Games between responsive behavioural rules

Abhimanyu Khan

Shiv Nadar University

8 December 2018

Games Between Responsive Behavioural Rules

Abhimanyu Khan*

December 8, 2018

Abstract

I study recurrent strategic interaction between two responsive behavioural rules in generic bi-matrix weakly acyclic games. The two individuals that play the game in a particular period choose their strategy by responding to a sample of strategies used by the co-players in the recent history of play. The response of a player is determined by his behavioural rule. I show that the game reaches a convention whenever the behavioural rule of each player is 'weakly responsive' to the manner in which strategies were chosen in the past by the co-players, and stays locked into the convention if the behavioural rules are 'mildly responsive'. Furthermore, amongst 'mildly responsive' behavioural rules, individuals described by the behavioural rule of 'extreme optimism' perform the best in the sense that their most preferred convention is always in the stochastically stable set; under an additional mild restriction that differentiates the behavioural rule of the other player from extreme optimism, the convention referred to above is the unique stochastically stable state.

JEL Classification: C73.

Keywords: evolution of conventions; behavioural rules; weakly responsive; mildly responsive; optimism

*Shiv Nadar University. E-mail: abhimanyu.khan@snu.edu.in

1 Introduction

This paper explores the evolution of outcomes in a recurrent strategic game. There are two distinct population of individuals, and in each period, one individual is randomly chosen from each population to play the game (for example, a bargaining game between share-croppers and landowners). It is often observed that conventions, or norms of behaviour, evolve in such situations – for instance, even though the process of bargaining between share-croppers and landowners may be represented by a bargaining game, it may be customary for share-croppers to receive half of the produce. In this paper, I answer two questions: firstly, positing that the recent history of the game influences current behaviour, what is the most general condition that can be made on how individuals in each population choose their strategy such that a convention evolves in the long-run, and that the behavioural rule of each individual further reinforces the established convention? Secondly, in situations where conventions do evolve in the long-run, and there are multiple conventions that may potentially evolve, is it possible to identify a behavioural rule that performs the best in the sense of obtaining its most preferred convention as the long-run outcome?

I analyse the questions posed above in a framework similar to adaptive play (Young (1993a)). There are two finite populations, and in each period, one individual is chosen from each population to play the game. The two populations have fixed roles in the strategic game: for example, randomly chosen individuals from population A always assume the role of the row player while randomly chosen individuals from population B assume the role of the column player. In order to choose a strategy in a particular period, the randomly chosen individual obtains information about (individuals of) the other population by drawing a sample from the finite history of strategies used by the latter in the previous periods, and then responds to the sample thus obtained. I associate this correspondence from the space of feasible samples to the strategies that are chosen with positive probability with a 'behavioural rule', and each population is characterised by a behavioural rule.

Restricting attention to generic weakly acyclic games (defined later), I show that a convention is established in the long-run (Proposition 1) and persists thereafter (Proposition 2) under very general conditions; in particular, this holds whenever, the behavioural rule is 'weakly responsive' and 'mildly responsive' to the information contained in the randomly drawn sample respectively. I argue (see Section 2) that conditional on individuals choosing a strategy by responding to the history of the game, it is reasonable to impose that any behavioural rule should satisfy 'weakly responsiveness' and 'mildly responsiveness'.

Following this, I ask, amongst the class of all 'mildly responsive' behavioural rules, is it possible to identify the behavioural rule that outperforms any other behavioural rule? For this purpose, I focus on one particular behavioural rule, namely 'extreme optimism'. The distinguishing feature of extremely optimistic individuals is that corresponding to any sample that may be drawn from the other population's previous period's strategies, they assume that amongst all strategies in the sample, their randomly chosen co-player will play the strategy that will make it possible for them to attain the most preferred outcome. Then, adding perturbation to the choice-making process of individuals (due to mistakes or experimentations), I show that the stochastically stable set always contains the most preferred convention of the population of extremely optimistic individuals. Furthermore, under an additional restriction on the behaviour of the other population that results in this other population being 'mildly different' from the population of extremely optimistic individuals, I show that the most preferred convention of the population of extremely optimistic individuals is the stochastically stable outcome.

