

Xu, Zibo

**Working Paper**

## Stochastic stability in finite extensive-form games of perfect information

SSE/EFI Working Paper Series in Economics and Finance, No. 743

**Provided in Cooperation with:**

EFI - The Economic Research Institute, Stockholm School of Economics

*Suggested Citation:* Xu, Zibo (2013) : Stochastic stability in finite extensive-form games of perfect information, SSE/EFI Working Paper Series in Economics and Finance, No. 743, Stockholm School of Economics, The Economic Research Institute (EFI), Stockholm

This Version is available at:

<https://hdl.handle.net/10419/122146>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Stochastic stability in finite extensive-form games of perfect information

Zibo Xu<sup>1</sup>

Center for the Study of Rationality, the Hebrew University of Jerusalem  
Department of Economics, Stockholm School of Economics

SSE/EFI Working Paper Series in Economics and Finance  
No 743  
March 2013

## Abstract

We consider a basic stochastic evolutionary model with rare mutation and a best-reply (or better-reply) selection mechanism. Following Young's papers [29] and [30], we call a state stochastically 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, the best-reply dynamic converges to a Nash equilibrium almost surely. Moreover, only Nash equilibria can be stochastically stable. We present a 'centipede-trust game', where we prove that both the backward induction equilibrium component and the Pareto-dominant equilibrium component are stochastically stable, even when the populations increase to infinity. For finite extensive-form games of perfect information, we give a sufficient condition for stochastic stability of the set of non-backward-induction equilibria, and show how much extra payoff is needed to turn an equilibrium stochastically stable.

*JEL classification:* C7, D7, C6.

---

<sup>1</sup>This working paper is a revised version of 'Evolutionary stability in general extensive-form games of perfect information' in Discussion Paper Series 631, the Center for the Study of Rationality, Hebrew University of Jerusalem. The author is grateful to Sergiu Hart and Jorgen Weibull for many suggestions and discussions. The author also wishes to thank Katsuhiko Aiba, Tomas Rodriguez Barraquer, Yosef Rinott, Bill Sandholm and Eyal Winter for their comments. The author would like to acknowledge financial support from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement No. 249159, and from the Knut and Alice Wallenberg Foundation.

*Keywords:* Evolutionary game theory, Markov chains, equilibrium selection, stochastic stability, games in extensive form, games of perfect information, backward induction equilibrium, Nash equilibrium components, best-reply dynamics.

# 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 call a state *stochastically stable* if its long-term relative frequency of occurrence is bounded away from zero as the mutation rate decreases to zero. This notion is proposed by Foster and Young in [5], by Young in [29] and [30]. It is also called ‘long-run equilibrium’ by Kandori et al. in [14], ‘in the support of the limit distribution’ by Samuelson in [25], by Fudenberg and Levine in [7], and ‘evolutionarily stable’ by Hart in [12].

Such dynamic evolutionary models have been extensively studied under various selection mechanism for games in strategic form. (See [14], [29], [26] and many other literatures.) They show that some Nash equilibria, e.g. the risk-dominant ones, are more stable than others under certain circumstance. In general normal-form games, the evolutionary process brings *inertia* to the model: individuals’ observations are limited and imperfect, and their understanding of the whole dynamic process is constrained. Both these factors support the assumption that in each period 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 an extensive-form game of perfect information under an evolutionary dynamic, some nodes are visited more frequently than the others, which cannot be represented in a normal-form game. One also needs to be careful when applying the inertia argument above. We say that a node is *disconnected* if the probability that node is reached is zero (see (2.1) for rigorous definition). 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 individuals may entirely depends on that single ‘connecting’ individual. If this ‘connecting’ individual diverts away from this node and makes it disconnected again in the next period, then the ones playing at this node may lose the incentive to adjust their strategies. Therefore,

the set of better/best reply strategies can change in every period. This is one basic observation in the dynamics in extensive-form games.

We adopt the dynamic model introduced in Hart's paper [12], which is a natural and simple process that provides for both selection and mutation. The model includes an adapted version of the best-response dynamic discussed in [14] and [15] for a  $k$ -player normal form game, and we apply it here in extensive-form games of perfect information. For symmetric coordination games discussed in [14], they deduce a robust stability result with a general Darwinian property which allows for multiple individuals in each population to change their strategies in each period, while, in our current paper, in each period only one individual in each population may change her strategy. We study the extreme case of all individuals subject to change in our extensive-form framework in another paper [27], which corresponds to the case that the adjustment probability is 1 for each individual in each period in [15]. We show in [27] that the stability result on an extensive-form game may be sensitive to the rate of replacing strategies. In the current paper, we instead focus on the effect of large population on stability result, which is not addressed in [14] or [15]: the stochastic stability result of a Nash equilibrium may be different for the case of large and small population size.

Our analysis approach of a *regular perturbed Markov process* is adapted from [29]. The evolutionary process applied on a normal-form game in [29] is similar to our selection-mutation process. However, Young's model is essentially equivalent to an evolutionary process with a single individual in each population, and a state defined there is a bounded history of play. Our model instead focuses on large populations and the selection is entirely based on the last state, which is the distribution of strategies played in all populations in the last period.

Hart [12] proves that the backward induction equilibrium is the unique stochastically stable state, 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 [10] relaxes this constraint on the product of population size and mutation rate.

However, Hart only considers a specific type of perfect-information games: each player can never play at more than one node in a game. We use the basic dynamic model of population games proposed in Hart's paper, but dispense with this constraint. Our first conclusion is that our best-reply dynamics always lead to a Nash equilibrium in finite extensive-form games of perfect information. We further deduce that, for all such games, regardless of population size, only Nash equilibria can be stochastically stable. It is a natural question whether Hart's stability result still holds in this general

context.

Our starting point is the concept of *Nash equilibrium components*. Here, we adopt the standard definition: a Nash equilibrium component is a maximal connected set of Nash equilibria. In a generic game, all equilibria in the same Nash equilibrium component give rise to identical induced probability distribution over terminal nodes, i.e., they differ only off the equilibrium path. (cf. [16] and [2].) Thus, a backward induction equilibrium component is a Nash equilibrium component whose final outcome is also the outcome of a backward induction equilibrium. When a state is in a backward induction equilibrium component, the strategy distribution of a population at a node off the equilibrium path can be various. 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 stochastically stable result. (See Example 4 in [28] for more details.)

In [12], Hart also shows an example that equilibria other than the backward induction equilibrium may be stochastically 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 unique stochastically stable component even when the populations go to infinity. Moreover, this is a minimum example with respect to the number of nodes in all finite extensive-form games of perfect information with large population. For more details, please refer to the end of [28], where we discuss why such an example needs more than three non-terminal nodes. The results hold for both the best reply and the better reply selection mechanisms.

The evolutionary approach may also explain the emergence of trust and fairness under bounded rationality modeled in a finite extensive-form game. Here we consider the case in which populations of individuals play a game repeatedly where each population 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 [14]. That is, not all individuals need to react instantaneously to their environment; when individuals react, they react myopically; there is a small probability that individuals change their strategies at random. Note that in our example in Section 6, the stochastically stable Nash equilibrium component not containing the backward induction equilibrium is in fact Pareto-efficient. Thus, this evolutionary approach shows that, under the condition of bounded rationality, players select Pareto-efficient equilibria in the long run as well, even if they are not the backward induction equilibrium. (See [23] for the stability result

of Pareto-efficient equilibria in symmetric  $2 \times 2$  coordination games.) Moreover, the Pareto efficiency can be viewed as resulted from trust (reciprocal play) built on a sufficiently long multiple-move game. Indeed, recall that every extensive-form game with no more than three non-terminal nodes has the backward induction equilibrium component as its unique stochastically stable component. We shall discuss it in details in Section 6.

In Section 7, we generalise our example and provide a sufficient condition for the set of non-backward-induction equilibria to be stochastically 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. The earlier example is a warm-up for this general theorem.

We can apply this sufficient condition for stochastic stability to solve the problem of improving stability by increasing payoffs. We firstly observe that, when a payoff vector gives global best payoff to both players, it must be the result of the backward induction equilibrium in the game, and the backward induction equilibrium component is stochastically 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 that equilibrium component become stochastically stable? 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 the current state is in the backward induction equilibrium component. Then very few individuals in population I are playing off the backward induction equilibrium path. As the number of individuals playing at each node is not fixed, a small change in the proportions of strategies in population I triggered by mutation may bring a huge effect to the relative proportions of strategies in population I in a subgame off the backward induction equilibrium path. That may further influence other individuals and other populations. 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’. We shall see

later that the standard technique of perturbed dynamical systems, proposed in [6] and applied as a ‘tree technique’ in [14] and [29], cannot lead to the stochastic stability result in the case of large populations for extensive-form games: that method is only effective for the case of fixed population size. We have to refer to the basic time-average asymptotic behavior of the Markov chain in our analysis.

## 2 The Model

### 2.1 An extensive-form game of perfect information

We adopt the standard definition of a finite extensive-form game of perfection information. (See [17], [11] and [21] for reference.)

Given a set  $N$  of finitely many nodes, we define a partial order binary relation  $\prec$  on  $N$  that represents *precedence*. We further suppose an initial node  $n^0$  as a predecessor of all other nodes in  $N$ . Such  $(N, \prec)$  defines a *tree*  $T$ , and we call  $n^0$  the *root* of  $T$ . We define the *immediate-predecessor function*  $\psi : N \rightarrow N$  such that

$$\psi(n') = \max\{n : n \prec n'\} \quad \forall n' \in N \setminus \{n^0\}$$

and  $\psi(n^0) = \emptyset$ . Let  $\Psi$  be the *predecessor function*  $\Psi : N \rightarrow 2^N$  with

$$\Psi(n') = \{n \in N : n \prec n'\}.$$

We denote  $\psi^{-1}$  to be the *immediate-successor function*. Thus,  $\psi^{-1}(n) = \{n' \in N : n = \psi(n')\}$  for all  $n$  in  $N$ . The *successor function*  $\Psi^{-1}$  can be similarly deduced. We call a node  $n$  a *terminal node* if  $\psi^{-1}(n) = \emptyset$ , and write  $N_t := \{n \in N, \psi^{-1}(n) = \emptyset\}$ .

We say that a sequence  $\{n_1, \dots, n_i\}$  of nodes is a *subplay* in the tree  $T$  if  $n_{j-1} = \psi(n_j)$  for all  $1 < j \leq i$ . If  $n_1 = n^0$  and  $n_i \in N_t$ , then it is called a *play*. Denote the set of all plays by  $H$ .

We define a  $k$ -player extensive-form game of perfect information on the finite tree  $(N, \prec)$ . Denote  $\mathcal{N} = \{\Lambda^0, \Lambda^1, \dots, \Lambda^k\}$  as a partition of  $N \setminus N_t$ , and call it the *assignment of decision nodes*. The members of  $\Lambda^0$  are called *chance nodes*; for each  $i \leq k$ , the members of  $\Lambda^i$  are called the *nodes of player  $i$* . Given a node  $n \in N$ , we put  $\lambda(n)$  as the indicator of which player moves on this node. So  $\lambda(n) = i$ , if  $n \in \Lambda^i$ . For chance nodes, define  $\tau : \psi^{-1}(\Lambda^0) \rightarrow [0, 1]$  to be a probability distribution function such that

$$\sum_{n' \in \psi^{-1}(n)} \tau(n') = 1 \quad \forall n \in \Lambda^0.$$



We define a vector  $\mathbf{v} = (v^1, \dots, v^k)$  such that each  $v^i : H \rightarrow \mathbb{R}$  is a Bernoulli function of player  $i$  for all  $1 \leq i \leq k$ . We call the quadruple  $(T, \mathcal{N}, \tau, \mathbf{v})$  an extensive-form game  $\Gamma$  of perfect information. When we apply such a game  $\Gamma$  to the dynamic process introduced later, we assume  $\Lambda^0 = \emptyset$  for simplicity, and thus  $\Gamma$  can be represented by a tripe  $(T, \mathcal{N}, \mathbf{v})$  later in our framework. (All results in this paper can be generalised to the case including chance nodes.)

For each player  $i$ , a (*pure*) *strategy*  $a^i$  assigns a successor to each node in  $\Lambda^i$ . So  $\psi(a^i(n)) = n$  for all  $n$  in  $\Lambda^i$ . Denote the set of pure strategies of player  $i$  by  $A^i$ , and the set of *pure-strategy profiles* by  $A = \prod_{i=1}^k A^i$ . We denote the probability distribution of play in game  $\Gamma$  for a pure strategy profile  $a$  to be a function  $\rho_a : H \rightarrow [0, 1]$  with  $\sum_{h \in H} \rho_a(h) = 1$ . (Note that  $H$  is finite.) Given a node  $\bar{n}$ , we denote  $H_{\bar{n}} := \{h \in H : \bar{n} \in h\}$  and say that the node  $\bar{n}$  is *connected* under a pure strategy profile  $a$  if

$$\sum_{h \in H_{\bar{n}}} \rho_a(h) > 0. \quad (2.1)$$

When  $\Lambda^0 = \emptyset$ , given such a pure strategy profile  $a$  in  $A$ , we can find a sequence of nodes  $\{n_0, n_1, \dots, n_m\}$  such that  $n_0 = n^0$  and

$$n_{i+1} = a^{\lambda(n_i)}(n_i) \quad \forall 0 \leq i < m.$$

It follows that the last node  $n_m \in N_t$ .

