Chapter 1

Learning from Personal Experience and the Selection of Equilibria

First version: July 1999
This version: June 2001

Abstract

We study learning from personal experience in a large, finite population. Every period players are paired at random to play a symmetric $2 \times 2$ coordination game. Each player retains a finite idiosyncratic history of his own past plays and chooses his strategy as the best response thereto. Although not leading to payoff-positive dynamics, this learning process does converge to a state in which all players agree on the same Nash equilibrium. When small-probability mistakes are superimposed on the learning process, the model select the risk-dominant equilibrium. This result accords with many recent evolutionary models (e.g., Kandori et al., 1993, Young, 1993). In our model, however, the stochastic stability of the risk-dominant equilibrium is more pronounced: the number of mistakes required to upset this equilibrium increases with the population size whereas the analogous number for the Pareto-dominant equilibrium is constant.

JEL Classification: C73, C72

Keywords: learning, large population, personal experience, stochastic stability, selection of equilibria, coordination game
1 Introduction

Although the capacity to learn from others’ experience makes us human, personal experience is pivotal to our learning. Thomas, a Jesus’ contemporary, often dubbed ‘Doubting,’ is a well-known exemplar of this truth. Despite his friends’ insistent pleas, Thomas did not start believing in Jesus until he had touched Jesus’ wound (John 20:27). All persuasion having failed, a single personal experience was enough to bring about Thomas’ conversion.

In spite of its importance, learning from personal experience has not been explored adequately in a large-population context. A majority of large-population models build on different modes of learning that rely on transmission of experience (information) among players. Young (1993a, 1993b), for instance, presupposes that players react to a sample of recent plays, making no distinction between their own plays and other players’ plays. In his model, players learn effectively from the experience of the whole society. In other models (e.g., Kandori, Mailath and Rob—henceforth KMR, 1993; Ellison, 1993; Robson and Vega-Redondo, 1996; Canals and Vega-Redondo, 1998) players choose the best response to the current distribution of strategies in the population, hence ignoring their past experience completely. In yet other models (Schlag 1998; Eshel, Samuelson and Shaked, 1998; Eshel, Sansone and Shaked, 1996) players’ strategy revision is driven by imitation. Canning (1992) is the only author we know who considers learning from personal experience, if as a special case of a more general dynamics. Owing to the generality of his model, however, his predictions for the long-run behavior of the population are quite limited.

To complement the existing literature on learning in a large population and to see how learning from personal experience differs from other modes of learning, we study a model where players learn exclusively from their personal experience. Above all, we are interested in how such learning, when perturbed by occasional mistakes, resolves the problem of selection among multiple (Nash) equilibria.

A finite population of players is paired at random, once in every period, to play a game. To keep things simple yet interesting, we confine our attention to symmetric $2 \times 2$ coordination games in which the Pareto-dominant equilibrium (PDE) differs from the risk-dominant equilibrium (RDE). Each player retains a finite idiosyncratic history of his own past plays and chooses his strategy as the best response to this history.

We show that this type of learning does not lead to payoff-positive dynamics on the level of the whole population. In other words, it may happen that the share of players using the best response to the current population strategy
profile declines from period to period.\(^1\) For this reason, learning from personal experience does not fall into the class of dynamics studied by KMR (1993) and their followers (Ellison, 1993, Robson and Vega-Redondo, 1996). Just as payoff-positive dynamics, however, learning from personal experience converges to a state in which all players coordinate one of the Nash equilibria of the stage game. Which of the two equilibria this is depends on both the initial profile of players’ histories and the realized pairing patterns.

We then allow that players make occasional mistakes. We suppose that with a small probability independent of his history each player fails to follow his learning rule, choosing his strategy at random. As customary in large-population learning models, we suppose that players’ proclivities to making mistakes are of comparable magnitudes.