If this strategic interaction between two populations is interpreted as a 'competition' between two behavioural rules, then this paper suggests that the behavioural rule of extreme optimism performs better than 'almost any other' behavioural rule. Depending on the nature of the game, the effect of extreme optimism on the other population may be beneficial or adversarial. I show that in strategic situations such as bargaining games, or Battle-of-the-sexes type coordination games, where there is a conflict of preferences, in the long-run, the other population is constrained to its least preferred convention. However, in games such as the minimum effort game, or Stag-hunt type coordination games, where the preferences are perfectly aligned, the other population also enjoys the benefits of obtaining its most preferred convention.

This paper contributes to the literature on how conventions may arise, and which conventions are more likely to emerge, in a decentralised environment where individuals myopically pursue their own interest. I build on the framework introduced in Young (1993a), where individuals best-respond to past information in an adaptive play model. Under a best-response dynamic, Young (1993a) shows that play converges almost surely to a convention, and then uses Foster and Young's (1990) concept of stochastic stability to analyse which conventions are more likely to be observed in the long-run. Kandori, Mailtah and Rob (1993) also study the stochastically stable outcome in coordination games with a best-response dynamic. Examples of other behavioural rules for which long-run outcomes of specific games have been studied include satisficing and aspirational play (Karandikar, Mookherjee, Ray and Vega-Redondo (1998) for 2×2 games) and imitation (Robson and Vega-Redondo (1996), Josephson and Matros (2004), Bergin and Bernhardt (2009)) However, in the afore-mentioned papers, all players share a particular behavioural rule. While Kaniovski, Kryazhimskii and Young (2000), Jospehson (2009) and Juang (2002) study heterogeneous behavioural rules, the heterogeneity is not across individuals but rather models the fact that the same individual may use different rules (for example, he may probabilistically best-respond or imitate); moreover, the scope of the analysis is restricted to specific games (for eg. 2×2 co-ordination games). In this paper, I look at a broad class of games with the intention of analysing which particular behavioural does better when it interacts with individuals of another behavioural rule. Relatedly, Axelrod (1984) studies the success of various strategies in the repeated Prisoner's Dilemma. While my paper shares the similarity of comparing the success of various rules/strategies, the focus here is on recurrent interaction in a broad class of games and I abstract away from repeated game considerations in order to isolate the effect of behavioural rules.

2 Model and Results

Let G represent a bi-matrix game. There are two finite populations, denoted by A and B. Time is discrete and in each time period, one individual from each population is randomly chosen, and these two individuals play that game in that period. Each individual of each population has positive probability of being chosen in any particular period. I denote the randomly chosen individual in population A and B in period t as $A(t)$ and $B(t)$ respectively. X_A and X_B denote the finite time-invariant strategy set of individuals belonging to population A and population B respectively. The strategy played by the randomly chosen individual from population J in period t is denoted by $x_J(t) \in X_J$, where $J = A, B$; the strategy chosen by the randomly chosen individual from the other population is denoted by $x_{-J}(t) \in X_{-J}$. Thus, the period t outcome is $x(t) = (x_A(t), x_B(t))$, and $X = X_A \times X_B$ is the space of outcomes. I emphasise that the situation is one of recurrent interaction, and I abstract away from any consideration that may arise out of repeated interactions. $\succeq_A$ and $\succeq_B$, each defined over X, is the binary rational preference relation of each individual in population A and population B respectively; for any two outcomes $x, y \in X, x \succeq_J y$ implies that outcome x is at least as good as outcome y for individuals in population J. $\succ_J$ and $\sim_J$ are the strict preference relation and indifference relation induced by $\succeq_J$.[1]