The set of *mixed strategies* for player  $i$  is defined as

$$X^i := \Delta(A^i) = \left\{ \sigma = (\sigma^i(a))_{a \in A^i} : \sigma^i(a) \geq 0 \quad \forall a \in A^i \text{ and } \sum_{a \in A^i} \sigma^i(a) = 1 \right\}. \quad (2.2)$$

So a mixed strategy  $x^i$  is a vector of probabilities assigned to each pure strategy in  $A^i$ . The set of *mixed-strategy profiles* is denoted as  $X = \prod_{i=1}^k X^i$ . We call the induced probability distribution of a mixed-strategy profile  $x$  over plays in  $T$  as the *outcome* of  $x$ . Note that a pure-strategy profile  $a$  generates a *payoff vector*  $\mathbf{u}(a) = \sum_{h \in H} \rho_a(h) \mathbf{v}(h)$ . We can linearly extend it to a mixed-strategy profile  $x$ :

$$\mathbf{u}(x) = \sum_{a \in \text{supp}(x)} \left( \prod_{a^i \in a} x^i(a^i) \right) \mathbf{u}(a). \quad (2.3)$$

A mixed-strategy profile  $x$  is a *Nash equilibrium* of the game  $\Gamma$  if

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

for every  $i \leq k$  and every  $y^i \in X^i$ , where  $x^{-i} := (x^j | 1 \leq j \leq k, j \neq i)$ .

A *subtree* rooted at a node  $n$  is the truncated tree  $(\Psi^{-1}(n) \cup \{n\}, \prec)$ . A *subgame* rooted at node  $n$  is the corresponding subtree with the induced assignment of decision nodes and the payoff function. We denote this subgame by  $\Gamma_n$ . A Nash equilibrium is a *backward induction equilibrium* (also called *subgame-perfect equilibrium*) if it induces a Nash equilibrium in all subgames. Kuhn proved in [18] that there always exists a pure backward induction equilibrium, constructed from the terminal nodes and going towards the root. In this paper, we consider a *generic* finite  $k$ -player game  $\Gamma$  in extensive form with perfect information. In the generic assumption, no player obtains the same payoff at any two terminal nodes. Thus, the backward induction equilibrium in  $\Gamma$  is unique.

For a mixed strategy profile  $x$ , we say that a node  $\bar{n}$  is connected under  $x$  if there exists a pure strategy profile  $a$  with non-zero probability in  $x$  such that node  $\bar{n}$  is connected under  $a$ . A realised play of a Nash equilibrium is also called an equilibrium path.

In an extensive-form game, two different pure strategies for the same player always induce the same probability distributions over plays, if they differ only at disconnected nodes (cf. Proposition 4.1 in [21]). This observation suggests a lower-dimensional representation of an extensive-form game. We call two pure strategies  $a_1^i$  and  $a_2^i$  for player  $i$  *outcome equivalent* and write  $a_1^i \sim a_2^i$  if, with every combination  $a^{-i}$  of strategies for the other players, the outcome generated by these two strategies are always the same, i.e.,

$$\rho_{(a_1^i, a^{-i})}(h) = \rho_{(a_2^i, a^{-i})}(h) \quad \forall h \in H \quad \forall a^{-i} \in A^{-i}. \quad (2.4)$$

Such relationship of outcome equivalence generates for each player  $i$  a partition  $B^i$  of the set  $A^i$ . That means

1. The union of all sets in  $B^i$  equals to  $A^i$ ;
2. Given any  $b^i$  in  $B^i$ , for any two strategies  $a_1^i, a_2^i \in b^i$ , (2.4) holds.

Thus, each  $b^i$  is an equivalence class, and we call  $B^i$  the set of *pure quasi strategies* of player  $i$ , and the set of *pure quasi strategy profiles* is defined as  $B := \prod_{i=1}^k B^i$ . Given a pure quasi strategy profile  $b = (b^1, \dots, b^k)$ , we can find a pure strategy profile  $a = (a^1, a^2, \dots, a^k)$  with  $a^i \in b^i$  for all  $1 \leq i \leq k$ , and we define the payoff vector of profile  $b$  as  $\mathbf{u}(b) := \sum_{h \in H} (\rho_a(h))v(h)$ . The set of *mixed quasi strategies* and the payoff vector of a mixed quasi strategy profile can be defined analogously to (2.2) and (2.3), respectively. Nash equilibria can also be defined with quasi strategies. We use the definition of quasi strategy in this paper. When no ambiguity, we simply refer to a quasi strategy as a strategy.

For instance, given the game  $\bar{\Gamma}$  in Figure 1, for a pure strategy of player I which includes the move  $\alpha_1$  at the root, it must also specify the move she would play at the bottom node. We, however, do not specify it for a quasi strategy, as it is impossible to reach the bottom node in that case. Hence, in this one-player game, there are only three quasi strategies corresponding to  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ , respectively, in our framework. See Appendix 10.1 for comparison between the partition generated from outcome equivalence and the standard representation of reduced normal form for an extensive-form game: outcome equivalence is defined on the realisation of outcome, while a reduced-normal-form strategy is concerning the payoff equivalence.

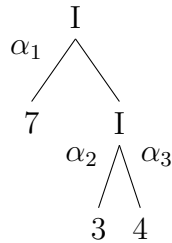


Figure 1: Game  $\bar{\Gamma}$

## 2.2 Evolutionary dynamics

We consider the following *population game* associated to  $\Gamma$ . Assume that for each player  $i$  with  $1 \leq i \leq k$  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  $i$ . An individual is characterised by a pure quasi strategy. (See Section 4.1 for more information.) We define a state  $w$  to be the collection of pure quasi strategies of all individuals in all populations, i.e.,  $w = (w^i)_{1 \leq i \leq k}$  where  $w^i := (w^i(q))_{q \in M(i)}$  and  $w^i(q) \in B^i$  for each  $i \leq k$  and each  $q \in M(i)$ . Let the state space  $W$  be  $\prod_{1 \leq i \leq k} (B^i)^{M(i)}$ . At a state  $w$ , for each  $i$ , let  $m_{b^i}(w)$  be the number of individuals playing strategy  $b^i \in B^i$ , and let  $x_{b^i}(w)$  be the proportion of population  $M(i)$  that plays the strategy  $b^i$ , i.e.,  $x_{b^i}(w) = m_{b^i}(w)/m$ . We put

$$x(w) := ((x_{b^i}(w))_{b^i \in B^i})_{1 \leq i \leq k} \quad (2.5)$$

and thus we may view each state  $w$  as a  $k$ -tuple mixed strategy in  $X$ .

**Comment:** A mixed strategy can be thought of as a distribution of potential participants in a large population, where each potential participant commits to a particular pure strategy. The game is played over and over

again by participants who are randomly drawn from large populations. This idea is suggested in [19], [22], [21], etc.

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 pure quasi strategy  $b$  of some player  $i$ , in period  $t$ , we also abbreviate  $m_b(Z_t)$  by  $m_b(t)$ , which is the random number of individuals in population  $M(i)$  playing the strategy  $a$  in the Markov chain in period  $t$ . 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 [12].

- Conditional on the current state, transitions are independent over populations, i.e.,

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

- In each period, for each  $i$  with  $1 \leq i \leq k$ , 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.,  $\tilde{w}^i(q^i) = b^i$  with probability  $1/|B^i|$  for each  $b^i$  in  $B^i$ .
- Conditional on selection, we define a set of ‘better strategies’, namely

$$BT^i(q^i, w) := \{b^i \in B^i : u^i(b^i, w^{-i}) > u^i(w^i(q^i), w^{-i})\}, \quad (2.6)$$

where  $w^{-i}$  indicates the collection of pure quasi strategies of all individuals in all populations other than  $M(i)$ . 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 randomly chosen in a *better-reply distribution*  $p^i(\cdot)$ :

$$p^i(\tilde{w}^i(q^i) = b^i | w^i(q^i) \notin BT^i) > \kappa \quad \forall b^i \in BT^i, \quad (2.7)$$

for a constant  $\kappa > 0$ . (This corresponds to the selection mechanism used in [14] and [15].) We further assume that  $p(\cdot) := (p^i(\cdot))_i$  is independent of period  $t$ . 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

$$BS^i(w) := \arg \max_{b^i \in B^i} u^i(b^i, w^{-i}). \quad (2.8)$$

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 in the following probability distribution:

$$p^i(\tilde{w}^i(q^i) = b^i | w^i(q^i) \notin BS^i(w)) > \kappa \quad \forall b^i \in BS^i(w) \quad (2.9)$$

for a constant  $\kappa > 0$ , and that  $p(\cdot) := (p^i(\cdot))_i$  is independent of period  $t$ . The only difference between the models with best-reply and better-reply mechanism is in this selection procedure.

### 3 Main Results

To give the formal definition of stochastic 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 [30]. Given two states  $w$  and  $w'$ , the resistance of the transition  $w \rightarrow w'$  defined in [30] is equivalent to the minimum number of mutations needed for a transition from  $w$  to  $w'$  in our context.  $\square$

For a fixed population size, a state  $w$  in  $W$  is called *m-stochastically 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  $w$  *stochastically 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  $w$  *stochastically 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 stochastically stable if and only if a subset in it is stochastically stable. For every set of mixed quasi 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  $BI = (s^i)_{1 \leq i \leq k}$ , and denote the  $\epsilon$ -neighbourhood of  $BI$  by

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

We shall first prove in Section 4 that our best-reply dynamic leads to a Nash equilibrium, and that only Nash equilibria are stochastically stable regardless of population size.

The combined results from Hart [12] and Gorodeisky [10] 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, 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. So  $NC = NE \setminus BC$ . We give a proof of the following theorem by game  $\Gamma_1$  in Section 6.

**Theorem 3.2** *Under some best-reply dynamic or better-reply dynamic, there exists a two-player game of perfect information and a positive number  $c$  such that for all finite  $m > 1$   $NC$  is  $m$ -stochastically stable and*

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

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

## 4 Stochastic Stability of Nash Equilibria

In this section, we prove that, under the best-reply or better-reply selection mechanism, only Nash equilibria are stochastically stable regardless of population size, after we introduce some notations and concepts. To this end, we need to study the relationship between backward induction equilibrium and general Nash equilibrium in subgames, and we need the two operations introduced below.

### 4.1 Definitions, notations and operations

Recall that each individual plays a pure quasi strategy in each period. If at some decision node  $n$ , strategy  $b$  of an individual  $q$  requires  $b(n) = n'$ , we then say that  $q$  moves toward  $n'$  at node  $n$  (or towards a node in  $\Psi^{-1}(n)$ ).

Consider a generic finite game  $G$  in extensive form of perfect information with the associated population game. Recall that a state describes the distribution of all individual's strategies in  $G$ . Given a state  $w$ , a node  $n$  is connected in  $G$  at  $w$  if and only if at every  $n'$  in  $\Psi(n)$ , i.e., on the subplay from the root of  $G$  to the immediate predecessor of  $n$ , there is at least one individual  $q$  in  $M(\lambda(n'))$  whose move is towards  $n$ . (Note  $M(\lambda(n'))$  is the corresponding population of the player who plays at node  $n'$ .) When a node  $n$  in  $G$  is connected at state  $w$ , we write  $R_w(n) = 1$ ; otherwise  $R_w(n) = 0$ .

Given a node  $n$  in  $G$ , we say an individual in  $M(\lambda(n))$  is *active* at node  $n$ , if she moves towards the node  $n$  at every predecessor node of  $n$  where that individual takes a move, i.e., at all nodes in  $\Psi(n) \cap \Lambda^{\lambda(n)}$ . If for a node  $n$  in  $G$ , player  $\lambda(n)$  does not play at any node in  $\Psi(n)$ , then every individual in  $M(\lambda(n))$  is active at  $n$ . For any individual, the set of all moves taken at the nodes where the individual is active essentially corresponds to a pure quasi strategy. For convenience, we recognise a move as a directed edge in the game tree. When we need to emphasize any notation defined with respect to the game  $G$ , we add  $(G)$  on it, e.g.,  $N(G)$ ,  $N_t(G)$ ,  $\lambda(n)(G)$ , etc.

We introduce below two operations on a node  $n$  in the set

$$L(G) := \psi^{-1}(N_t(G)) = \{n \in N(G) : \psi^{-1}(n) \subseteq N_t(G)\}$$

in a generic extensive-form game  $G$  of perfect information.

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

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

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

We delete the edge from  $\psi(n)$  to  $n$ , and check whether  $\psi(n)$  is a new terminal node after deletion. If yes, then we also delete the edge from  $\psi^2(n)$  to  $\psi(n)$ , and check whether  $\psi^2(n)$  is a new terminal node after deletion. This process continues until we find a predecessor  $\bar{n} \in \Psi(n)$  with survival terminal nodes after deletion. We denote the such modified game by  $G^{-n}$ .

## 4.2 Constraints on the histories

Given a state  $w$  in  $W$ , if  $x(w)$  is in  $NE$ , we simply write  $w \in NE$ . (See (2.5) for the definition of  $x(w)$ .)

Given a finite generic extensive-form game  $\Gamma$  of perfect information and an associated Markov chain defined in Section 2.2, 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 crucial step to show that only  $NE$  can be stochastically stable in the evolutionary process. We shall construct such a finite history under the following constraints.