To investigate the implications of players’ mistakes on learning dynamics, we invoke the concept of stochastic stability, developed independently by Young (1993a) and KMR (1993).\(^2\) A state (a profile of players’ histories) is stochastically stable if the Markov process describing the learning dynamics perturbed by mistakes places a positive weight on this state when the probability of mistakes tends to zero. Stochastically stable states are the most difficult to upset by players’ mistakes.

We prove that given the sufficient length of players’ histories there exists a unique stochastically stable state: the state where all players play the RDE. This finding accords well with the simulations run by Di Gioacchino (1992) and with an example offered by Canning (1992, pp. 459-463). It also reinforces Young’s (1993a) and KMR’s (1993) findings on equilibrium selection in \(2 \times 2\) coordination games. Under the learning-from-personal-experience dynamics, however, stochastic stability of the RDE is more pronounced. While in both Young’s and KMR’s models the number of mistakes required to upset the RDE is proportional to the loss from deviating from the RDE (and the analogue is true of the PDE), in our model this is not the case. Specifically, the number of mistakes required to upset the RPE is very large and generally increases with population size.

Since both Young’s (1993a) model (in which players share their experience completely) and our model (in which players rely on their personal experience only) predict the RDE to be stochastically stable, we conjecture that the same

---

\(^1\)Weibull (1998) makes a distinction between payoff-positive dynamics and weakly payoff-positive dynamics. The former arises when all strategies performing above average (given the current population strategy profile) exhibit positive growth rates, the latter when at least one strategy performing above average has this property. Since in \(2 \times 2\) games only one strategy can perform above average, in our case both definitions coincide.

will be true of any model between them (a model in which players rely on a mix of their personal experience and the experience of their fellows). In fact, the way we prove stochastic stability of the RDE points strongly towards the conclusion that RDE will be stochastically stable whenever each player chooses the best response to some finite, continually updated history.

The rest of the paper is organized as follows. In the next section we describe how players choose their strategies (i.e., how they learn). In section 3 we discuss the dynamics of players’ learning in the absence of mistakes. In section 4 we allow the learning dynamics to become perturbed by players’ mistakes. In section 5 we analyze stochastic stability under the perturbed learning dynamics and derive the main finding on equilibrium selection. A discussion and conclusion close the paper.

2 The Model: Social Adaptive Play

Time elapses in discrete units, $t = 1, 2, \ldots$. An even number $N$ of players are paired at random in each period. The interaction within a pair of players takes the form of a symmetric $2 \times 2$ coordination game $\Gamma$

$$
\begin{array}{c}
\text{II} \\
\text{I} \\
\Gamma: \\
C & D \\
\hline
C & a, a & b, c \\
D & c, b & d, d
\end{array}
$$

(1)

Here $C$ and $D$ are pure strategies of the game, $a, b, c$ and $d \in \mathbb{R}$ and $a \geq c, b \leq d$ so that both $(C, C)$ and $(D, D)$ are Nash equilibria. In addition, we assume that $a > d$ and $a - c < d - b$. Then, $(C, C)$ is the Pareto-dominant equilibrium (PDE) and $(D, D)$ the risk-dominant equilibrium (RDE) in the sense of Harsanyi and Selten (1988). We shall find it useful to refer to $a - c$ as a loss from deviating from $(C, C)$, in short LD from $(C, C)$, and to $d - b$ as a loss from deviating from $(D, D)$, in short LD from $(D, D)$.

Players approach each interaction myopically (they do not believe their decision can substantially influence the other players’ future decisions, because,

---

$^3$By random pairing we mean that each possible pairing pattern $\mu$—a pair-wise partition of $\{1, \ldots, N\}$—is realized with positive probability. This probability does not have to be the same for different pairing patterns.

$^4$In game $\Gamma$, strategy $D$ is a risk-dominant strategy of player 1 if it yields greater expected payoff to this player than strategy $C$, given that player 2 plays both his strategies with equal probabilities. Otherwise, strategy $C$ is a risk-dominant strategy of player 1. The analogue holds for player 2. Numerically, strategy $D$ is a risk-dominant strategy of both players if $\frac{1}{2}c + \frac{1}{2}d > \frac{1}{2}a + \frac{1}{2}b$, i.e., if $a - c < d - b$.  