The state of the game at the beginning of period $t+1$ (when each individual is yet to decide on his strategy) is given by the finite M period history of strategies i.e. $\omega(t+1) = (x(t), x(t-1), \ldots, x(t-M+1)) = ((x_A(t), x_B(t)), (x_A(t-1), x_B(t-1)), \ldots, (x_A(t-M+1), x_B(t-M+1)))$; $\omega_A(t+1) = (x_A(t), x_A(t-1), \ldots, x_A(t-M+1))$ and $\omega_B(t+1) = (x_B(t), x_B(t-1), \ldots, x_B(t-M+1))$ denote the strategies used by the randomly chosen individuals in population A and population B respectively in the preceding M periods. The state space is denoted by $\Omega = (X_A \times X_B)^M$. In the beginning of period $t+1$, individual $J(t+1)$ from population J draws a random sample (without replacement) of strategies of exogenously fixed size S from $\omega_{-J}(t)$, and I assume that every feasible sample has positive probability of being drawn. Let $s_J(t+1)$ denote the sample drawn by $J(t+1)$ in period $t+1$, and $supp(s_J(t+1)) = \{x : x \in s_J(t+1)\}$ be the support of sample $s_J(t+1)$. Individuals

[1] That is, for any two outcomes $x, y \in X, x \succ_J y \Leftrightarrow x \succeq_J y$ and not $y \succeq_J x$, while $x \sim_J y \Leftrightarrow x \succeq_J y$ and $y \succeq_J x$.

in population J are associated with a time-invariant behavioural rule, or a time-invariant response correspondence, $R_J : X^S_{-J} \rightrightarrows P(X_J)$ that goes from the space of feasible samples to the power set of the strategy set X_J that contains all the non-empty subsets of X_J. If $s_J(t+1)$ is the sample drawn by individual $J(t+1)$, then each element of the response set $R_J(s_J(t+1))$ has strictly positive probability of being played by him in the current period $t+1$.

Now, for any $y \in X_B$, I define $C_A(y) = \{x \in X_A : (x, y) \succeq_A (z, y), \forall z \in X_A\}$, i.e. assuming that the co-player of the population A individual plays strategy y, the set $C_A(y)$ comprises of all those strategies that lead to outcomes that are not strictly dominated by an outcome that arises from playing a strategy outside this set $C_A(y)$ for the population A individuals. Similarly, for any $x \in X_A$, $C_B(x) = \{y \in X_B : (x, y) \succeq_B (x, z), \forall z \in X_B\}$.

I restrict attention to the class of generic bi-matrix games G' that are weakly acyclic. I define a bi-matrix game to be generic if for any two disparate outcomes $(x, y) \in X_A \times X_B$ and $(x', y') \in X_A \times X_B$, with $x \neq x'$ or $y \neq y'$ (or both), either $(x, y) \succ_J (x', y')$ or $(x', y') \succ_J (x, y)$ for any $J \in \{A, B\}$. Next, in order to define a weakly acyclic game (in the sense of Young (1993a)), I first introduce the concepts of an improving path and a sink. An improving path exists between two outcomes $x' = (x'_A, x'_B)$ and $x'' = (x''_A, x''_B)$ if either (i) $x'_A = x''_A$, $x'_B \neq x''_B$, and $x''_B \in C_B(x'_A)$, so that $(x''_A, x''_B) \succeq_B (x'_A, x'_B)$, or (ii) $x'_A \neq x''_A$, $x'_B = x''_B$, and $x''_A \in C_A(x'_B)$, so that $(x''_A, x''_B) \succeq_A (x'_A, x'_B)$; I represent this improving path by $x' \to x''$. An outcome is defined to be a sink if there does not exist an improving path from it. Then, a game is weakly acyclic if every outcome in X is either a sink, or there exists a sequence of improving paths $x^0 = (x^0_A, x^0_B) \to \ldots \to x^n = (x^n_A, x^n_B)$ from every outcome x^0 that is not a sink to an outcome x^n that is a sink, and where $x^0, \ldots, x^n \in X$.[2] Some prominent examples of weakly acyclic games include the (Nash) bargaining game, co-ordination games, the Stag-Hunt game and the minimum effort game.

The behavioural rule of population J, R_J, is *weakly responsive* if $supp(s_J(t+1)) = \{x\} \in X_{-J}$ implies $C_J(x) \cap R_J(s_J(t+1)) \neq \emptyset$. That is, whenever a weakly responsive individual draws a sample of the other population's strategies that only consists of one particular strategy x, then at least one strategy that leads to an outcome that is at least as good as any other outcome (given that the other player chooses x) is always in his response set. Weak responsiveness does not impose $C_J(x) \subset R_J(s_J(t+1))$, or $C_J(x) = R_J(s_J(t+1))$ when $s_J(t+1) = \{x\}$, for some $x \in X_{-J}$ i.e. not all strategies in $C_J(x)$ need appear in $R_J(s_J(t+1))$, and $R_J(s_J(t+1))$ may contain strategies other than the strategies in $C_J(x)$. Neither does weak responsiveness place any restriction on the behaviour of any individual for any other sample apart from the ones which are comprised of only one particular strategy.