1. **No-mutation Constraint:** In each period, the chosen individual in each population either picks her best-reply strategy or keep using her 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  $1 - \mu - \sigma$  in each period.
2. **Disconnected-moves-unchanged Constraint:** When a chosen individual switches to her best-reply strategy in period  $t$ , for all nodes  $n$  with  $R_{w_t}(n) = 0$  and the property that the individual is active at  $n$  in both periods  $t$  and  $t + 1$ , her move at  $n$  will be the same in these two periods. In other words, when an individual’s current strategy consists of a move at a disconnected node  $n$ , her move at  $n$  remains unchanged if she is still active at  $n$  in the next period.

For instance, in the game in Figure 6, suppose that in period  $t$  no individual in  $M(1)$  is playing  $s_1$ , and thus the middle left node  $n_{lf}$  is disconnected. Note that all individuals in  $M(2)$  are active at node  $n_{lf}$  in any period. So the move of any individual in  $M(2)$  at node  $n_{lf}$  keeps unchanged in period  $t + 1$ .

3. **Fixed-move Constraint:** It is concerning about a node  $n$  in  $\Gamma$  and an edge  $e$  departing from  $n$ , and it takes effect from a certain period  $t$ : from  $t$  on, whenever node  $n$  is disconnected and an individual turns active at  $n$ , her move at  $n$  is always the edge  $e$ . This is possible, since, when an individual changes to her best-reply strategy, her moves on disconnected nodes do not affect her payoff.

We still take the game in Figure 6 as an example. We can apply this constrain from some period  $t$  that, whenever an individual in  $M(1)$  takes a move at the disconnected bottom node, she picks the right edge, which corresponds to the strategy  $s_3$ .

### 4.3 Construction of a finite history ending at $NE$

Denote  $\Gamma$  as the extensive-form game we study. Given any initial state  $w_0$ , we construct a finite history  $\alpha = (w_0, w_1, \dots, w_t)$  in the stochastic process under the best-reply selection mechanism (without mutation) such that  $w_t \in NE$ .

To this end, we shall define below a sequence of extensive-form games  $(G_{n_1}, \dots, G_{n_r})$  indexed by  $n_i \in N(\Gamma)$  for all  $i$  with  $1 \leq i \leq r$ . We require  $N(G_{n_l}) \supset N(G_{n_{l+1}})$  for all  $l$  with  $1 \leq l < r$ . We also let  $G_{n_1} = \Gamma$  and  $G_{n_r}$  be a one-node game. Furthermore, for  $i \geq 1$ , game  $G_{n_{i+1}}$  is obtained from pruning or cutting at a node  $n_{i+1}$  in  $L(G_{n_i})$  in game  $G_{n_i}$ . Each  $G_{n_i}$

will be associated with a finite sequence of states  $(w_i^0, w_i^1, \dots, w_i^{\beta(i)})$  under the best-reply selection mechanism. (These states are defined with respect to  $\Gamma$ .) In doing so, we could fix each specific move in turn in a Nash equilibrium, and, for each connected node in the equilibrium, all individuals active at that node could follow that move from some period till the end of history  $\alpha$ .

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

**Case II:** There exists one such finite process. We adopt one such process  $y$  and denote it as  $(w_i^0, w_i^1, \dots, w_i^l)$ . We denote the equivalent node to the terminal node  $n_{i+1}^{\max}$  in  $G_{n_i}$  by  $n_{i+1}^e$  in  $\Gamma$ . (\*) (Recall that  $n_{i+1}^e$  gives player  $\lambda(n_{i+1})(G_{n_i})$  the maximum payoff in  $G_{n_i}$  conditional on reaching node  $n_{i+1}$ . Note that  $n_{i+1}^e$  is a terminal node in  $\Gamma$ , or there exists a game  $G_l$  such that  $l < i$  and  $G_l$  is pruned at node  $n_{i+1}^e$ .) Suppose at state  $w_i^l$  there are  $q$  active individuals at node  $n_{i+1}$  whose move is not towards  $n_{i+1}^e$ . In the next  $q$  periods, when applying the best-reply selection mechanism, we always pick one such individual in population  $M(\lambda(n_{i+1}))$  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  $q$  periods: when there is only one individual moving towards  $n_{i+1}$  at some node  $n' \in (\Psi(n_{i+1}) \setminus \Lambda\lambda(n_{i+1}))$  in the subplay from the root of  $\Gamma$  to  $n_{i+1}$ , we always avoid picking this individual in population  $M(\lambda(n'))$ . 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  $q$  periods.

Denote the period 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  $n_{i+1}^e$  since period  $t^{n_{i+1}}$ . We prune  $G_{n_i}$  at node  $n_{i+1}$  and obtain  $G_{n_{i+1}} := G_{n_i}^{\sim n_{i+1}}$ , 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  $L(G_{n_{i+1}})$  and check the existence of a qualified finite transition process as for game  $G_{n_i}$ . This process continues until we obtain a one-node game  $G_{n_r}$  and the associated finite transition process  $(w_r^0, \dots, w_r^{\beta(r)})$ , where  $G_{n_r}$  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  $\alpha$ . Recall that in the construction process, the sequence of generated extensive-form games  $(G_{n_1}, \dots, G_{n_r})$  has the property  $N(G_{n_l}) \supset N(G_{n_{l+1}})$  for all  $l$  with  $1 \leq l < r$ . Recall also that the associated sequence  $(w_i^0, w_i^1, \dots, w_i^{\beta(i)})$  to each game  $G_{n_i}$  is finite. Because the population size  $m$  and  $|N(\Gamma)|$  are both finite, we may infer that  $\alpha$  is finite. The transition history  $\alpha$  is in the form

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

where, for any  $i > 0$  and any  $j$  with  $l_i < j < l_{i+1}$ ,  $w_j^1$  is not defined in the above history. It is because that we cut  $G_{n_j}$  at node  $n_{j+1}$ , or we prune  $G_{n_j}$  at node  $n_{j+1}$  but the associated sequence of states  $(w_j^t)_{0 < t \leq \beta(j)}$  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  $\alpha$ . Thus, a node  $n$  is inactive if there is either a predecessor node of  $n$  or  $n$  itself cut in the construction of  $\alpha$ .

For the constructed finite history  $\alpha = (w_i)_i$ , 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  $\alpha$  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  $\alpha$  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  $\bar{n}$  in  $\psi^{-1}(n_i)$  from some period on. Therefore, from that period on, whenever  $n_i$  is connected and the best-reply strategy of an individual in  $M(\lambda(n_i))$  consists of a move at node  $n_i$ , the move must be towards that fixed node  $\bar{n}$ .*

When we talk about an active individual  $q$  at an active node  $n$  in the proof below, we mean that node  $n$  is an active node and the individual  $q$  is active at node  $n$ .

**Proof.** Recall the definition of  $n_i^e$  and  $t^{n_i}$  in Case II in the above construction process when the active node  $n_i$  is considered in the game  $G_{n_{i-1}}$  (see (\*) and (\*\*), respectively). We show below by induction that every active individual at an active node  $n_i$  takes a move towards  $n_i^e$  from period  $t^{n_i}$  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  $\alpha$  and  $\psi^{-1}(n) \neq \emptyset$ . Given a node  $n$  in  $\Gamma$ , denote the subgame rooted at  $n$  by  $\Gamma_n$ . The **induction hypothesis** is that for all active nodes  $n_r$  in  $\Gamma_{n_i}$  after its root, every active individual at node  $n_r$  takes a move towards  $n_r^e$  from period  $t^{n_r}$  on. We first prove the following claim.

**Claim:** In period  $t^{n_i}$  all active individuals at the active node  $n_i$  are taking the move towards  $n_i^e$  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 period when it reaches the state  $w_{n_{i-1}}^l$  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  $\Gamma_{n_i}$  except the root takes a move towards  $n_j^e$  in any period  $t \geq \bar{t}$ . Note that if a node  $n_r$  is inactive, then after any period  $t^{n_j}$  with  $j > r$ , node  $n_r$  remains disconnected. Recall that the specified move sequence starting from  $n_j^e$  leads to the backward-induction payoff vector in  $\Gamma_{n_j}$  if all inactive nodes are excluded. It follows that, in all periods  $t \geq \bar{t}$  when  $n_i$  is connected, the local best-reply strategy in  $\Gamma_{n_i}$  of an individual in  $M(\lambda(n_i))$  consists the move from  $n_i$  to  $n_i^e$ . Consider those  $q$  active individuals at node  $n_i$  whose move is different. We know from Case II in the construction of  $\alpha$  that, when anyone of those  $q$  active individuals is chosen after period  $\bar{t}$ ,  $n_i$  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 the move towards  $n_i^e$ . That completes the proof of the above claim.

With the same argument above, in any period  $t > t^{n_i}$ , when  $n_i$  is connected, an individual in  $M(\lambda(n_i))$  has a best-reply strategy  $s$  with the move towards  $n_i^e$ , 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 to another edge departing from  $n_i$ . When an individual's best reply consists of a move at  $n_i$  but with  $n_i$  disconnected, Constraint 3 requires that move to be towards  $n_i^e$ . We have completed the proof of the lemma.  $\square$

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

**Proof.** Consider the last active node, the root of  $\Gamma$  in the construction of  $\alpha$ . At state  $w_t = w_r^{\beta(r)}$ ,  $G_{n_r}$  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  $w_t$  in  $\alpha$ . So  $w_t \in NE$ .  $\square$

**Comment:** To deduce the result that the backward induction equilibrium is always stochastically stable, Hart proved in (3.4) in [12] 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.

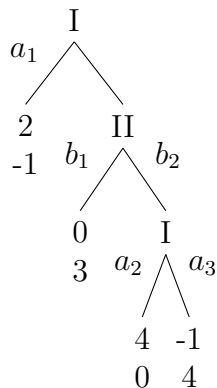


Figure 2: A game with non-acyclic  $Q_0$

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.

**Lemma 4.3** *In a finite  $k$ -player game in extensive form of perfect information, we suppose that each player has no more than  $\zeta$  pure quasi strategies, i.e.,  $\zeta \geq \max\{|B^1|, \dots, |B^k|\}$ . Under the selection mechanism of best reply or better reply, if from any initial state the process can enter NE in finite periods with probability  $p$ , then it can also enter NE in no more than  $\zeta^{mk}$  periods with probability no less than  $p$ , where  $m$  is the population size.*

**Proof.** We firstly observe the size of the state space  $W$  is  $(\zeta^m)^k$ . For any partial transition history  $(w_0, w_1, \dots, w_t)$  with positive probability  $p$ , if there exist two periods  $i$  and  $j$  such that  $i < j$  and  $w_i = w_j$ , then the shortened partial history  $(w_0, \dots, w_{i-1}, w_i, w_{j+1}, \dots, w_t)$  happens 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  $i \neq j$ . Recall the size of the state space  $W$ .  $\square$

We have the following corollary from this lemma and the Markov property.

**Corollary 4.4** *The best-reply dynamic defined in Section 2.2 (without mutation) converges to a Nash equilibrium almost surely in a finite extensive-form game of perfect information.*

We can further prove that the adapted play with finite memory and sample size in [29] converges to a Nash equilibrium. In the context of random matching in [23], the convergence result above still holds. Since a best reply dynamic is a kind of uncoupled dynamic, it is an open question in how many classes of games and under what selection mechanism the convergence result holds. Recall the negative result in [13].

**Theorem 4.5** *Under the best-reply selection mechanism, for a generic finite extensive-form game of perfect information, we have the following property:*

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

for all  $m$  in  $\mathbb{N}$ .

**Proof.** We suppose that each player has no more than  $\zeta$  pure quasi strategies. From Lemma 4.2 and Lemma 4.3, it follows that under the best-reply selection mechanism as well as the Constraint 1 in Section 4.2, a state  $w$  not in  $NE$  can enter  $NE$  in no more than  $(\zeta^m)^k$  periods 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  $t \leq \zeta^{mk}$ , 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 in Section 2.2 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 in each period the chosen individual in each population either moves to pick her best-reply strategy or keeps using her best-reply strategy against the current distribution of other populations. So in each period, the transition probability to the next state where  $l$  populations change due to selection and the other  $k - l$  populations unchanged is no less than

$$m^{-k} (\kappa \bar{\sigma})^l (1 - \mu)^{k-l}$$

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

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

and

$$\frac{t}{c(w, w', t)} \leq \frac{\zeta^{mk}}{\left( 2 \left( \frac{\kappa \bar{\sigma}}{2m} \right)^k \right)^{\zeta^{mk}}} =: f_{\bar{\sigma}}(m).$$

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

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

It follows that

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

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

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

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 2.2. To see this, note that, if some player 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 [12]. One may attempt to deduce the conclusion from Theorem 3.1 in [30]. 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.

Consider a best-reply selection mechanism that, whenever the chosen individual is not playing the best strategy, she selects each one current best strategy with equal probability, i.e., (2.9) is replaced by

$$p^i(\tilde{w}^i(q^i) = b^i | w^i(q^i) \notin BS^i(w)) = 1/|BS^i(w)| \quad (4.2)$$

for each  $b^i \in BS^i(w)$ . In this case,

$$f_{\bar{\sigma}}(m) = \frac{\zeta^{mk}}{\left(2 \left(\frac{\bar{\sigma}}{2m\zeta}\right)^k\right)^{\zeta^{mk}}} \quad (4.3)$$

in the proof above. We will use this constant in the later proofs. The analogous better-reply selection mechanism has (2.7) replaced by

$$p^i(\tilde{w}^i(q^i) = b^i | w^i(q^i) \notin BT^i) = 1/|BT^i| \quad (4.4)$$

for each  $b^i \in BT^i$ .