---
perhaps, \( N \) is quite large) and they decide which strategy to choose solely on the basis of their own personal experience (past interactions). All players have limited memory: they remember only their last \( m \) interactions. The informational set of player \( i \in \{1, \ldots, N\} \) at the beginning of period \( t \) can therefore be characterized by the history

\[
h^i(t) = (s^i(t-m), s^i(t-m+1), \ldots, s^i(t-1)),
\]

where \( s^i(t-\tau) \in C, D \) denotes the strategy used by \( i \)'s opponent \( \tau \) periods ago.\(^5\) Players perceive histories as anonymous (i.e., for no \( i \in \{1, \ldots, N\} \) does the identity of their past opponent matter) and they assign the same weight to each observation, however recent it is. Under these conditions, game \( \Gamma \) may be succinctly characterized by a pair of positive integers \( n_c \) and \( n_d \), where \( n_c \) is the minimum number of Cs in a history that makes a player play \( C \), and \( n_d \) is the minimum number of Ds in a history that makes a player play \( D \). It is easy to see that\(^6\)

\[
\begin{align*}
n_c &= \left\lfloor \frac{d - b}{(d - b) + (a - c)} \right\rfloor m; \\
n_d &= \left\lfloor \frac{a - c}{(d - b) + (a - c)} \right\rfloor m,
\end{align*}
\]

and

\[
n_c + n_d = m + 1. \tag{4}
\]

The assumption \( a - c < d - b \) implies that \( n_c \geq n_d \). Note that for large \( m \), \( n_c \) is almost proportional to LD from \((D, D)\), whereas \( n_d \) is almost proportional to LD from \((C, C)\).

Each player \( i \) chooses his strategy \( s^i(t) \) to play the game \( \Gamma \) at period \( t \) as the best response to \( h^i(t) \). To avoid the uninteresting degenerate case, we assume that \( s^i(t) \) can always be chosen in a unique way, i.e., that \( n_c \neq n_d \). This implies that \( n_c > n_d \). By virtue of (4), it follows that

\[
n_d \leq m - 1.
\]

Since there is no uncertainty about players’ strategy choice, the evolution of the process depends only on a particular pairing pattern. The two following definitions will be helpful in depicting this evolution in algebraic terms (the first follows Young, 1993a):

\(^{5}\)We assume that at \( t = 1 \) players are endowed with some ‘fictitious’ history that might be interpreted as their initial belief about how the game had been played in the past.

\(^{6}\)Here and in what follows \( \lfloor x \rfloor \) denotes the smallest integer greater than or equal to \( x \).
SUCCESSOR: We call history \( \hat{h} \) a successor of history \( h \) if \( \hat{h} \) was created by deleting the left-most element of \( h \) and adjoining a new right-most element. We call \( h(t) = (h^1(t), h^2(t), \ldots, h^N(t)) \) the population history at time \( t \), and denote the space of all population histories as \( \mathcal{H} \).

FOLLOWER: We call population history \( \tilde{h} = (\tilde{h}^1, \tilde{h}^2, \ldots, \tilde{h}^N) \) the follower of population history \( h = (h^1, h^2, \ldots, h^N) \) if

(i) Each \( \tilde{h}^i \) is the successor of \( h^i \).

(ii) There exists a pairing pattern \( \mu(h, \tilde{h}) \) such that for every pair of players \( \{i, j\} \in \mu(h, \tilde{h}) \), the right-most element of \( \tilde{h}^i \) is the player \( j \)'s best response to \( h^j \), and the right-most element of \( \tilde{h}^j \) is the player \( i \)'s best response to \( h^i \).