I argue that this is the weakest condition that must be satisfied by any behavioural rule

[2]As an example, consider a game where the strategies x_A and x_B are dominated by the strategies x'_A and x'_B for population A and population B individuals, and that (x'_A, x'_B) is a sink outcome. Then, one example of an improving path from (x_A, x_B) to (x'_A, x'_B) is $(x_A, x_B) \to (x'_A, x_B) \to (x'_A, x'_B)$.

that responds to past information; after all, conditional on individual's response depending on the obtained sample, the situation where a sample comprises of only one particular strategy is the simplest decision making situation for an individual; if the above condition does not hold for a particular behavioural rule, i.e. if $C_J(x) \cap R_J(s_J(t+1)) = \emptyset$ for some $supp(s_J(t+1)) = \{x\} \subset X_{-J}$, then the behavioural rule is not responsive to the information contained in the sample even in the simplest possible case. It is in this sense that I term weak responsiveness as the weakest or most general condition on responsive behavioural rules.

This strategic situation, where the two randomly chosen individuals from each population play the game by choosing a strategy from their respective response sets that is generated by their behavioural rule, can be described by a Markov process Q on the state space Ω. A convention is a state $\omega \in \Omega$ if the same outcome, say $(\bar{x}, \bar{y})$, has been realised in all the past M periods, and in addition, the outcome $(\bar{x}, \bar{y})$ is a sink; this particular convention is denoted by $\omega_{\bar{x},\bar{y}}$. $Co(G')$ denotes the set of conventions of the game G', and $Co(G')_J = \{\omega_{x,y} \in Co(G')$ such that $(x, y) \succeq_J (x', y'), \forall \omega_{x',y'} \in Co(G')\}$ is the set of conventions that are not strictly dominated by other conventions according to the preference relation of population J. Since I restrict attention to generic games, $Co(G')_J$ is a singleton and I refer to it as the most preferred convention for individuals of population J. The game is said to reach a convention almost surely if the probability of a transition to a convention from any initial state is strictly positive; the game is said to converge almost surely to a convention if the game reaches a convention almost surely and then stays in the convention thereafter. Then Proposition 1 follows.

Proposition 1. *Suppose that the sampling is sufficiently incomplete, and the behavioural rule of each population is weakly responsive. Then any bi-matrix game weakly acyclic game reaches a convention almost surely.*

Proof. Let $\omega(t+1)$ denote the state at the beginning of time period $t+1$. Then, the following events occur with positive probability:
(i) In all time periods $t+1, \ldots, t+S$, the randomly chosen individual from either population draws the strategies played by individuals of the other population in the periods $t-S+1, \ldots, t$; since the same sample is drawn by individuals of a particular population, the population A individuals play the same strategy $\bar{x}_A$ while the population B individuals play the same strategy $\bar{x}_B$ in all these periods. Hence, the outcome $(\bar{x}_A, \bar{x}_B)$ obtains in all the time periods $t+1, \ldots, t+S$.
(ii) Now, suppose that the outcome $(\bar{x}_A, \bar{x}_B)$ is a sink. Then, in each time period from $t+S+1$ onwards, individuals from either population draw a sample that comprises of the strategies of the preceding S periods. So, the sample drawn by a population A (similarly, population B) individual comprises only of $\bar{x}_B$ (similarly, $\bar{x}_A$); as $(\bar{x}_A, \bar{x}_B)$ is a sink outcome, population A (similarly, population B) individuals continue to play $\bar{x}_A$ (similarly, $\bar{x}_B$), so that in all time

