bar{c}_{\bar{\sigma}}} - e^{-\bar{c}_{\bar{\sigma}}}) \forall w \in BC. \quad (10.24)$$

Denote  $s_4$  to be a pure strategy of player II which is the best response to  $s_3$ . (Thus  $(s_3, s_4)$  is a pure Nash equilibrium.) The mutant playing  $s_3$  attracts  $M(2)$  to divert to a strategy in  $\mathcal{N}_2(\bar{N})$ .

For any two stages  $\tau_1$  and  $\tau_2$  with  $\tau_1 < \tau_2$ , denote the difference of number of individuals playing a strategy in  $s_4$  and  $\mathcal{N}_2(\bar{N})$  by

$$\Delta_{s_4}(\tau_1, \tau_2) := (m_{s_4}(\tau_2) - m_{s_4}(\tau_1))$$

and

$$\Delta_{\bar{N}}(\tau_1, \tau_2) := \sum_{s \in \mathcal{N}_2(\bar{N})} (m_s(\tau_2) - m_s(\tau_1)),$$

respectively.

Given two stages  $\tau_1 < \tau_2$ , denote the event of no mutation in both populations between stage  $\tau_1$  and  $\tau_2$  by  $G_{\tau_1, \tau_2}$ , and the event of

$$2\Delta_{s_4}(\tau_1, t) \geq \Delta_{\bar{N}}(\tau_1, t)/(\varsigma_2 - 1) \quad \forall t \text{ with } \tau_1 < t < \tau_2$$

by  $H_{\tau_1, \tau_2}$ . We take  $m - \sum_{s \in \mathcal{N}_2(\bar{N})} m_s(u_2)$  as number  $n$  in Corollary 10.15, and find that, given any  $t > u_2$ , (for example, take  $t$  as  $\min\{t' > u_2 : \sum_{s \in \mathcal{N}_2(\bar{N})} m_s(t') = m\}$ )

$$P \left( H_{u_2, t} \mid G_{u_2, t}, m_{s_3}(t') = \sum_{s \in \mathcal{N}_1(\bar{N})} m_s(t') = 1 \quad \forall t' \text{ with } u_2 < t' < t \right) \geq \varsigma_2^{-\bar{s}}. \quad (10.25)$$

We show that it is with positive probability that the mutant will keep playing  $s_3$  for sufficiently long time.

For the original Markov chain  $(Z_t)_t$ , we take  $k = 2\rho/(\bar{\sigma}(\varsigma_2 - 2 - \rho))$  in Lemma 10.7 and conclude that the probability is greater than  $2^{-1-2k}$  that the mutant will keep playing strategy  $s_3$  between stage  $u_2$  and  $u_2 + \lceil km \rceil$ .

(\*) Taking the same  $k$  and  $\lambda = 1 - \frac{\rho}{\varsigma_2 - 2}$  in Lemma 10.8, we find

$$P \left( \sum_{s \in \mathcal{N}_1(\bar{N})} m_s(u_2 + \lceil km \rceil) \geq \frac{\rho m}{\varsigma_2 - 2} \mid E_{h>l}, G_{u_2, u_2 + \lceil km \rceil} \right) \geq \frac{1}{2} \quad (10.26)$$

where

$$E_{h>l} := \{m_{s_3}(t') = \sum_{s \in \mathcal{N}_1(\bar{N})} m_s(t') > 0 \quad \forall u_2 < t' \leq u_2 + \lceil km \rceil\}.$$

**Comment 1:** Suppose the current state is in  $BC$ . From Condition 2, it follows that no individual in population I is moving at a node in  $\Gamma_{\bar{N}}$ . If every

individual in  $M(2)$  is playing plays a strategy in  $S^2$ , then by Lemma 10.11, no individual in  $M(2)$  is moving at a node in  $\Gamma_{\bar{N}}$ .