Let for each pair \((h, \tilde{h})\), \( \tilde{h} \) being the follower of \( h \), \( M(h, \tilde{h}) \) denote the set of pairing patterns \( \mu(h, \tilde{h}) \) from the previous definition. The process whereby players choose their strategies may be described as a finite Markov chain on space \( \mathcal{H} \), with the probability of transition between \( h \) and \( \tilde{h} \) given by

\[
P^0_{h, \tilde{h}} = \begin{cases} 
\sum_{\mu \in M(h, \tilde{h})} P(\mu) & \text{if } \tilde{h} \text{ is the follower of } h; \\
0 & \text{if } \tilde{h} \text{ is not the follower of } h,
\end{cases}
\]

where \( P(\mu) \) is the probability of realization of the pairing pattern \( \mu \). To adhere to the terminology coined by Young (1993a) and Canning (1992) we call \( P^0 \) social adaptive play with memory \( m \).

3 Convergence of Social Adaptive Play without Mistakes

Social adaptive play is but a group variant of fictitious play with bounded recall. What differentiates social adaptive play from fictitious play is the randomness of interaction. While fictitious play has been studied extensively (see, for instance, Robinson, 1951; Shapley, 1964; Monderer and Shapley, 1996, or Young and Foster, 1998), social adaptive play has received much less attention. The only theoretical paper we know that is concerned with the long-run behavior of social adaptive play is by Canning (1992). Even Canning’s analysis, however, is silent about whether our variant of social adaptive play will converge to an equilibrium of stage game \( \Gamma \) or whether cycles occur ad perpetuum.\(^7\)

\(^7\)Canning shows merely that the frequency with which various population states are realized stabilizes if the process is recorded for a sufficiently long time.
To start the analysis of long-run behavior of social adaptive play, we shall prove that social adaptive play does not belong to the category of large-population models with payoff-positive dynamics (and hence it differs from the dynamics considered by KMR).

**CLAIM:** In social adaptive play it may happen that the proportion of players who use the best response to the (current) population strategy profile declines from period to period.

**Proof:** To prove the claim we shall construct a counter-example. Put \( m = 4 \), \( N = 6 \), and choose \( \Gamma \) as follows:

\[
\begin{array}{c|c|c}
\text{I} & \text{C} & \text{D} \\
\text{C} & 40, 40 & 0, 21 \\
\text{D} & 21, 0 & 21, 21 \\
\end{array}
\]

It is easy to check that for this \( \Gamma \) and \( m, n_c = 4 \) and \( n_d = 3 \). In other words, if a player’s history contains at least two \( D \)s, he chooses \( D \); otherwise, he chooses \( C \). Further, if three or more players play \( D \), then \( D \) is the unique best response to the population strategy profile.

Consider now the following situation:

\[
\begin{align*}
\text{h}_1(t) &= \text{h}_2(t) = \text{h}_3(t) = (D, D, C, C), \\
\text{h}_4(t) &= \text{h}_5(t) = \text{h}_6(t) = (C, C, C, C).
\end{align*}
\]

Given these histories, players 1 to 3 play at period \( t \) strategy \( D \), while players 4 to 6 play strategy \( C \). Thus \( D \) is the best response to the population profile. Suppose that the pairing pattern \( \{1, 6\}, \{2, 5\}, \{3, 4\} \) is realized at \( t \). Then, players’ histories at \( t + 1 \) will assume the form of

\[
\begin{align*}
\text{h}_1(t+1) &= \text{h}_2(t+1) = \text{h}_3(t+1) = (D, C, C, C); \\
\text{h}_4(t+1) &= \text{h}_5(t+1) = \text{h}_6(t+1) = (C, C, C, D).
\end{align*}
\]

Given these histories, all players will play \( C \) from \( t + 1 \) on. The share of \( D \)-players thus declined to zero from \( t \) to \( t + 1 \).\(^8\)

The intuition behind the claim is simple. If players do not yet cling firmly to one strategy (i.e., if their histories are such such that the better strategy is not

---

\(^8\)Note that the claim says nothing about whether the histories that allow the share of the best-response players to decrease can arise naturally throughout the process of learning. It may be the case that there exist \( T \in \mathbb{R} \) such that for any initial population history \( h(1) \) and any realized patterns of pairing at \( t = 1, 2, \ldots, T-1 \), from \( T \) on the social adaptive guarantees the payoff-positive dynamics. We do not think this is very likely, though.
much better than the worse one), a particularly unfortunate pairing pattern can make some players abandon the best response to the current population strategy profile. As a result, the number of the best-response players decreases. This is because players do not condition their behavior on the current state of the population but on their past encounters.