periods from $t+1, \ldots, t+M$, the history of strategies is made up of population A (similarly, population B) individuals playing $\bar{x}_A$ (similarly, $\bar{x}_B$). In any period that follows, any sample drawn by a population A (similarly, population B) individual comprises only of $\bar{x}_B$ (similarly, $\bar{x}_A$), and so, these individuals choose $\bar{x}_A$ and $\bar{x}_B$. A convention is thus reached almost surely. (iii) Suppose, on the other hand, that the outcome $(\bar{x}_A, \bar{x}_B)$ is not a sink. Since the game is weakly acyclic, there exists a sequence of improving paths from this outcome to a sink. Let $x^0 \to \ldots \to x^n$ be such a sequence of improving paths, where x^i is the outcome (x^i_A, x^i_B) for $i = \{0, \ldots, n\}$, $x^0 = (\bar{x}_A, \bar{x}_B)$, and the outcome x^n is a sink. Without loss of generality, the improving path $x^0 \to x^1$ involves outcome x^1 such that $x^0_A \neq x^1_A$, $x^0_B = x^1_B$, and $(x^1_A, x^0_B) \succeq_A (x^0_A, x^0_B)$. Now, with positive probability, in the time periods $t+S+1, \ldots, t+2\,S$, the population A individuals' sample comprises of the strategies $(x_B(t+1), \ldots, x_B(t+S)) = (\bar{x}_B, \ldots, \bar{x}_B)$ while the population B individuals' sample comprises of the strategies $(x_B(t-S+1), \ldots, x_B(t))$. As a result, with positive probability, the population A individuals play x^1_A and the population B individuals play x^0_B in all the time periods $t+S+1, \ldots, t+2\,S$, due to which, at the end of period $t+2\,S$, the S-period history comprises only of the outcomes $x^1 = (x^1_A, x^0_B)$. If this outcome is a sink, then (ii) above applies; if not, then using the argument outlined above, it can be show that in all of the next S periods, the outcome corresponds to outcome x^2 of the sequence of improving paths; with positive probability, the process proceeds in this manner till the there is a S-period run of the sink outcome x^n, at which point (ii) above applies. Hence, again, a convention is reached almost surely. ■

The proposition above shows that for the game to converge almost surely to a convention, the only requirement is that each behavioural rule is weakly responsive. However, weak responsiveness is not sufficient for the persistence of an established convention – weak responsiveness in generic games implies that if $supp(s_J(t+1)) = \{x\}$, then $C_J(x) \cap R_J(s_J(t+1)) \neq \emptyset$. So, if $R_J(s_J(t+1)) \setminus C_J(x)$ is non-empty, then a weakly responsive individual responds with a strategy in the set $R_J(s_J(t+1)) \setminus C_J(x)$ with positive probability, causing a transition from the convention. Now, I define the behavioural rule R_J to be *mildly responsive* if $supp(s_J) = \{x\}$ implies $R_J(s_J(t+1)) = C_J(x)$ for any $x \in X_{-J}$. Hence, when a mildly responsive individual draws a sample that comprises only of a particular x, his response set is equivalent to $C_J(x)$; however, mild responsiveness places no other restriction on the behaviour of an individual when other samples are drawn. Then, because a population of mildly responsive individuals is also weakly responsive, Proposition 2 below follows directly from Proposition 1 and the definition of mild responsiveness, and hence, is stated without proof. Further, for the rest of the paper, I assume that the behavioural rule of either population is mildly responsive.

Proposition 2. *Suppose that the sampling is sufficiently incomplete, and the behavioural rule of each population is mildly responsive. Then any generic weakly acyclic game converges almost surely to a convention.*