However, when without this  $S^2$  constraint, there may be some individuals in  $M(2)$  playing a strategy in  $\mathcal{N}_2(\bar{N})$ , or even any reduced strategy. It's the average payoff to  $M(1)$  such that no individual prefers to deviate from  $S^1$ , and hence the state stays in  $BC$ . So, if a mutation happens to be a pure strategy in  $SC$  of an individual in  $M(1)$ , then the following scenario might be true. In the case that a significant proportion of  $M(2)$  is playing a strategy consistent with the path  $pa_{\bar{v}}$ , a pure strategy of player I containing a local backward induction strategy in  $\Gamma_{\bar{N}}$  might be the best response for  $M(1)$  at some stage after  $u_1$ . That will make  $M(2)$  prefer a strategy with a local backward induction strategy in  $\Gamma_{\bar{N}}$  and then two populations will begin to convert to a trajectory back to  $BC$ . See the example below in Figure 8.

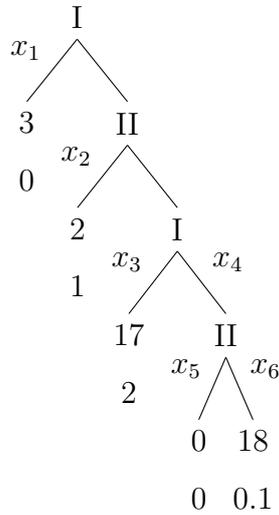


Figure 8: Example in Comment 1

We consider the case of large populations with very small mutation rate. Suppose  $Z_0 \in BC$  where  $m_{x_1} = m$ ,  $m_{x_2} = \lceil \frac{15m}{16} \rceil$ ,  $m_{x_5} = 0$  and  $m_{x_6} = \lfloor \frac{m}{16} \rfloor$ . We further suppose that a mutation happens at stage 1 such that  $m_{x_3} =$

$1 > m_{x_4} = 0$ , and that this mutation will stay at  $x_3$  for sufficiently long time. Then,  $M(2)$  is under selection pressure, and an individual playing  $x_2$  would like to choose  $x_5$  or  $x_6$  randomly. Let's assume a stage that  $m/256$  individuals have changed their strategies from  $x_2$  to  $x_5$  or  $x_6$ . By strong law of large numbers, it would be very close to the result of  $m/512$  to  $x_5$  and  $m/512$  to  $x_6$ . We now check the average payoff of strategy  $x_4$  of player I against this distribution of  $M(2)$

$$u_{x_4} = 2 \left( \frac{15}{16} - \frac{1}{256} \right) + 18 \left( \frac{1}{512} + \frac{1}{16} \right) = \frac{775}{256} \approx 3.027$$

and the average payoff of  $x_3$

$$u_{x_3} = 2 \left( \frac{15}{16} - \frac{1}{256} \right) + \frac{17}{256} = \frac{495}{256} < 3$$

So  $x_4$  is the best response at that time, and  $M(1)$  is under the selection pressure to  $x_4$ . If the bottom node is connected by such drift of  $M(1)$ , then  $x_5$  is dominated by  $x_6$ . Recall that  $(x_3, x_5)$  is the pure Nash equilibrium in  $NC$ . So the transition triggered by the mutant playing  $x_3$  will move back to  $BC$  with probability close to 1.

**Comment 2:** We continue with the analysis in Comment 1. For a state in  $BC$ , the proportion profile of  $M(2)$  playing each strategy has many possibilities. If one can force the majority of  $M(2)$  drifting away from  $\mathcal{N}_2(\bar{N})$ , then it is with bounded probability that the state enters  $SC$  after the second mutation of an individual in  $M(1)$  adopting  $s_3$ .

A mutation of an individual in  $M(1)$  to  $s_1$  may trigger such desired drift of  $M(2)$ . From Lemma 10.13, we know that the best response of  $M(2)$  against  $s_1$  is some strategy in  $\mathcal{N}_2 \setminus \mathcal{N}_2(\bar{N})$ .