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

We study the problem of stochastic stability by the approach below.

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

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

The Markov chain is ergodic, and property (5.1) holds regardless of the initial state.

To obtain the invariant distribution between  $BC$  and  $NC$ , we analyse the long-term relative proportion of visiting time in  $NC$ .

From Theorem 4.5, 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 10.2 for more details. Recall that we consider a process in which the mutation



rate decreases to zero. When the mutation rate is much smaller than the inverse of the bound above, the expected interval between two sequential mutations 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 stochastically 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 stochastically 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  $C_2$  for no less than  $1/(k\mu)$  periods before the next mutation, where  $k$  is the number of players. Since  $C_1$  is stochastically stable, the relative frequency of visits at  $C_1$  is positive. If the process has stayed at  $C_1$  for  $l/(k\mu)$  periods, then in expectation it has visited  $C_2$  for no less than  $lp/(k\mu)$  periods. (We ignore the period when the state is not in  $NE$ .) From (5.1), we can conclude that  $C_2$  is also stochastically stable. Hence, to show  $NC$  being stochastically 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 notation  $U(BC, NC)$ , which is the maximum expected transition time from a state in  $BC$  to  $NC$ . The rigorous definition of  $U(BC, NC)$  is introduced in Section 10.2.

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  $NC$  in some period  $t > 0$ . Given a state  $s$ , we denote the event  $F$  with  $Z_0 = s$  by  $F_s$ . Note that we can in fact start from any period  $\bar{t} > 0$ , 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 period  $\bar{t}$ ,  $F$  includes that the first mutant in population  $M(1)$  picks strategy  $a_2$  and the second mutation happens  $2\mu^{-1}$  periods 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 need to realise  $F$  is true (or false), respectively. That is,

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

and

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

Note that  $g_{\bar{\sigma}}(m)$  in the lemma below is defined in (10.1). To understand and use this lemma, we can view the condition  $\mu < g_{\bar{\sigma}}(m)$  as  $\mu$  being enough small.

**Lemma 5.1** *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  $NC$  in some period  $t > 0$ . We further suppose that there exist two positive numbers  $l$  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

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

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

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

The proof is in Section 10.3.

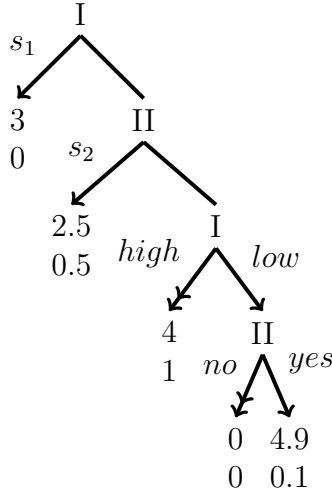


Figure 3: A centipede-trust game  $\Gamma_1$

## 6 A centipede-trust game

We call game  $\Gamma_1$  in Figure 3 a centipede-trust game, since it can be viewed as a combination of a trust game (see [1]) and a centipede game (see [24]).

The backward induction equilibrium in the game  $\Gamma_1$  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.

We say an extensive-form game of perfect information is a *finite stopping game* if, for each non-terminal node, the number of its immediate successor nodes not in  $N_t$  is no more than 1, i.e.,

$$|\psi^{-1}(n) \setminus N_t| \leq 1, \forall n \in N \setminus N_t.$$

By the finite stopping game  $\Gamma_1$  in Figure 3, we prove Theorem 3.2 under a best-reply selection mechanism with the condition (4.2). Note that the plays 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 a finite stopping game in [28].) 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 [8] for analysis of replicator-dynamic mechanism and [20] for experimental results on ultimatum games.) In  $\Gamma_1$ , player II has limited bargaining power in the whole game. In period 1, player I has a pie of size 3. She can either stop the game and eat the whole pie by strategy  $s_1$ , or pass the pie to player II. If she chooses to continue the game, player II can then secure 1/6 of the pie to himself by stopping the game, or pass the pie back to player I. If no one decides to stop the game by that time, the pie then grows to size 5, and player I is going to make a final offer to player II in period 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 in period 3 is to demand almost the whole pie to himself. If so, player II can decide whether to accept or reject this tiny part of the pie. If she rejects it, then both players leave with nothing.

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

From another perspective, we may transform  $\Gamma_1$  into an agent normal-form game  $\Gamma_2$ , where each player can only move at one node. In  $\Gamma_2$  below, player I and player III has identical payoffs at all terminal nodes, and the same payoffs applies for player II and player IV. The difference between  $\Gamma_2$  and  $\Gamma_1$  is 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 her (quasi) strategy. There are always  $m$  individuals playing at the third node in the associated population game of  $\Gamma_2$ , while, in the population game of  $\Gamma_1$ , if all individuals in  $M(1)$  are playing quasi strategy  $s_1$ , then none is playing at the third 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

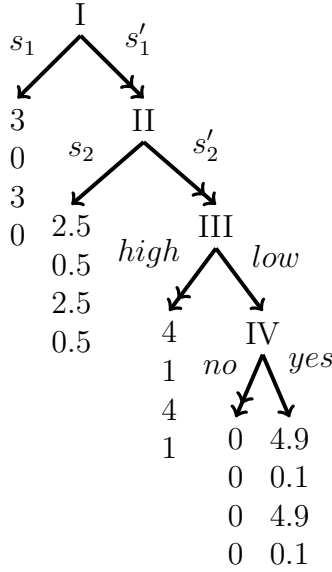


Figure 4:  $\Gamma_2$ : transformation of  $\Gamma_1$  to an agent normal-form game

$e_n := (s'_1, s'_2, high, no)$  is included in  $NC$ . One can show that the backward induction equilibrium is the only stochastically stable state when populations increase to infinity. To see this claim, from Theorem 4.5, 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 no better payoff than the one of her backward induction strategy. The only way to transit from  $BC$  to  $NC$  is by consecutive mutations ( $dm$  times for some constant  $d$ ), 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  $dm$  (e.g.  $d=5$ ) periods and hence it stays unchanged there during these periods. Thus the probability is also positive that a significant proportion (greater than  $40/49$ ) of  $M(4)$  has moved to  $yes$  in these  $dm$  periods. That will further put  $M(3)$  under the selection pressure to  $low$ , and will then change the best-reply strategy for  $M(2)$  and  $M(1)$  in order. Finally all populations converge to their respective backward induction equilibrium strategy. That is, the evolutionary dynamic leads to  $e_b$ .

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*, and 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 time, and that has to be *high* if the distribution 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 stochastically 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 three economics motivations 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 of reciprocal interactions. (Recall that Hart in [12] shows only backward induction equilibria are stochastically stable for large populations in a game where each player can only move once.) For game  $\Gamma_1$ , if, in period 1, player I resists the temptation to eat the whole pie immediately and chooses to continue the game with a promise of payoff 1 in period 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 in period 1 can be viewed as a 'carrot' to attract player II to be more patient. If player II tends to stop the game in period 2, then player I gradually learns it in the long run and will finally simply stop the game in period 1, i.e., the result of backward induction equilibrium with 0 payoff to player II. If player II indeed continues the game in period 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 her promise. If player I is greedy in this period and try to abuse her first-mover advantage, then it is reasonable to expect that player II might reject the offer to punish player I for cheating. She may simply do so to show her dignity or indifference between nothing and crumbs of the pie. The threat of saying no in the last period 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.

From another point of view, this non-backward induction equilibrium is

fairer than the backward induction equilibrium in game  $\Gamma_1$ . Here we adopt the characterisation of fairness in [4]: in response to an act of player I that is favourable for player II, II is willing to take costly actions to return at least part of the favour (positive reciprocity), and in response to an act that is perceived as harmful by II, II is willing to take costly actions to reduce I's material payoff (negative reciprocity). The trust scheme and fairness might 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 stochastically 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 vector  $(4, 1)$ . We know that the classical paper [14] 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, regardless of population size.

For the third point, it is interesting to see that once we include the outside options  $s_1$  and  $s_2$  for both players, the extended mini ultimatum game may lead to an outcome as predicted. Indeed, in real life, people are usually not forced to participate a certain programme or investment. They may suffer opportunity cost if they commit to the game. The set of payoff vectors in the mini ultimatum game is not the complete payoff set in the centipede-trust game with outside options. The result of the backward induction equilibrium in the original mini-ultimatum game is no longer a Nash equilibrium in the extended version. See [20] for the experimental results of mini-ultimatum games where the backward induction equilibrium is often not selected under the circumstance that the offer to the responder is small.

### 6.3 The proposition and an easy extension

We prove Theorem 3.2 by Proposition 6.1 below, whose proof is included in Section 10.7.

**Proposition 6.1** *Given the game  $\Gamma_1$  in Figure 3, for the best-reply selection with the condition (4.2) or the better-reply selection mechanism with the condition (4.4), there exists a positive number  $c$  such that for all finite  $m > 1$ ,*

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

**Comment 1:** The stability result of  $NC$  depends on the exact condition (2.7) or (2.9) in the selection mechanism.

**Comment 2:** In [12], Hart also shows an example that equilibria other than the backward induction equilibrium may be stochastically 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.

It is not difficult to find a four-node game with a stochastically 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.

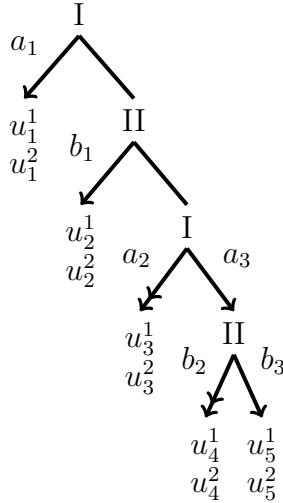


Figure 5: A general four-node game with  $NC$  stochastically stable

By Lemma 10.7 and the analogous arguments as in the proof of Proposition 6.1, one can find the alternative Nash equilibrium component with payoff vector  $(u_3^1, u_3^2)$  stochastically 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 stochastically stable

We apply the best-reply selection mechanism with the condition (4.2) in this section, and show a sufficient condition for that  $NC$  is stochastically stable in  $\Gamma$  for any population size. This result is later 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  $\Gamma$ , recall that the subgame of  $\Gamma$  rooted at  $n$  is denoted as  $\Gamma_n$ . Given a Nash equilibrium component  $EC \subseteq NE$ , we denote the play of  $EC$  by  $p_{EC}$ , which is from the root of  $\Gamma$  to the terminal node of  $EC$ . Given a node  $n$  in  $\Gamma$ , we denote the subplay from the root to  $n$  by  $p(n)$ . Recall that the *depth* of a node is the length of the subplay 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  $p_{BC}$  and  $p_{EC}$  by  $\tilde{n}_{BC,EC}$ . 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  $\tilde{n}_{BC,EC}$  as player I. (Thus the other player is denoted by player II.) Given a play  $h$  which terminates at node  $n$ , we sometimes denote  $\mathbf{v}_n$  to be the payoff vector  $\mathbf{v}(h)$ , since  $h$  uniquely determines the terminal node  $n$ . For a Nash equilibrium component  $EC$ , we denote the payoff vector of the equilibrium path  $p_{EC}$  by  $\mathbf{v}_{EC} = (v_{EC}^1, v_{EC}^2)$ . Denote the terminal node of  $EC$  by  $n_{EC}$ . Suppose that there are  $\lambda$  terminal nodes in the generic game  $\Gamma$ . We enumerate the attached payoffs to each player in order, respectively, as  $(v_{(1)}^1, \dots, v_{(\lambda)}^1)$  and  $(v_{(1)}^2, \dots, v_{(\lambda)}^2)$  such that the superscripts correspond to the players and  $v_{(m)}^i > v_{(n)}^i$  for all  $m < n$ ,  $i \in \{1, 2\}$ .

In our model, player I and player II has  $\varsigma_1 := |B^1|$  and  $\varsigma_2 := |B^2|$  quasi strategies in  $\Gamma$ , respectively.

**Condition 1:** suppose that a component  $SC$  has the property  $v_{SC}^2 = v_{(1)}^2$  and

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

We denote the terminal node with the highest payoff  $v_{(1)}^1$  to player I by  $n_{(1)}^1$ , and assume  $n_{SC} \neq n_{(1)}^1$ . Denote the deepest node on both subplay  $p(n_{SC})$  and  $p(n_{(1)}^1)$  by  $n_{(1)}^1(SC)$ .

**Condition 2:** All nodes in subgame  $\Gamma_{n_{(1)}^1(SC)}$  are not on  $p_{BC}$ .

**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  $m$ -stochastically stable and*

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

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 subplay from the root to  $n_{(1)}^1$ , then  $NC$  is stochastically 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 relax the constraint 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  $n_{SC}$  is  $(v_{SC}^1, v_{SC}^2)$  with  $v_{SC}^2 = v_{(1)}^2$  and

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

## 8 Extra value to reach stochastic stability

As in the previous section, we apply the best-reply selection mechanism with the condition (4.2). We also apply in the dynamic process the Disconnected-moves-unchanged Constraint, which is introduced in Section 4.2.