In order to capture the feature that conventions, at times, may be borne out of exogenous shocks or historical accidents, I now introduce the possibility of mistakes or experimentation in the decision-making process of an individual. With independent probability $\varepsilon > 0$, the randomly chosen individual from population J makes a mistake, or experiments, by playing a strategy $x_J(t+1) \notin R_J(s_J(t+1))$ in time period $t+1$ for any feasible sample $s_J(t+1)$ that can be drawn from $\omega_{-J}(t+1)$, i.e. an individual plays a strategy that is not dictated by the behavioural rule of his parent population. Due to experimentations/mistakes, it is now possible to transit from one convention to another. This perturbed Markov process, represented by $Q(\varepsilon)$, is ergodic; for each $\varepsilon > 0$, there exists a unique stationary distribution $\mu(\varepsilon)$ given by $\mu(\varepsilon)Q(\varepsilon) = \mu(\varepsilon)$. The stochastically stable set – see Foster and Young (1990), Kandori, Mailath and Rob (1993) and Young (1993a) for details– is the set of states that receive positive weight in the limiting stationary distribution $\mu^* = lim_{\varepsilon \downarrow 0} \mu(\varepsilon)$. For the purpose of this paper, the identification of this set is determined by the relative difficulty/ease of transition from one set of conventions to the complementary set of conventions. According to Theorem 3 in Ellison (2000), if it is possible to transit into a particular convention from any other convention with a single experimentation, then the said convention is in the stochastically stable set; if, in addition, a transition from the particular convention needs strictly more than one experimentation, then it is the unique stochastically stable state.

I now focus on one particular behavioural rule, namely 'extreme optimism'. The individuals in population J are said to be *extremely optimistic* if any individual from the population assesses, on drawing a sample $s_J(t+1)$ in any time period $t+1$, that his randomly chosen co-player from the other population will choose a strategy from the *assessment set* $AS(s_J(t+1)) = \{x \in s_J(t+1) : \exists y \in X_J$ such that $(y,x) \succeq_J (y',x'), \forall y' \in X_J$ and $\forall x' \in s_J(t+1)\}$; consequently, in period $t+1$, he plays a strategy in the set $C_J(x)$, for some $x \in AS(s_J(t+1))$. Thus, a extremely optimistic individual believes that he is going to face only the most favourable of circumstances – amongst all the strategies in the drawn sample, his co-player will play a strategy that will lead to the most favourable outcome for him – and he acts accordingly. The next proposition shows the long-run advantage of this behavioural trait.

Proposition 3. *Suppose that individuals of a particular population are extremely optimistic, and all individuals of the other population are mildly responsive. If sampling is sufficiently incomplete, then in any generic weakly acyclic game, the most preferred convention of the extremely optimistic population is always in the stochastically stable set.*

Proof. Without loss of generality, let population A individuals be extremely optimistic. Consider a convention $\omega_{x,y} \in Co(G')_A$; since it is a sink outcome, at least one experimentation is needed to transit from (the basin of attraction) of $\omega_{x,y}$. (That is, the radius of $\omega_{x,y}$ is at least equal to one.) So, in order to prove the proposition, it is sufficient to show that a transition to a convention $\omega_{x,y}$ from any other convention $\omega_{x',y'} \in Co(G')$ is possible with a single mistake/experimentation (Theorem 3, Ellison (2000).)

Suppose that the initial convention in period t is $\omega_{x',y'}$, and that in period t, the randomly chosen individual in population B experiments, or makes a mistake, by playing the strategy y. Then in all periods $t+1$ to $t+M$, with positive probability, the population A individual draws a sample that contains this strategy y; extreme optimism implies $y \in AS(s_A(t+k))$, for all $k = 1, \ldots, M$. Since $\omega_{x,y} \in Co(G')_A$, $x \in R_A(s_A(t+1))$ whenever $y \in A(s_A(t+1))$, and so, $x_A(t+k) = x$, for all $k = 1, \ldots, M$. Also, with positive probability, in all time periods $t+S+1$ to $t+M$, the randomly chosen individual from population B draws the sample from the previous S periods; this sample comprises only of x, and because the outcome (x, y) is a sink, the population B individual plays y in all these periods. Thus, the outcome in all the S periods from $t+S+1$ to $t+M$ is given by (x, y). Then, in all the $M-S$ periods from $t+M+1$ to $t+2M-S$, the randomly chosen individuals from either population (with positive probability) draw a sample from the previous S periods – these samples comprise only of x for population B individuals , and y for population A individual. As a result, the strategies x and y are played in all these periods by the population A and population B individuals respectively. It follows that at the end of period $t+2\,M-S$, the M period history comprises only of the outcomes (x, y). The game converges to the convention $\omega_{x,y}$, and the transition to this convention from any other convention has been shown to be possible by a single experimentation, thus proving the proposition. ■