Denote by \( D \in H \) the population history composed of players’ histories that contain only \( Ds \) (\( D \) has been played as far as everybody remembers). Denote by \( C \in H \) the analogous population history for strategy \( C \). Both \( D \) and \( C \) are steady states of \( P^0 \): if each player’s history contains only \( Ds \), each player plays \( D \) in the present and will do so in the future; the analogue is true of strategy \( C \). Besides, no other steady states of \( P^0 \) apart from \( C \) and \( D \) exist. The argument for this goes as follows. Let \( h \in H \) be a steady state of \( P^0 \) different from both \( C \) and \( D \). If \( h \) compels all players to play \( D \), then in at most \( m \) periods the system will reach \( D \neq h \), which contradicts the assumption that \( h \) is a steady state. Hence, there has to be at least one player who—on the basis of his history in \( h \)—plays \( C \) and, for similar reason, at least one player who plays \( D \). With positive probability, these two players are paired for at least \( m \) periods in a row. Since both \( n_c \) and \( n_d \) are smaller than \( m \), at least one player must change his strategy during these \( m \) periods. But, if he does so, he deviates from his history in \( h \), and so \( h \) could not be a steady state.\(^9\)

The following lemma and theorem demonstrate that, irrespective of the initial condition, some steady state will be reached with probability one if the process evolves for a sufficiently long time. Due to random pairing and “contagion” among players (each player “infects” the history of his opponents with his strategy), the process embarks on a path toward a pure strategy equilibrium of the game \( \Gamma \). In both the following lemma and theorem we assume that the game \( \Gamma \) and the history length \( m \) are fixed. Since Lemma 1 is a standard result for “traditional” fictitious play,\(^10\) we relegated its proof to the Appendix.

**Lemma 1:** If \( N = 2, m > 1 \) and \( n_c > n_d \), the social adaptive play (in this case reduced to fictitious play) converges in at most \( 3m \) periods to one of the two steady states \( C \) or \( D \). Hence \( C \) and \( D \) are the only absorbing states of \( P^0 \).

Note that Lemma 1 relies on the assumption that \( n_c \neq n_d \). Relaxing this assumption would allow in the case of odd \( m \) for never-ending miscoordination

\(^9\)Similarly it may be proven that social adaptive play based on a more complex stage game \( \Gamma \) has no other steady states except for those in which all players play the same Nash equilibrium strategy (and have done so for at least \( m \) periods).

\(^10\)See, e.g., Ellison, 1997, p. 187 for essentially the same claim.
between players. An instructive example of such miscoordination sought for
a non-symmetric game is given by Young (1993a, p. 65).

**Theorem 1:** If \( N \geq 4, m > 1 \) and \( n_c > n_d \), the social adaptive play
converges with probability one to either \( C \) or \( D \). Hence \( C \) and \( D \) are the
only absorbing states of \( P^0 \).

**Proof:** We shall show that for any population history \( h \) there exist a positive
probability \( p \) and a natural number \( Q \), both dependent only on \( m, N \) and \( \Gamma \),
such that the probability of \( P^0 \) converging from \( h \) to either \( C \) or \( D \) in \( Q \) periods
is at least \( p \). Then, the probability of not converging to either \( C \) or \( D \) in \( TQ \)
periods is bounded above by \( (1 - p)^T \), which tends to 0 as \( T \) approaches infinity.
Depending on history \( h \) we shall examine several cases.