One may note that this mutation may lead a transition to a Nash equilibrium not in  $SC$  or  $BC$ , i.e., to some Nash equilibrium in  $NC \setminus BC$ . For

our purpose, we only need to consider the case that it still stays in  $BC$  at stage  $u_2 - 1$ , and that will be shown in the Second Period.

### 10.8.2 Proof of Theorem 7.1 by event $F$

We check why the transition is in a trajectory if all events above, i.e.,  $D_9$ , (10.25), (\*) and (10.26), are true. Recall that in (10.20), there are at most  $\rho m/\varsigma_2$  individuals in  $M(2)$  playing a strategy in  $\mathcal{N}_2(\bar{N})$  just before the second mutation. Consider the state at stage  $\hat{t} := u_2 + \lceil km \rceil$ . From (10.26) at least  $\rho m/(\varsigma_2 - 2)$  individuals are playing a strategy in  $\mathcal{N}_2(\bar{N})$  at  $\hat{t}$ . So there are at least

$$\frac{\rho m}{\varsigma_2 - 2} - \frac{\rho m}{\varsigma_2} = \frac{2\rho m}{\varsigma_2^2 - 2\varsigma_2}$$

individuals converting to a strategy in  $\mathcal{N}_2(\bar{N})$  between stage  $u_2$  and  $\hat{t}$ . From (10.25), we may further infer that at least  $\frac{\rho m}{\varsigma_2^2 - 2\varsigma_2}$  among them are playing  $s_4$  at stage  $\hat{t}$ . Denote by  $\hat{T}$  the minimum stage that the number of individuals in  $M(2)$  playing a strategy in  $\mathcal{N}_2(\bar{N})$  is  $\lfloor \rho m/(\varsigma_2 - 2) \rfloor$ . Then  $\hat{T} \leq \hat{t}$ .

Suppose an individual  $q$  in  $M(1)$  is playing a strategy in  $\mathcal{N}_1(\bar{N})$  at stage  $\hat{T}$ . Let's consider its average payoff against the induced distribution of  $M(2)$  which play a strategy in  $\mathcal{N}_2(\bar{N})$  at stage  $\hat{T}$ . If  $q$  plays  $s_3$ , then the average payoff is  $u_{SC}^1$ . If  $q$  plays a strategy consistent with  $pa(\bar{V})$ , e.g.  $s_1$ , then the average payoff is at most  $\max(\frac{u_i^1}{\varsigma_2^2} + \frac{(\varsigma_2^2 - 1)u_{(1)}^1}{\varsigma_2^2})$  for all  $i$  with  $u_i^1 \neq u_{(1)}^1$  or  $u_{SC}^1$ . By Condition 1, we know that  $s_3$  gives the best payoff which can be generated by a strategy in  $\mathcal{N}_1(\bar{N})$  at  $\hat{T}$ .

Meanwhile, from (10.15), we know that a pure strategy of player I in  $BC$  is still the (global) best reply for  $M(1)$  at least until stage  $\hat{T}$ . After  $\hat{T}$ , with  $M(2)$  drifting to strategies in  $\mathcal{N}_2(\bar{N})$ , pure strategies of player I in  $SC$  will become the best replies for population  $M(1)$  from some stage at least until  $u_3$ , if (10.25) is true. If  $F_9$  is true, then it will converge to a state in  $SC$  before

$u_3$ . In summary, none strategy consistent with  $pa(\bar{V})$  is a best response for  $M(1)$  at any stage between  $u_2$  and  $u_3$ .