While the above proposition hints at the long-run advantage that extreme optimism may confer, it does not exclude the possibility that conventions that are least preferred by the extremely optimistic population are also in the stochastically stable set. I now argue that generic weakly acyclic games, if the other population's behavioural rule is mildly different from extreme optimism, then the most preferred convention of the extremely optimistic population is the unique stochastically stable outcome. For this purpose, I now define what it means for the other population to be mildly different from the extremely optimistic population. Without loss of generality, I assume that the behavioural rule of population A is given by extreme optimism. Let $T = \{x \in X_A : \exists y \in X_B \text{ such that } \omega_{x,y} \in Co(G'_A)\}$ be the set of strategies for population A so that for every strategy in this set, there exists a strategy y for population B such that $\omega_{x,y}$ belongs to the set of most preferred conventions for population A individuals. Population B is said to be *mildly different* from population A population if the following condition holds: whenever a population B individual draws the specific sample s that comprises of $S-1$ occurrences of a particular strategy $x \in T$ and a single occurrence of any other strategy, then $R_B(s) = C_B(x)$. Here, I stress on the fact that the proportion with which the 'other' strategy occurs in the sample can be arbitrarily small, as the S (the size of the sample) can be arbitrarily large – it is in this sense that population B is mildly different from extremely optimistic population. I now show that the most preferred convention of the extremely optimistic population is the unique stochastically stable outcome.

Proposition 4. *Suppose that the response correspondence of population A exhibits extreme optimism, and population B's response correspondence is mildly different from it. If sampling is sufficiently incomplete, then the convention that is most preferred by the population of extremely optimistic individuals is the uniquely stochastically stable state of any generic weakly acyclic game.*

Proof. Let population A's most preferred convention be denoted by $\omega_{x,y}$. Proposition 3 shows that $\omega_{x,y}$ is always in the stochastically stable set because a transition into it from any other convention is possible with a single experimentation/mistake. So, to complete the proof, it is sufficient to show that a transition out of $\omega_{x,y}$ is not possible by a single experimentation/mistake.
Suppose that the current convention is $\omega_{x,y}$, and in a particular time period t, either an individual in population A experiments/makes a mistake by playing $x' \neq x$, or an individual in population B experiments/makes a mistake, by playing $y' \neq y$. In either case, any sample that a population A can draw in any of the succeeding time periods comprises of at least $S-1$ instances of strategy y and at most one occurrence of strategy y'; similarly, any sample that a population B can draw in any of the succeeding time periods comprises of at least $S-1$ instances of strategy x and at most one occurrence of strategy x'. Since the individuals in population A are extremely optimistic, the response to any sample that they can draw in any time period in the future is x; because population B is mildly different from population A, the response of any population B individual to any sample that they can draw in any time period in the future is y. As a result, the only outcome in any time period in the future that arises from the population A and population B individuals choosing a strategy by their behavioural rule is (x, y), and so, after M time periods following the time period when the mistake/experimentation appears, the convention $\omega_{x,y}$ is re-established. Hence, it is not possible to transit from $\omega_{x,y}$ with one experimentation/mistake. ■

Example. Consider the bi-matrix coordination game, with the payoff matrix below, where individuals from population A assume the role of the row player and individuals from population B assume the role of the column player. $X_A = \{U, D\}$ and $X_B = \{L, R\}$. I assume $a > e, g > c, b > d$ and $h > f$. Thus, (U, L) and (D, R) are the sink outcomes (and the Nash equilibria). Suppose without loss of generality that $a > g$; then population A, whose response correspondence displays extreme optimism, prefers the outcome (U, L) over (D, R). This is representative of coordination games, such as the Stag-hunt game (a necessary condition for this $b > h$ – here the individuals have the same preferences over the two sink outcomes) or Battle of the Sexes (this occurs when $h > b$ – here the two individuals differ in their preference over the two sink outcomes).