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 stochastically stable result for large populations, i.e.

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

Note that stochastic stability can be viewed as a Nash refinement. A natural question is that, if we increase the payoff vector of a non-backward-induction equilibrium component, how much extra value do we need to make its equilibrium path a stochastically 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 stochastically 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 equilibrium path a stochastically stable result for large populations?

**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  $I$  by  $(BC, SC)$  as in Section 7. We say that component  $SC$  is dominated by  $BC$  on player  $I$  if and only if  $v_{BC}^1 = v_{(1)}^1$ .*

We first present a negative result by the following game  $\Gamma_3$  in Figure 6. It shows that, if one non-backward-induction equilibrium component  $SC$  is dominated by  $BC$  on player  $I$ , then it is possibly not enough to make it stochastically stable for large populations, by simply raising one payoff in the payoff vector of  $SC$  to be the global maximum and the other payoff in it to be the second best.

In  $\Gamma_3$ , the backward induction equilibrium is  $(s_1, (a_1, a_4))$ . (Note that a quasi strategy of player  $II$  in  $\Gamma_3$  needs two moves.) The other equilibrium component, which we denote by  $SC$ , contains the pure Nash equilibrium  $(s_2, (a_2, a_3))$ . 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  $\bar{\mathbf{v}}_{SC}$  to be the increased vector of  $\mathbf{v}_{SC}$ . For convenience, we put  $\bar{\mathbf{v}}_n = \mathbf{v}_n$  for all  $n \neq n_{SC}$  and then denote the modified payoff vector sequence by  $(\bar{\mathbf{v}})$ . We also denote the such modified game by  $\bar{\Gamma}_3$ .

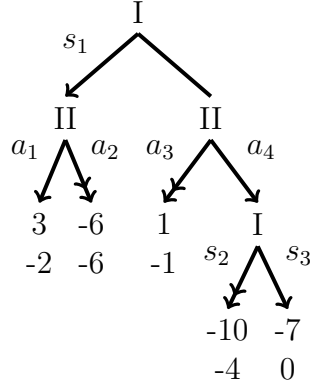


Figure 6: Game  $\Gamma_3$

One can see that only  $BC$  is stochastically stable for large populations in  $\Gamma_3$  by an adapted proof in [10] or [14]. We can further prove that to make  $SC$  stochastically stable with added values on  $v_{SC}$ , it needs  $\bar{v}_{SC}^1 = \bar{v}_{(1)}^1$  and  $\bar{v}_{SC}^2 = \bar{v}_{(1)}^2$ .

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

(1)  $\bar{v}_{SC}^1 = \bar{v}_{(1)}^1$  and  $\bar{v}_{SC}^2 = \bar{v}_{(2)}^2$ . The path of  $SC$  in  $\bar{\Gamma}_3$  is the same as in  $\Gamma_3$ . This case is similar to the game shown in the Section ‘Further comments’ in [28].

(2)  $\bar{v}_{SC}^1 = \bar{v}_{(2)}^1$  and  $\bar{v}_{SC}^2 = \bar{v}_{(1)}^2$ . The path of  $SC$  in  $\bar{\Gamma}_3$  is the same as in  $\Gamma_3$  as well. Once the state is in  $BC$ , player I has no incentive to deviate from its path  $p_{BC}$  in the dynamic process.]

Under the circumstance that, in a general game  $\Gamma$ , 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 stochastically stable result in  $\Gamma$  for large populations. If  $SC$  is not dominated by  $BC$  on player I in  $\Gamma$ , then there exists a positive  $\epsilon(\Gamma)$  with the following proposition. If we raise  $\mathbf{v}_{SC}$  to  $\bar{\mathbf{v}}_{SC}$  such that  $\bar{v}_{SC}^2 = \bar{v}_{(1)}^2$  and  $\bar{v}_{SC}^1 > \bar{v}_{(1)}^1 - \epsilon(\Gamma)$ , then the equilibrium component in  $\bar{\Gamma}$  with path  $p_{SC}$  becomes stochastically stable for any population size.*

We relegate the proof in Appendix.

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

## 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  $\Gamma_1$ , during the periods 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 situations in real life. (However, for some situations in epidemiology and epizology, one ‘bad’ mutation can indeed finally influence all related populations.) So one might try to smooth the selection process and make it continuous. For instance, Hart suggests to use the payoff difference to a positive power less than 1. (See (5.1) in [12].)

We only consider the case that  $\mu$  is extremely small with respect to  $m$  in the limiting process. (Recall Lemma 10.4.) 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  $c > 0$ . A more extreme 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 stochastically stable outcome. An interesting result is that the stochastic stability is affected by background limiting parameters under some tree structures. For instance, in the game  $\Gamma_1$ , 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  $(4,1)$  is not stochastically stable in this limiting process.

## 10 Appendix

### 10.1 Terminology of reduced normal form

There are two other well-known normal form representations of an extensive-form game. (cf. [21].) A normal-form game is *semi-reduced* (so called *pure-strategy reduced* or *quasi-reduced*) if for all  $a_1^i, a_2^i \in A^i$  and for all players  $i = 1, \dots, k$

$$\mathbf{u}(a_1^i, a^{-i}) = \mathbf{u}(a_2^i, a^{-i}), \quad \forall a^{-i} \in A^{-i} \Rightarrow a_1^i = a_2^i.$$

A normal-form game is *mixed-strategy reduced* (or simply called *reduced*) if for all  $a^i \in A^i$  and for all players  $i = 1, \dots, k$

$$\mathbf{u}(a^i, a^{-i}) = \sum_{\bar{a}^i \in A^i} \sigma^i(\bar{a}^i) \mathbf{u}(\bar{a}^i, a^{-i}), \quad \forall a^{-i} \in A^{-i} \Rightarrow \sigma^i(a^i) = 1.$$

(Recall  $\sigma^i$  in (2.2).) Given an extensive-form game  $\Gamma$ , we denote the set of strategies for player  $i$  in semi-reduced normal form and mixed-strategy reduced normal form by  $S^i$  and  $M^i$ , respectively. Recall the set of quasi strategies for player  $i$  is  $B^i$ . From their definitions, we could let  $B^i \supseteq S^i \supseteq M^i$  for any player  $i$  in any extensive-form game  $\Gamma$ .

Recall that  $\Lambda^0 = \emptyset$  implies no chance node.

**Lemma 10.1** *Given a generic extensive-form game of perfect information with  $\Lambda^0 = \emptyset$ , for each player  $i$ ,  $B^i \supseteq S^i$  implies  $B^i = S^i$ .*

**Proof.** When  $\Lambda^0 = \emptyset$ , for every player  $i$  and every two different quasi strategies  $b_1^i, b_2^i \in B^i$ , there exist two different plays  $h_1, h_2$  and one strategy combination  $b^{-i}$  of other players such that

$$\rho_{(b_1^i, b^{-i})}(h_1) = \rho_{(b_2^i, b^{-i})}(h_2) = 1,$$

i.e.,  $(b_1^i, b^{-i})$  and  $(b_2^i, b^{-i})$  generate different plays  $h_1$  and  $h_2$ , respectively, in  $\Gamma$ . (Rigorously, given a quasi strategy profile  $b = (b^1, \dots, b^k)$  and a pure strategy profile  $a = (a^1, \dots, a^k)$  such that  $a^i \in b^i$  for all  $i$  with  $1 \leq i \leq k$ , we have  $\rho_b(h) = \rho_a(h)$  for all play  $h$  in  $H$ .) Since  $\Gamma$  is generic, it follows that

$$\mathbf{u}(b_1^i, b^{-i}) \neq \mathbf{u}(b_2^i, b^{-i}).$$

It is true for every pair of quasi strategies  $b_1^i, b_2^i \in B^i$ . Recall that the set of quasi strategies is a partition of pure strategies, and we reach the conclusion  $B^i = S^i$ .  $\square$

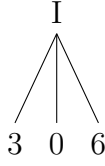


Figure 7:  $\Gamma'$  with  $B^i \neq M^i$

Even if  $\Lambda^0 = \emptyset$ , it is not true that  $B^i = M^i$  for all players  $i$  in all generic extensive-form game of perfect information. In the one-player game  $\Gamma'$  below, we denote  $B^1 = \{b_1, b_2, b_3\}$  where  $u^1(b_1) = 3$ ,  $u^1(b_2) = 0$  and  $u^1(b_3) = 6$ . However, strategy  $b_1$  can be replaced by a mixed strategy  $x^1 = \{\sigma^1(b_1) = 0, \sigma^1(b_2) = 1/2, \sigma^1(b_3) = 1/2\}$ . Thus,  $M^1 \neq B^1$ .

For the general case without the constraint on  $\Lambda$ , the conclusion in Lemma 10.1 is no longer true. In the one-player game  $\Gamma''$  below, we denote  $B^1 = \{b_1, b_2\}$ . The strategy  $b_1$  and  $b_2$  leads to the same expected payoff, and hence can be replaced by each other. So  $B^1 \neq S^1$ .

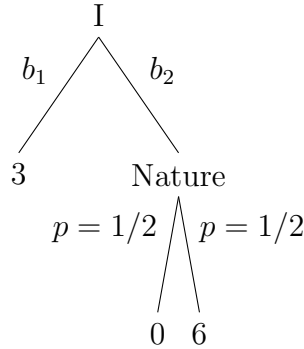


Figure 8:  $\Gamma''$  with  $B^i \neq S^i$

## 10.2 Results on transition time in Markov chain

The following corollary of Theorem 4.5 estimates the expected transition time when no mutation is 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 10.2** *Given a Markov chain  $(Z_t)_{t \geq 0}$  with  $\mu = 0$  and the associated  $f_\sigma(m)$  defined in (4.3) for a  $\sigma \geq \bar{\sigma} > 0$ , we have  $U_{0,m}(W, NE) < f_{\bar{\sigma}}(m)$ .*

**Proof.** The result follows from the definition of  $f_{\bar{\sigma}}(m)$  in (4.3).  $\square$

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

**Lemma 10.3** *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

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

$\square$

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

**Lemma 10.4** *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 (4.3), for every  $\mu < c_{\bar{\sigma}}/(32f_{\bar{\sigma}}(m))$ , it satisfies  $U_{\mu,m}(W, NE) < c_{\bar{\sigma}}\mu^{-1}$ .*



For the proof later, we define

$$g_{\bar{\sigma}}(m) := c_{\bar{\sigma}}/(32f_{\bar{\sigma}}(m)). \quad (10.1)$$

**Proof.** From Corollary 10.2,  $U_{0,m}(W, NE) < f_{\bar{\sigma}}(m)$ , where  $f_{\bar{\sigma}}(m)$  is defined in (4.3). 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 period that a mutation happens by  $t_{\mu} > 0$ . As this is a two-player game, we apply  $p = 2\mu - \mu^2$  in Lemma 10.3 and obtain  $E[t_{\mu}] > 1/(2\mu)$ . It follows from Lemma 10.3 that

$$P(t_{\mu} > 1/(2\mu) > 2f_{\bar{\sigma}}(m)) > 1/8. \quad (10.2)$$

From Corollary 10.2 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 (10.2), the process enters  $NE$  by period  $2f_{\bar{\sigma}}(m)$  is bounded as follows.

$$\begin{aligned} & P(T_{\mu,m}(w, NE) < 2f_{\bar{\sigma}}(m)) \\ & \geq P(t_{\mu} > 2f_{\bar{\sigma}}(m)) P(T_{\mu,m}(w, NE) < 2f_{\bar{\sigma}}(m) | t_{\mu} > 2f_{\bar{\sigma}}(m)) \\ & > 1/16 \quad \forall w \in W. \end{aligned} \quad (10.3)$$

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 10.5**

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

**Proof.** By Lemma A.1 in [10], for any  $w$  in  $S$ ,

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

which completes the proof.  $\square$

It is also important to know in expectation how long the Markov process 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 period 0 as  $t_0^e(w)$ . For each  $i \geq 0$ , we mark the period  $\min\{t > t_i^e(w) : Z_t = w\}$  as  $t_{i+1}^s(w)$  and then the period  $\min\{t > t_{i+1}^s(w) : Z_t \neq w\}$  as  $t_{i+1}^e(w)$  in turn. We define  $\bar{U}(w)$  to be the expected length of the time during which the dynamic process stays at  $w$  in the following way,

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

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

$$\bar{U}(w) = E[t_i^e(w) - t_i^s(w)] \quad \forall i \geq 1.$$

**Lemma 10.6** *Given a two-player extensive-form game of perfect information, for the induced Markov chain with the mutation rate  $\mu$ , we have*

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

for all  $w$  in  $NE$ .

**Proof.** From period 0, we enumerate the periods when at least one mutation happens as  $u_1, u_2, \dots$ . In every period, the probability of a mutation in any population is no more than  $2\mu$ . It follows that

$$E[u_{k+1} - u_k] \geq 1/(2\mu) \tag{10.4}$$

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.  $\square$

### 10.3 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 10.5, 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 10.4 and initial assumptions, we find  $U(BC, NC) \leq p^{-1}(l + c_{\bar{\sigma}})\mu^{-1}$ . By Lemma 10.5 again, we arrive at

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

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 10.6 then completes the proof.