If on the basis of their histories in \( h \) all players play identical strategies, then
the convergence will occur for sure in at most \( m \) periods (\( Q = m \)). Suppose,
therefore, that there exists \( K_d \subset \{1, \ldots, N\} \), \( N > |K_d| \geq 2 \), a group of players
who play \( D \). Denote by \( K_c \) the complement to this group in \( \{1, \ldots, N\} \).

Suppose first that both groups contain an even number of players. Then,
with a positive probability players are for \( m \) periods in a row paired in such a
way that everybody meets only with members of his own group. After these
\( m \) periods, each group when considered separately reaches a steady state. Now
let for \( n_d + m \) periods members of \( K_d \) be paired with members of \( K_c \) and the
remaining players with their own kin; this surely happens with a positive prob-
bability. After the first \( n_d \) periods any member of \( K_c \) who has been paired with
a member \( K_d \) “converts” to strategy \( D \). In the next \( m \) periods his history will
become completely free of \( C \)s. As a result, after \( n_d + m \) periods the number
of \( D \)-players doubles, while the number of \( C \)-players shrinks accordingly. Pro-
ceeding in a similar manner one can show that \( D \) is reached at the latest in
\( Q = m + (n_d + m) + \cdots + (n_d + m) \) periods, where the number of bracketed
terms \( n \) is the smallest integer that makes \( 2^n |K_d| \) reach or exceed \( N \) (thus,
\( n \geq \log_2 \frac{N}{|K_d|} \)).

Suppose now that \( |K_d| = 1 \). Then, with a positive probability the member of
\( K_d \) meets with the same player for \( 3m \) periods in a row. As proven in Lemma 1,
after these \( 3m \) periods both players will play the same strategy. If this strategy
is \( C \), then the whole population has converged to \( C \) (because during these \( 3m \)
periods the histories of the other players who played \( C \) at the beginning have
undoubtedly been cleared of \( D \)s as well). Hence \( Q = 3m \). If this strategy is \( D \),
\( |K_d| \) has increased to 2, and the proof continues as above. In this case, \( Q \) is at
most \( 3m + m + \lceil \log_2 \frac{N}{2} \rceil (n_d + m) \).
Suppose finally that \(|K_d|\) is odd, \(|K_d| \geq 2\). Then, with a positive probability players are paired for \(3m\) periods in a row so that everybody who can is paired with his kin and the “odd-men-out” of both groups are paired with each other. As proven in Lemma 1, after these \(3m\) periods the “odd-men-out” will be playing the same strategy and hence both \(K_d\) and \(K_c\) will have an even number of members. From what we proved above for this case it follows that \(Q\) is at most \(3m + m + \lfloor \log_2 \frac{N}{|K_d|^{-1}} \rfloor (n_d + m)\).

Theorem 2 says nothing about the case of \(m = 1\). This case, however, is of little interest for it implicitly violates the assumption that \(n_c \neq n_d\). Besides, when \(m = 1\) both the number of \(C\)-players and number of \(D\)-players stays constant: if players of identical strategies meet, neither of them changes his strategy; if players of different strategies meet, they “convert” each other. Hence no convergence can occur.

4 Social Adaptive Play with Mistakes

If players make occasional mistakes in their strategy choice, Theorem 1 holds no longer because the perturbations move the process constantly away from the steady states. As we shall show, however, a small probability of mistakes guarantees that the process will spend most of the time in one steady state (a stochastically stable state in the sense of Young, 1993b).

The way we model players’ mistakes follows closely Young (1993a). Fix \(m\) and \(\Gamma\). Assume that at any period \(t\), with probability \(\lambda_i \varepsilon\) player \(i\) makes a mistake: instead of choosing the best response to his history \(h^i(t)\), he chooses \(s^i(t)\) at random, drawing it from the probability distribution \(q_i\). Assume that for each \(i\), \(q_i\) places a positive probability on both \(C\) and \(D\). Assume further that players’ mistakes are independent.\(^{11}\)

In the expression for total probability of player \(i\)’s mistake, \(\lambda_i\) describes the idiosyncratic propensity of this player to err, while \(\varepsilon\) captures a general, whole-population tendency toward erroneous strategy choices. A priori knowledge of \(\lambda_i\) is, of course, hard to justify. Our result, however, does not depend on a particular pattern of \(\lambda_i\)’s as far as they are constant and independent of \(\varepsilon\).