In this game, extreme optimism implies that if an individual from population A draws a sample that contains at least one instance of strategy L being played, then he assesses that his randomly matched co-player will play L; consequently, he plays U. Population B is said to be *mildly different* from population A population if for the specific sample comprising of $S-1$ occurrences of strategy U and a single occurrence of strategy D, the response set of any individual in the population B does not contain the strategy R. Then, Proposition 4 implies that the unique stochastically stable state is the convention $\omega_{U,L}$. The general message of this example extends to $N \times N$ coordination games (i.e. coordination games where each population has N strategies): in case of Stag-Hunt type co-ordination games, where the preferences of individuals of either population are perfectly aligned, the commonly preferred outcome is the uniquely stable outcome in the long-run; however, in the case of Battle of the Sexes type coordination games, where preferences over the conventions are oppositely aligned, the other population is constrained to its least preferred outcome.

	L	R
U	a, b	c, d
D	e, f	g, h

■

3 Conclusion

In this paper, I analyse the long-run stochastically stable outcome when individuals belonging to two different populations play a game with/against each other in a recurrent framework. In each period of this recurrent game, an individual is randomly chosen from each population and these two individuals play the game in that period. In order to choose a strategy, individuals randomly draw a sample from the finite history of strategies used by individuals of the other population, and respond to this sample in a manner dictated by the behavioural rule of their parent population. A convention is established whenever the behavioural rule of either population satisfies a very general condition, namely weak responsiveness, and a mild strengthening of weak responsiveness, namely mild responsiveness, ensures that the game stays locked into that established convention.

The next objective is examine the relative success of various mildly responsive behavioural rules by looking at the outcomes obtained in the stochastically stable set. Here, I firstly show that if a population exhibits extreme optimism, then the convention that is most preferred by it is always in the stochastically stable set. Next, I show that under a reasonable restriction on the behavioural rule of the other population that differentiates it from a population exhibiting extreme optimism, the convention that is most preferred by the latter is unique stochastically stable state. Thus, extreme optimism imposes an ‘externality’ on the other behavioural rules, and depending on the nature of the game, the 'externality' might benefit or constrain the other population. In games where the interests of the two populations are oppositely aligned – for example, fixed-sum games such as the bargaining game, or Battle of the Sexes type of co-ordination games – the other population is constrained to the worst possible outcome; in this case, one might say that extreme optimism outperforms any other behavioural rule. However, in games where there is no conflict of interest - for example, Stag-hunt games or minimum effort games – the other population shares the spoils of the most preferred convention of the population displaying extreme optimism emerging as the only stable outcome in the long-run.

References

1. Axelrod R (1984). The Evolution of Cooperation. New York: Basic Books

2. Bergin J and Bernhardt D (2009). Cooperation through imitation. Games and Economic Behaviour 67 (2): 376-388

3. Ellison G (2000). Basins of Attraction, Long Run Stochastic Stability, and the Speed of Step-by-Step Evolution. Review of Economic Studies 67 (1): 17-45

4. Foster D and Young HP (1990). Stochastic Evolutionary Games Dynamics. Journal of Theoretical Population Biology 38 (2): 219 - 232.

5. Juang WT (2002). Rule Evolution and Equilibrium Selection. Games and Economic Behaviour 39 (1): 71-90

6. Josephson J and Matros A (2004). Stochastic imitation in finite games. Games and Economic Behaviour 49 (2): 244-259

7. Josephson J (2009). Stochastic adaptation in finite games played by heterogeneous populations. Journal of Economic Dynamics and Control 33 (8): 1543-1554

8. Kandori M, Mailath G and Rob R (1993). Learning, Mutation, and Long Run Equilibria in Games. Econometrica 61 (1): 29-56

9. Kaniovski YM, Kryazhimskii AV and Young HP (2000). Adaptive Dynamics in Games Played by Heterogeneous Populations. Games and Economic Behaviour 31 (1): 50-96

10. Karandikar R, Mookerjee D, Ray D and Vega-Redondo F (1998). Evolving Aspirations and Cooperation. Journal of Economic Theory 80 (2): 292-331

11. Robson AJ and Vega-Redondo F (1996). Efficient Equilibrium Selection in Evolutionary Games with Random Matching. Journal of Economic Theory 70 (1): 65-92

12. Young HP (1993a). The evolution of conventions. Econometrica 61 (1): 57-84

13. Young HP (1993b). An evolutionary model of bargaining. Journal of Economic Theory 59 (1): 145-168