### 10.4 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 10.7** *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*

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

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

$$P(kB_t \geq C_t \forall 0 < t \leq n) = P(S_t/t \geq (1-k)/(k+1) \forall 0 < t \leq n). \quad (10.5)$$

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 the 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 (10.5), we complete the proof.  $\square$

## 10.5 Preliminary lemmas for the dynamics

**Lemma 10.8** *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  $M(i) \ni q$  in some period  $t$ , then for any positive number  $k$ , the probability that  $q$  is not chosen in  $M(i)$  for all periods between  $t$  and  $t + \lceil km \rceil$  is greater than  $2^{-2k-1}$ .*

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

**Lemma 10.9** *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  $(Z_t)$  of a finite extensive-form game of perfect information, suppose that population  $M(i)$  is under selection pressure moving away quasi strategy  $b^i$  for all periods  $t > \bar{t}$ . 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 period  $t_1$  and  $t_2$ .

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

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

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

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

It follows that

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

By Markov inequality, we find

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

which completes the proof.  $\square$

## 10.6 A sufficient condition for 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 periods  $\tau_1$  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.

**Lemma 10.10** *Consider the Markov chain associated to the game  $\Gamma_1$  with the best-reply selection mechanism. Given two periods  $t_1$  and  $t_2$  with  $t_1 < t_2$ , we assume that event  $G_{t_1, t_2}$  happens, i.e., no mutation happens between  $t_1$  and  $t_2$ . We further assume that  $m_{high}(t_1) = 1$ ,  $m_{low}(t_1) = 0$ , and that population  $M(2)$  satisfies the following proposition in period  $t_1$ .*

$$m_{s_2}(t_1) \geq 2m/3 \text{ and } 2.5m_{s_2}(t_1) + 4.9m_{yes}(t_1) \leq 3m. \quad (10.6)$$

If

$$1.01\Delta_{no}(t_1, t) \geq \Delta_{yes}(t_1, t) \geq 0 \quad \forall t \text{ with } t_1 < t \leq t_2, \quad (10.7)$$

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

Note that the quasi strategy  $s_1$  is a best reply for  $M(1)$  in period  $t_1$ , if and only if  $Z_{t_1}$  satisfies (10.6).

**Proof.** We denote  $u_{low}(t)$  as the payoff to an individual in population  $M(1)$  who plays *low* in period  $t$ . By (10.6), 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* in period  $t_1$  are  $p_1 = x_{s_2}(Z_{t_1})$ ,  $p_2 = x_{no}(Z_{t_1})$  and  $p_3 = x_{yes}(Z_{t_1})$ , respectively. Then, the payoff of an individual in population  $M(1)$  playing strategy *low* in period  $t_1$  is

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

Then, given a period  $t$  with  $t_1 < t \leq t_2$ , with the possible move of population  $M(2)$  under the constraint (10.7), the payoff of an individual in population  $M(1)$  playing strategy *low* in period  $t$  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 period  $t_1$  and period  $t$ . Thus  $u_{low}(t) \leq u_{low}(t_1) \leq 3$ , and hence strategy *low* is weakly dominated by strategy  $s_1$  in period  $t$ . Recall  $m_{low}(t_1) = 0$  and the absence of mutation between period  $t_1$  and  $t$ . We may then infer that there is still be no individual playing at strategy *low* in period  $t$ .  $\square$

## 10.7 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  $NC$  in some period  $t > 0$ . 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  $F = \bigcap_{i=1}^4 F_i$  (or  $F_4 | \bigcap_{i=1}^3 F_i$ ) in a dynamic process starting from period 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 10.7, 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 the definition of  $g_{\bar{\sigma}}(m)$  in (10.1)). 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  $Z_0 \in BC$  in period 0. We shall see that the argument below is independent to the exact initial state in  $BC$ . We enumerate the periods that at least one mutation happens as  $u_1, u_2, \dots$ . Suppose that  $q^1$  is the individual chosen in  $M(1)$  in period  $u_1$ .

**Event  $F_1$**  :=  $\{Z_{u_1}^1(q^1) = \text{high}, m_b(u_1) = m_b(0) \forall b \in \{s_2, \text{no}, \text{yes}\}\}$ . [The first mutation is in population  $M(1)$  only and that yields a *high* strategy.]

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

$$P(F_1 | Z_0 = w) = (1 - \mu)/6 > 1/7. \quad (10.8)$$

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

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

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

$$P(F_2 | Z_0 = w) > 1/16 - 1/200. \quad (10.9)$$

It is straightforward to see

$$P(u_2 - u_1 > \mu^{-1}) \geq (1 - 2\mu)^{\mu^{-1}} = \left( (1 - 2\mu)^{\frac{1}{2\mu}} \right)^2 > 4^{-2} = 1/16, \quad (10.10)$$

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 period after  $u_1$  when the state is in  $NE$ , i.e.,  $t_1 = \min\{t > u_1 : Z_t \in NE\}$ .

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

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

$$P(F_3|D_2, Z_0 = w) \geq 1 - e^{-\bar{c}_{\bar{\sigma}}2^{-\bar{s}}}. \quad (10.11)$$

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 (4.3). From Corollary 10.2 and Markov inequality,

$$P(U_{0,m}(Z_{u_1}, NE) > e^{\bar{c}_{\bar{\sigma}}2^{\bar{s}}}f_{\bar{\sigma}}(m)) < e^{-\bar{c}_{\bar{\sigma}}2^{-\bar{s}}}. \quad (10.12)$$

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 period  $u_1$  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 period between the first and the second mutation.]

**Claim:**

$$P(F_3 \cap F_4|D_2, Z_0 = w) \geq 2^{-\bar{s}}(2^{-\bar{c}_{\bar{\sigma}}} - e^{-\bar{c}_{\bar{\sigma}}}) \forall w \in BC. \quad (10.13)$$

Note that, in any period  $t > u_1$ , 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 periods  $\tau_1$  and  $\tau_2$  with  $\tau_1 < \tau_2$ , recall the definition of  $\Delta_{no}(\tau_1, \tau_2)$  and  $\Delta_{yes}(\tau_1, \tau_2)$  in Section 10.6 and  $G_{\tau_1, \tau_2}$  in Lemma 10.9. We denote the event of  $1.01\Delta_{no}(\tau_1, t) \geq \Delta_{yes}(\tau_1, t)$  for all periods  $t$  with  $\tau_1 < t < \tau_2$  by  $H_{\tau_1, \tau_2}$ . We take  $m_{s_2}(u_1)$  as number  $n$  in Lemma 10.7, and find that, given any  $t > u_1$ , (for example, take  $t$  as  $\min\{t' > u_1 : m_{s_2}(t') = 0\}$ )

$$P(H_{u_1, t}|G_{u_1, t}, m_{high}(t') \geq m_{low}(t') = 0 \forall t' \text{ with } u_1 < t' < t) \geq 2^{-\bar{s}}. \quad (10.14)$$

From Lemma 10.10 and the definition of  $F_1$ , if both  $H_{u_1, t}$  and  $G_{u_1, t}$  happens, then it follows that *low* is not a best reply of  $M(1)$  up to period  $t$



and no individual in  $M(1)$  is playing *low* at  $t$ . 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  $x_{s_2}^2 \leq 1/3$ . We show below that the following event happens with positive probability conditional on  $1.01\Delta_{no}(u_1, t) \geq \Delta_{yes}(u_1, t)$  and no mutation involved for all periods  $t > u_1$ : 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 10.8 and conclude that the probability is greater than  $2^{-1-2k}$  that the mutant will keep playing strategy *high* between period  $u_1$  and  $u_1 + \lceil km \rceil$ . (\*) Recall that when the mutant is playing *high* and no individual in  $M(1)$  is playing *low*,  $M(2)$  is under selection pressure and moving away from  $s_2$ . Taking the same  $k$  and  $\lambda = 1/3$  in Lemma 10.9, we find

$$P\left(m_{s_2}(u_1 + \lceil km \rceil) \leq \frac{m}{3} \mid F_{h>l}, G_{u_1, u_1 + \lceil km \rceil}\right) \geq \frac{1}{2} \quad (10.15)$$

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 (\*), (10.14) and (10.15) 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 (10.11), we have proved the claim.

Finally, from (10.8), (10.9) and (10.13), it follows that

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

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

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

By the definition of  $F_3$ ,  $E[t_1 - u_1 \mid F, Z_0 = w] < 1/(2\mu)$ , 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  $BC$ ,

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

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. The proof is valid, as long as all best-reply quasi strategies are assigned with the same probability for each selection step.

## 10.8 Preliminary results to prove Theorem 7.1

Recall that in  $\Gamma$  the sets of all nodes where player I and II play are  $\Lambda^1$  and  $\Lambda^2$ , respectively.

**Lemma 10.11** *Condition 1 implies that player I moves at node  $n_{(1)}^1(SC)$ .*

**Proof.** Denote the immediate successor node of  $n_{(1)}^1(SC)$  on path  $p_{SC}$  by  $\hat{n} := \psi^{-1}(n_{(1)}^1(SC)) \cap p_{SC}$ . (Recall the definition of  $\psi^{-1}$  in Section 2.1.) It is straightforward to see the subplay from  $\hat{n}$  to  $n_{SC}$  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  $n_{(1)}^1(SC)$ . Thus the path of the global backward induction equilibrium is from the root to  $n_{SC}$ . This contradicts with the assumption that  $p_{SC} \neq p_{BC}$ .  $\square$

**Lemma 10.12** *Denote by  $N_{BC}^2$  the set of nodes where player II takes a move according to her backward induction quasi strategy. If  $\Gamma$  satisfies both Condition 1 and Condition 2, then  $\Gamma_{n_{(1)}^1(SC)}$  contains no node in  $N_{BC}^2$ .*

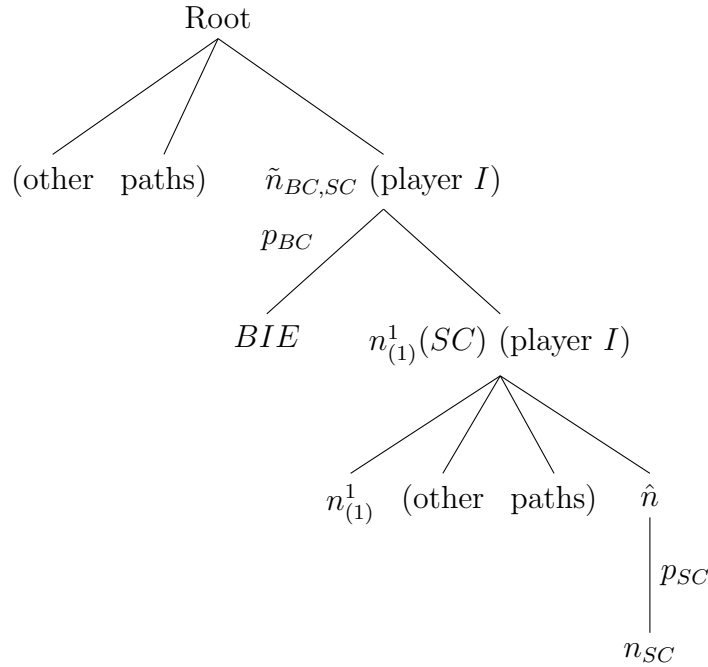


Figure 9: Tree structure of  $\Gamma$

**Proof.** We prove it by contradiction. See Figure 9 for the tree structure of  $\Gamma$ .

**Step 1:** Suppose the node  $\Gamma_{n_{(1)}^1(SC)}$  where player II moves in her backward induction quasi strategy, then player II moves towards  $n_{(1)}^1(SC)$  in her backward induction strategy at all nodes in  $p(n_{(1)}^1(SC)) \cap \Lambda^2$ .

**Step 2:** Suppose the node  $\hat{n} = \psi^{-1}(n_{(1)}^1(SC)) \cap p_{SC}$  exists. (If not, the analysis in this step can be ignored.) Because the payoff vector  $(v_{SC}^1, v_{SC}^2)$  gives the maximum payoff for each player in  $\Gamma_{\hat{n}}$ , the path of the local backward induction equilibrium starts from  $\hat{n}$  and ends at  $n_{SC}$ . Hence, at each node in  $\Psi(n_{(1)}^1(SC)) \cap p_{SC}$ , the player moves in the direction towards  $n_{SC}$  in her backward induction strategy.

**Step 3:** From Lemma 10.11, we know that player I moves at node  $n_{(1)}^1(SC)$ . If she moves towards  $\hat{n}$ , then with the backward induction path in  $\Gamma_{\hat{n}}$  taken into account, player I can secure payoff  $v_{SC}^1$ , conditional on that player II plays her backward induction strategy.