Define the event that the state enters  $NC$  before the third mutation by  $F := \{\exists t < u_3 \text{ s.t. } Z_t \in NC\}$ . Then  $F \supset ((D_2 \cap F_3^c) \cup D_{10})$ . From (10.16), (10.17), (10.19), (10.20), (10.21), (10.22) and (10.24), it follows that

$$\min_{w \in BC} P(F|Z_0 = w) > \frac{1}{(4\zeta_2)^2} \left( \frac{1}{16} - \frac{1}{200} \right)^2 c_5(\bar{\sigma}) \zeta_2^{-\bar{s}} \frac{2^{-\bar{c}\bar{\sigma}} - e^{-\bar{c}\bar{\sigma}}}{2}.$$

Now we calculate the maximum expected transition time from  $BC$  to  $NC$  conditional on  $F$ : for each  $w$  in  $BC$ ,

$$T_w(F) \leq E[u_1|F, Z_0 = w] + E[u_2 - u_1|F, Z_0 = w] + E[t_2 - u_2|F, Z_0 = w].$$

(Recall  $T_w(F)$  is defined in 5.2.) For all  $w$  in  $BC$ , by the definition of  $F_2$ ,  $E[u_2 - u_1|F, Z_0 = w] < 100/\mu$ ; by the definition of  $F_9$ ,  $E[t_2 - u_2|F, Z_0 = w] < 1/(2\mu)$ . So,

$$\max_{w \in BC} T_w(F) < 1/(2\mu) + 100/\mu + 1/(2\mu) = 101/\mu.$$

It follows from the definition of  $F_2$  and  $F_8$  that, for all  $w$  in  $BC$ ,

$$T_w(F^c) \leq (E[u_1|F, Z_0 = w] + 100/\mu + 100/\mu) < 1/(2\mu) + 100/\mu + 100/\mu.$$

We complete the proof by Lemma 5.1.

## Acknowledgements

The author is grateful to Sergiu Hart for many suggestions and discussions. The author would also like to thank Katsuhiko Aiba, Tomas Rodriguez Baraquero, Yosef Rinott and Eyal Winter for their comments.

## References

- [1] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*, MIT Press, Cambridge, 2003.
- [2] R. Cressman and K.H. Schlag, The dynamic (in)stability of backward induction, *J. Econ. Theory* 83 (1998) 260–285.
- [3] D. Fudenberg, D.K. Levine, *The Theory of Learning in Games*, MIT Press, 1998.
- [4] D. Foster and H.P. Young, Stochastic evolutionary game dynamics, *Theor. Pop. Biol.* 38 (1990) 219–232.
- [5] J. Gale, K. Binmore and L. Samuelson, Learning to be imperfect: The ultimatum game, *Games Econ. Behav.* 8 (1995) 56–90.
- [6] H. Gintis, R. Cressman and T. Ruijgrok, Subgame Perfection in Evolutionary Dynamics with Recurrent Perturbations, in *Handbook of Research on Complexity* (ed. J. Barkley Rosser), 353-368, Edward Elgar Publishing, Northampton, MA, 2009.
- [7] Z. Gorodeisky, Evolutionary stability for large populations and backward induction, *Mathematics of Operations Research* 31 (2006) 369–380.
- [8] S. Hart, Evolutionary dynamics and backward induction, *Games Econ. Behav.* 41 (2002) 227–264.
- [9] M. Kandori, G. Mailath and R. Rob, Learning, mutation, and long-run equilibrium in games, *Econometrica*, 61 (1993) 29–56.

- [10] L. Samuelson, *Evolutionary Games and Equilibrium Selection*, MIT Press, 1997.
- [11] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, MIT Press, 2010.
- [12] Z. Xu, *Evolutionary stability in finite stopping games under a fast best-reply dynamics*, Center for the Study of Rationality, Hebrew University of Jerusalem, Mimeo.
- [13] Z. Xu, *The instability of backward induction in evolutionary dynamics*, Center for the Study of Rationality, Hebrew University of Jerusalem, Mimeo.
- [14] H.P. Young, *The evolution of conventions*, *Econometrica*, 61 (1993) 57–84.
- [15] H.P. Young, *Individual Strategy and Social Structure*, Princeton University Press, 1998.