**Step 4:** Since player II moves along the whole  $p_{SC}$  in her backward induction strategy, player I's best reply against it, i.e., her backward induction strategy, should guarantee a payoff at least  $v_{SC}^1$ . However, by Condition 2,  $p_{BC}$  does not go through  $n_{(1)}^1(SC)$ , and hence player I receives a worse payoff in  $BC$ . Contradiction.  $\square$

**Lemma 10.13** *If  $\Gamma$  satisfies both Condition 1 and Condition 2, then in subgame  $\Gamma_{n_{(1)}^1(SC)}$ , the path of the local backward induction equilibrium is from  $n_{(1)}^1(SC)$  to  $n_{(1)}^1$ .*

**Proof.** We prove it by contradiction. Suppose that the path of the local backward induction equilibrium in  $\Gamma_{n_{(1)}^1(SC)}$  is not from  $n_{(1)}^1(SC)$  to  $n_{(1)}^1$ . Because the only terminal node not dominated by  $n_{SC}$  is  $n_{(1)}^1$ , the path of the local backward induction equilibrium in  $\Gamma_{n_{(1)}^1(SC)}$  must be from  $n_{(1)}^1(SC)$  to  $n_{SC}$ , and the path  $p_{BC}$  in  $\Gamma$  must be from the root to  $n_{SC}$ . This contradicts Condition 2.  $\square$

**Lemma 10.14** *Consider a quasi strategy  $b_1$  of player I which satisfies the following propositions.*

1. *At every node in  $\Psi(n_{(1)}^1(SC)) \cap \Lambda^1$ , player I moves towards the node  $n_{(1)}^1(SC)$ .*
2. *Player I plays a local backward induction strategy in  $\Gamma_{n_{(1)}^1(SC)}$ .*

*Then, the best-response quasi strategy  $b_2$  of player II against  $b_1$  includes no move in  $\Gamma_{n_{(1)}^1(SC)}$ .*

**Proof.** From Lemma 10.13, it follows that player I's induced strategy of  $b_1$  in  $\Gamma_{n_{(1)}^1(SC)}$  is consistent with the path of local backward induction equilibrium in  $\Gamma_{n_{(1)}^1(SC)}$ , which is from  $n_{(1)}^1(SC)$  to  $n_{(1)}^1$ .

We prove the statement by contradiction. Suppose that player II's best response  $b_2$  consists of some move at a node in  $\Gamma_{n_{(1)}^1(SC)}$ . It follows that the induced strategy of  $b_2$  in  $\Gamma_{n_{(1)}^1(SC)}$  is the local backward induction strategy in  $\Gamma_{n_{(1)}^1(SC)}$ . We may also infer that player II moves towards  $n_{(1)}^1$  at every node in  $p(n_{(1)}^1(SC)) \cap \Lambda^2$  in strategy  $b_2$ , since  $b_2$  includes a move in  $\Gamma_{n_{(1)}^1(SC)}$ . So the equilibrium path generated by  $(b_1, b_2)$  starts from the root and goes through node  $n_{(1)}^1(SC)$ .

Under this circumstance, we claim that  $(b_1, b_2)$  is the backward induction strategy profile in  $\Gamma$ . To see this, firstly recall that  $b_1$  and  $b_2$  are consistent with the path of backward induction equilibrium in  $\Gamma_{n_{(1)}^1(SC)}$ . Suppose a node  $n$  in  $\{n_{(1)}^1(SC)\} \cup \Psi(n_{(1)}^1(SC))$  is such that  $b_1$  and  $b_2$  are consistent with the path of the local backward induction equilibrium in subgame  $\Gamma_n$ . If the node  $\psi(n)$  is played by player I, then the move towards  $n_{(1)}^1$  gives him the highest possible payoff. So,  $b_1$  induces the backward induction strategy of player I in  $\Gamma_{\psi(n)}$ . If  $\psi(n)$  is played by player II, we know from the paragraph above that the best response in  $\Gamma_{\psi(n)}$  against  $b_1$  conditional on reaching node  $\psi(n)$  is moving towards  $n_{(1)}^1$ . In this way, we conclude that  $(b_1, b_2)$  is the backward induction strategy profile in  $\Gamma$ . Thus, path  $p_{BC}$  goes through node  $n_{(1)}^1(SC)$ , which contradicts Condition 2.  $\square$

**Lemma 10.15** *Suppose that  $\Gamma$  satisfies Condition 1 and 2. In subgame  $\Gamma_{\bar{N}}$ , there are at least two induced quasi strategies of player II. In the whole game  $\Gamma$ ,  $\varsigma_2 > 2$ .*

**Proof.** If there is only one quasi strategy in  $\Gamma_{n_{(1)}^1(SC)}$ , then  $p_{SC}$  is no longer the path of a Nash equilibrium component. Lemma 10.14 then completes the proof.  $\square$

We present below a generalised version of Lemma 10.7.

**Corollary 10.16** *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  $i \in \mathbb{N}$ . Given a natural number  $l$  with  $1 \leq l \leq r$ , we define  $N_t^l = \sum_{i=1}^t \mathbb{1}_{Y_i=l}$  for all natural numbers  $t$ . Then, given any  $k > 1$ , there exists a natural number  $s_k$  such that, for all  $n > 0$ , we have for any  $l$  with  $1 \leq l \leq r$*

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

## 10.9 Sketch proof of Theorem 7.1

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

Denote  $\bar{B}^1$  and  $\bar{B}^2$  to be the set of all quasi strategies of player I and player II which requires at least one move in subgame  $\Gamma_{n_{(1)}^1(SC)}$ , respectively.

Take a pure quasi strategy of player I in  $SC$  as  $b_3$ . Recall that the payoff vector of  $BC$  is  $(v_{BC}^1, v_{BC}^2)$ . For the equilibrium component  $BC$ , we enumerate all pure Nash equilibria in it as  $((b_1^1, b_1^2), (b_2^1, b_2^2), \dots, (b_i^1, b_i^2))$ . Denote  $B_{BC}^1 := (b_1^1, b_2^1, \dots, b_i^1)$  and  $B_{BC}^2 := (b_1^2, b_2^2, \dots, b_i^2)$ . We define  $\bar{u}$  to be the best payoff to player I when her uses the strategy  $b_3$  against a strategy in  $B_{BC}^2$  of player II, i.e.,  $\bar{u} := \max_{b \in B_{BC}^2} u^1(b_3, b)$ . Note that  $\bar{u} < v_{BC}^1 < v_{(1)}^1$ . So there exists a positive number  $\rho < 1$  with

$$(1 - \frac{\rho}{\varsigma_2 - 2})\bar{u} + \frac{\rho}{\varsigma_2 - 2}v_{(1)}^1 < v_{BC}^1. \quad (10.16)$$

(Recall Lemma 10.15 that  $\varsigma_2 > 2$ .) We fix such a  $\rho$  for the proof.

Our goal is 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 10.4 from now on. We take  $k = 1.5$  in Corollary 10.16, 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 (10.1) 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  $Z_0 \in BC$  in period 0. We shall see that the argument below is independent to the exact initial state in  $BC$ . We enumerate the periods that at least one mutation happens as  $u_1, u_2, \dots$

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

[Event  $D_6 = \bigcap_{i=1}^6 F_i$  will push the majority of  $M(2)$  to a strategy in  $B^2 \setminus \bar{B}^2$ . If the mutant at  $u_2$  happens to be an individual in  $M(1)$  playing a pure quasi strategy in  $SC$ , then the distribution of the converted  $M(2)$  on  $\bar{B}^2$  will possibly be approximately uniform all the time from  $u_2$  to  $u_3$ . Thus  $M(1)$  will favour a pure quasi strategy in  $SC$  from some time after  $u_2$ . See Comment 1 at the end of this section for more details.]

Recall the strategy  $b_1$  defined in Lemma 10.14. Suppose that  $q^1$  is the individual chosen in  $M(1)$  in period  $u_1$ .

**Event**  $F_1 := \{Z_{u_1}^1(q^1) = b_1, Z_{u_1}^2(q) = Z_0^2(q) \forall q \in M(2)\}$ . [The first mutation is in population  $M(1)$  only and that yields a  $b_1$  strategy.]

With the similar analysis as in (10.8), we conclude that, when  $\mu < 1/2$ , for all  $w$  in  $BC$ ,

$$P(F_1|Z_0 = w) = (1 - \mu)/(2\varsigma_1) > 1/(4\varsigma_1). \quad (10.17)$$

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

With the similar analysis as in (10.9), we conclude that, when  $\mu < 1/4$ , for all  $w$  in  $BC$ ,

$$P(F_2|Z_0 = w) > 1/16 - 1/200. \quad (10.18)$$

**Event**  $F_3 := \{Z_{u_2-1} \in BC\}$ . [After the transition 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 period after  $u_1$  when the mutant  $q^1$  changes her 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 period  $u_1$ , it takes less than  $1/(2\mu)$  for  $q^1$  to change her strategy.]

Conditional on  $D_3$ ,  $Z_{u_1+1/(2\mu)}$  is in  $NE$  if  $F_4$  happens. With the similar analysis as in (10.11), we conclude that, for all  $w$  in  $BC$ ,

$$P(F_4|D_3, Z_0 = w) \geq 1 - c_{\bar{\sigma}} = 1 - \exp\left(-\frac{200(\varsigma_2 - \rho)}{\rho\bar{\sigma}}\right). \quad (10.19)$$

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

With the similar analysis as in  $F_4$  in Example 1 in [28], 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  $BC$ ,

$$\begin{aligned} & P(D_5, Z_0 = w) \\ & \geq P(D_3, Z_0 = w) \left( \exp\left(-\frac{4(\varsigma_2 - \rho)}{\rho\bar{\sigma}} - 1\right) - \exp\left(-\frac{200(\varsigma_2 - \rho)}{\rho\bar{\sigma}}\right) \right). \end{aligned} \quad (10.20)$$

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

**Event**  $F_6 := \{\sum_{b \in \bar{B}^2} m_b(u_2) \leq \rho m/\varsigma_2\}$ . [At 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 an adapted analysis on  $F_5$  in Example 1 in [28], we conclude that, for all  $w$  in  $BC$ ,

$$P(D_6, Z_0 = w) \geq P(D_5, Z_0 = w)/2. \quad (10.21)$$

**Second phase:** events  $(F_7, \dots, F_{10})$  between period  $u_2$  and  $u_3$   
[Transition to  $SC$ .]

Recall that  $b_3$  is a pure quasi strategy of player I in  $SC$ . Suppose that  $\bar{q}^1$  is the individual chosen in  $M(1)$  in period  $u_1$ .

**Event**  $F_7 := \{Z_{u_2}^1(\bar{q}^1) = s_3, Z_{u_2}^2(q) = Z_{u_2-1}^2(q), \forall q \in M(2)\}$ . [The second mutation is in population  $M(1)$  only and that yields an  $b_3$  strategy.]

With the similar analysis as in (10.8), we conclude that, when  $\mu < 1/2$ , for all  $w$  in  $BC$ ,

$$P(F_7|D_6, Z_0 = w) = (1 - \mu)/(2\varsigma_1) > 1/(4\varsigma_1). \quad (10.22)$$

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

With the similar analysis as in (10.9), we conclude that, when  $\mu < 1/4$ , for all  $w$  in  $BC$ ,

$$P(F_8|D_7, Z_0 = w) > 1/16 - 1/200. \quad (10.23)$$

We denote  $t_2$  to be the first period after  $u_2$  when the state is in  $NE$ , i.e.,  $t_2 = \min\{t > u_2 : Z_t \in NE\}$ . We reset  $\bar{c}_{\bar{\sigma}} := 2 + \frac{4\rho}{\bar{\sigma}(\varsigma_2-2-\rho)}$  from now on.

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

With the similar analysis as in (10.11), we conclude that, for all  $w$  in  $BC$ ,

$$P(F_9|D_8, Z_0 = w) \geq 1 - e^{-\bar{c}_{\bar{\sigma}}\varsigma_2^{-\bar{s}}}. \quad (10.24)$$

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

**Claim:**

$$P(F_9 \cap F_{10}|D_8, Z_0 = w) \geq \varsigma_2^{-\bar{s}}(2^{-\bar{c}_{\bar{\sigma}}} - e^{-\bar{c}_{\bar{\sigma}}}) \forall w \in BC. \quad (10.25)$$

Denote  $b_4$  to be a best response of player II to  $b_3$  such that  $(b_3, b_4)$  is a pure Nash equilibrium. The mutant playing  $s_3$  attracts  $M(2)$  to divert to a strategy in  $\bar{B}^2$ .

For any two periods  $\tau_1$  and  $\tau_2$  with  $\tau_1 < \tau_2$ , denote the difference of number of individuals playing the strategy  $b_4$  between these two periods by

$$\Delta_{b_4}(\tau_1, \tau_2) := (m_{b_4}(\tau_2) - m_{b_4}(\tau_1))$$

and we further denote

$$\Delta_{\bar{B}^2}(\tau_1, \tau_2) := \sum_{b \in \bar{B}^2} (m_b(\tau_2) - m_b(\tau_1)).$$

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

$$2\Delta_{b_4}(\tau_1, t) \geq \Delta_{\bar{B}^2}(\tau_1, t) / (\varsigma_2 - 1) \quad \forall t \text{ with } \tau_1 < t < \tau_2$$

by  $H_{\tau_1, \tau_2}$ . ( $\varsigma_2 - 1$  corresponds to Lemma 10.14.) We take  $m - \sum_{b \in \bar{B}^2} m_b(u_2)$  as number  $n$  in Corollary 10.16, and find that, given any  $t > u_2$ , (for example, take  $t$  as  $\min\{t' > u_2 : \sum_{b \in \bar{B}^2} m_b(t') = m\}$ )

$$P \left( H_{u_2, t} \mid G_{u_2, t}, m_{b_3}(t') = \sum_{b \in \bar{B}^1} m_b(t') = 1 \quad \forall t' \text{ with } u_2 < t' < t \right) \geq \varsigma_2^{-\bar{s}}. \quad (10.26)$$

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 10.8 and conclude that the probability is greater than  $2^{-1-2k}$  that the mutant will keep playing strategy  $b_3$  between period  $u_2$  and  $u_2 + \lceil km \rceil$ . (\*) Taking the same  $k$  and  $\lambda = 1 - \frac{\rho}{\varsigma_2 - 2}$  in Lemma 10.9, we find

$$P \left( \sum_{b \in \bar{B}^2} m_b(u_2 + \lceil km \rceil) \geq \frac{\rho m}{\varsigma_2 - 2} \mid \hat{F}, G_{u_2, u_2 + \lceil km \rceil} \right) \geq \frac{1}{2} \quad (10.27)$$

where event

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

**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 a quasi strategy in  $B_{BC}^2$ , then by Lemma 10.12, no individual in  $M(2)$  is moving at a node in  $\Gamma_{\bar{N}}$ .

However, if this is the case, then there may be some individuals in  $M(2)$  playing a strategy in  $\bar{B}^2$ , or any other quasi strategy. It's the average payoff



to  $M(1)$  such that no individual prefers to deviate from  $B_{BC}^1$ , and hence the state stays in  $BC$ . If a mutation happens to be a pure quasi strategy in  $SC$  at 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 subplay  $p(n_{(1)}^1)$ , a pure quasi strategy of player I containing a local backward induction strategy in  $\Gamma_{n_{(1)}^1(SC)}$  might be the best response for  $M(1)$  for some time period after  $u_1$ . That would make  $M(2)$  prefer a strategy with a local backward induction strategy in  $\Gamma_{n_{(1)}^1(SC)}$  and then two populations would begin to convert to a trajectory back to  $BC$ . See the example below in Figure 10.

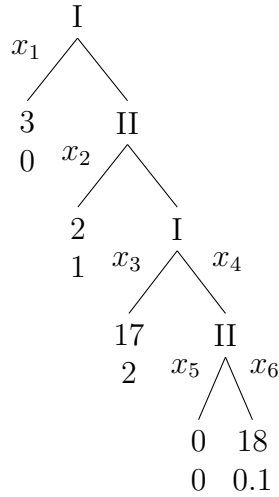


Figure 10: Example in Comment 1

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 in period 1 such that  $m_{x_3} = 1 > m_{x_4} = 0$ , and that 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 period that  $m/256$  individuals have changed their strategies from  $x_2$  to  $x_5$  or  $x_6$  with equal probability. 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, and show why event  $D_6$  above is necessary for the general two-player extensive-form games of perfect information. For a state in  $BC$ , the proportion profile of  $M(2)$  playing each strategy can be various. If one can force the majority of  $M(2)$  drifting away from  $\bar{B}^2$ , then it is with bounded probability that the process enters  $SC$  after the second mutation of an individual in  $M(1)$  playing  $s_3$ .

A mutation of an individual in  $M(1)$  to  $bs_1$  may trigger such desired drift of  $M(2)$ . From Lemma 10.14, we know that the best response of  $M(2)$  against  $b_1$  is some strategy in  $B^2 \setminus \bar{B}^2$ .

**Comment 2:** One may note that the first mutation may lead a transition to a Nash equilibrium in neither  $SC$  nor  $BC$ , i.e., to some other Nash equilibrium component. For our purpose, we only need to consider the case that it still stays in  $BC$  at period  $u_2 - 1$ , and that is shown in the Second Period.

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

We check why the transition is in a trajectory to  $SC$  if all events above, i.e.,  $D_9$ , (10.26), (\*) and (10.27), are true. Recall that in (10.21), there are at most  $\rho m / \varsigma_2$  individuals in  $M(2)$  playing a strategy in  $\bar{B}^2$  just before the second mutation. Consider the state in period  $\hat{t} := u_2 + \lceil km \rceil$ . From (10.27) at least  $\rho m / (\varsigma_2 - 2)$  individuals are playing a strategy in  $\bar{B}^2$  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  $\bar{B}^2$  between period  $u_2$  and  $\hat{t}$ . From (10.26), we may further infer that at least  $\frac{\rho m}{\varsigma_2^2 - 2\varsigma_2}$  among them are playing  $b_4$  in period  $\hat{t}$ . Denote by  $\hat{T}$  the minimum period that the number of individuals in  $M(2)$  playing a strategy in  $\bar{B}^2$  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  $\bar{B}^1$  in period  $\hat{T}$ . Let's consider its average payoff against the induced distribution of  $M(2)$

which play a strategy in  $\bar{B}^2$  in period  $\hat{T}$ . If  $q$  plays  $b_3$ , then the average payoff is  $u_{SC}^1$ . If  $q$  plays a strategy consistent with  $p(n_{(1)}^1)$ , e.g.  $b_1$ , then the average payoff is at most  $\max(\frac{v_{(3)}^1}{\varsigma_2^2} + \frac{(\varsigma_2^2-1)v_{(1)}^1}{\varsigma_2^2})$ . By Condition 1, we know that  $b_3$  gives the best payoff in all strategies in  $\bar{B}^1$  at  $\hat{T}$ .

Meanwhile, from (10.16), we know that a pure quasi strategy of player I in  $BC$  is still the (global) best reply for  $M(1)$  at least until period  $\hat{T}$ . After  $\hat{T}$ , with  $M(2)$  drifting to strategies in  $\bar{B}^2$ , pure quasi strategies of player I in  $SC$  will become the best replies for population  $M(1)$  from some period at least until  $u_3$ , if (10.26) 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  $p(n_{(1)}^1)$  is a best response for  $M(1)$  in any period between  $u_2$  and  $u_3$ .

Define the event that the state enters  $NC$  before  $u_3$  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 (10.17), (10.18), (10.20), (10.21), (10.22), (10.23) and (10.25), it follows that

$$\min_{w \in BC} P(F|Z_0 = w) > \frac{1}{(4\varsigma_2)^2} \left( \frac{1}{16} - \frac{1}{200} \right)^2 c_5(\bar{\sigma}) \varsigma_2^{-\bar{s}} \frac{2^{-\bar{c}\bar{\sigma}} - e^{-\bar{c}\bar{\sigma}}}{2}.$$

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.

## 10.10 Proof of Theorem 8.2

Recall the node  $n_{(1)}^1(SC)$  in  $\Gamma$  denoted in Condition 1 in Section 7 and that node  $\tilde{n}_{BC,SC}$  is the deepest node on both  $p_{SC}$  and  $p(n_{(1)}^1)$ . For simplicity, we abbreviate the notation  $\tilde{n}_{BC,SC}$  to  $\tilde{n}$ . We prove the theorem by considering the following two cases.

1. **Case I:**  $\tilde{n} \in \Gamma_{n_{(1)}^1(SC)}$ . We show that, when  $\bar{v}_{SC}^2 = \bar{v}_{(1)}^2$  and  $\bar{v}_{SC}^1 = \bar{v}_{(2)}^1$ ,  $p(n_{SC})$  becomes the path of the new backward induction equilibrium in  $\bar{\Gamma}$ . Firstly note that  $\bar{v}_{SC}^1 > \bar{v}_p^1$  and  $\bar{v}_{SC}^2 > \bar{u}_p^2$  for any play  $p$  which ends at a terminal node in subgame  $\bar{\Gamma}_{n_{(1)}^1(SC)}$ . So, the path of the backward induction equilibrium in  $\bar{\Gamma}_{\tilde{n}}$  leads to the payoff vector  $\bar{\mathbf{v}}_{SC}$ . If we turned the node  $\tilde{n}$  to a terminal node with payoff vector  $\mathbf{v}_{BC}$  (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  $\bar{\mathbf{v}}_{SC}$  dominates  $\mathbf{v}_{BC}$ , and thus the new backward induction equilibrium path proceeds towards  $n_{SC}$ . It is straightforward to see the play  $p(n_{SC})$  corresponds to a stochastically stable component for any population size in  $\bar{\Gamma}$ . (One mutation is enough to trigger a transition from any other equilibrium component to this new backward induction equilibrium component.)
2. **Case II:**  $\tilde{n} \notin \Gamma_{n_{(1)}^1(SC)}$ . We divide it into two sub-cases.

- (a) **Sub-case I:** the backward induction equilibrium in the subgame  $\Gamma_{n_{(1)}^1(SC)}$  terminates at  $n_{(1)}^1$ . It follows that the path of the backward induction equilibrium in  $\Gamma$  remains the same after raising  $\mathbf{v}_{SC}$ .

As before, we suppose player I and player II has  $\varsigma_1$  and  $\varsigma_2$  quasi 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.

Note that  $BC$  is the only stochastically stable component in  $\Gamma$ , so we can skip event  $F_3$  in the proof of Theorem 7.1: if  $Z_{u_2-1} \in NE$ , then  $Z_{u_2-1}$  must be in  $BC$ . (Detailed analysis needs the Disconnected-moves-unchanged constraint.)

- (b) **Sub-case II:** the backward induction equilibrium in subgame  $\Gamma_{n_{(1)}^1(SC)}$  does not terminate at  $n_{(1)}^1$ . By the similar arguments as in Case I, we find that  $p(n_{SC})$  becomes the path of the new backward induction equilibrium in  $\bar{\Gamma}$ , and the corresponding equilibrium component is stochastically stable for any population size.

## References

- [1] J. Berg, J. Dickhaut and K. McCabe, Trust, Reciprocity and Social History, 10 (1995) 122–142.
- [2] R. Cressman, Evolutionary Dynamics and Extensive Form Games, MIT Press, Cambridge, 2003.
- [3] R. Cressman and K.H. Schlag, The dynamic (in)stability of backward induction, J. Econ. Theory 83 (1998) 260–285.
- [4] A. Falk, E. Fehr and U. Fischbacher, On the Nature of Fair Behavior, Economic Inquiry, 41 (2003) 20–26.
- [5] D. Foster and H.P. Young, Stochastic evolutionary game dynamics, Theor. Pop. Biol. 38 (1990) 219–232.
- [6] M. I. Freidlin and A.D. Wentzell, Random Perturbations of Dynamical Systems, New York: Springer Verlag, 1984.
- [7] D. Fudenberg and D.K. Levine, The Theory of Learning in Games, MIT Press, 1998.
- [8] J. Gale, K. Binmore and L. Samuelson, Learning to be imperfect: The ultimatum game, Games Econ. Behav. 8 (1995) 56–90.
- [9] 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.
- [10] Z. Gorodeisky, Evolutionary stability for large populations and backward induction, Mathematics of Operations Research 31 (2006) 369–380.
- [11] S. Hart, Games in Extensive and Strategic Forms. In R.J. Aumann and S. Hart, editors, Handbook of Game Theory with Economic Applications, Elsevier, edition 1, volume 1, North Holland, 1992.
- [12] S. Hart, Evolutionary dynamics and backward induction, Games Econ. Behav. 41 (2002) 227–264.
- [13] S. Hart and A. Mas-Colell, Uncoupled dynamics do not lead to Nash equilibrium, American Economic Review, 93 (2005) 1830–1836.

- [14] M. Kandori, G. Mailath and R. Rob, Learning, mutation, and long-run equilibrium in games, *Econometrica*, 61 (1993) 29–56.
- [15] M. Kandori, R. Rob, Evolution of Equilibria in the Long Run: A General Theory and Applications, *Journal of Economic Theory*, 65 (1995) 383–414.
- [16] E. Kohlberg and J. Mertens, On the Strategic Stability of Equilibria, *Econometrica*, 54 (1986) 1003–1037.
- [17] D. Kreps and R. Wilson, Sequential Equilibria, *Econometrica*, 50 (1982) 863–894.
- [18] H. W. Kuhn, Extensive games and the problem of information. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games II. Annals of Mathematics Studies*, Vol. 28, Princeton University Press, 1953.
- [19] J. F. Nash, *Non-Cooperative Games*, PhD Dissertation at Princeton University, 1950.
- [20] H. Oosterbeek, R. Sloof and G. Van De Kuilen, Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis, *Experimental Economics*, 7 (2004) 171–188.
- [21] K. Ritzberger, *Foundations of Non-Cooperative Game Theory*, Oxford University Press, 2002.
- [22] K. Ritzberger and J. Weibull, Evolutionary Selection in Normal-form Games, *Econometrica*, 63 (1995) 1371–1399.
- [23] A. J. Robson and F. Vega-Redondo, Efficient equilibrium selection in evolutionary games with random matching, *Journal of Economic Theory*, 70 (1996) 65–92.
- [24] R. Rosenthal, Games of Perfect Information, Predatory Pricing, and the Chain Store, *Journal of Economic Theory* 25 (1998) 92-100.
- [25] L. Samuelson, *Evolutionary Games and Equilibrium Selection*, MIT Press, 1997.
- [26] W. H. Sandholm, *Population Games and Evolutionary Dynamics*, MIT Press, 2010.

- [27] Z. Xu, Fast evolution and a robustness index of backward induction, Center for the Study of Rationality, Hebrew University of Jerusalem, Mimeo.
- [28] Z. Xu, The instability of backward induction in evolutionary dynamics, Center for the Study of Rationality, Hebrew University of Jerusalem, Mimeo.
- [29] H.P. Young, The evolution of conventions, *Econometrica*, 61 (1993) 57–84.
- [30] H.P. Young, *Individual Strategy and Social Structure*, Princeton University Press, 1998.