Let us now characterize the perturbed process \(P^\varepsilon\). Fix the population history \(\mathbf{h}\) and assume that only a subset of players \(J \subseteq \{1, \ldots, N\}, J \neq \emptyset\), make a mistake.

\(^{11}\)As long as players’ mistakes are independent, one can allow for \(q_i\) to be conditional on \(h^i(t)\). We refrain from doing so to simplify the notation.
**J-FOLLOWER:** We call population history $\widehat{h} = (\widehat{h}^1, \widehat{h}^2, \ldots, \widehat{h}^N)$ the *J*-follower of population history $h = (h^1, h^2, \ldots, h^N)$ if

(i) Each $\widehat{h}^i$ is the successor of $h^i$.

(ii) There exists a pairing pattern $\mu_J(h, \widehat{h})$ with the following property.

Let $\Upsilon(\mu) \subseteq \{1, \ldots, N\}$ denote the set of players that are paired with players from $J$ when the pairing pattern is $\mu_J(h, \widehat{h})$. Then, for every pair of players $\{i, j\} \in \mu_J(h, \widehat{h})$ such that $i \notin \Upsilon(\mu)$, the right-most element of $\widehat{h}^i$ is the player $j$’s best response to $h^j$.

Let $M_J(h, \widehat{h})$ denote the set of $\mu_J(h, \widehat{h})$ from the previous definition. Then, the transition probability from $h$ to $\widehat{h}$, conditional on players from $J$ making mistakes and nobody else doing so, can be expressed as

$$Q^J_{h, \widehat{h}} = \begin{cases} \sum_{\mu \in M_J(h, \widehat{h})} P(\mu) \prod_{\{i, j\} \in \mu} q_j(\hat{s}^i) & \text{if } \widehat{h} \text{ is the } J\text{-follower of } h \text{ and for all } i \notin \Upsilon(\mu) \\
0 & \text{if } \widehat{h} \text{ is not the } J\text{-follower of } h. \end{cases}$$

For any $J \subseteq \{1, \ldots, N\}$, $J \neq \emptyset$, the probability of players from $J$ making mistakes and nobody else doing so is given by $\prod_{j \in J} \lambda_j \prod_{j \notin J} (1 - \lambda_j \varepsilon)$. Hence, the transition probabilities of the perturbed process $P^\varepsilon$ are

$$P^\varepsilon_{h, \widehat{h}} = \left( \prod_{i \in \{1, \ldots, N\} \atop i \notin J} (1 - \lambda_i \varepsilon) \right) P^0_{h, \widehat{h}} + \sum_{J \subseteq \{1, \ldots, N\} \atop J \neq \emptyset} \left( \prod_{j \in J} \lambda_j \varepsilon \right) \left( \prod_{j \notin J} (1 - \lambda_j \varepsilon) \right) Q^J_{h, \widehat{h}}. \tag{6}$$

We call $P^\varepsilon$ *perturbed social adaptive play* with memory $m$, probabilities of mistakes $\lambda_i \varepsilon$ and distributions of mistakes $q_i(\cdot)$.

## 5 Convergence of Perturbed Social Adaptive Play

Little is known about the long-run behavior of perturbed social adaptive play. For learning dynamics of which social adaptive play is a special case, Canning (1992) proved that the invariant distribution of the perturbed process approaches the invariant distribution of the unperturbed process as the probability of mistakes tends to zero. Canning also constructed an example of learning dynamics similar to social adaptive play in which all players react to idiosyncratic histories but play against the whole population. In this example, he showed that the state in which all players play the RDE is stochastically stable. Also simulations run by Di Gioacchino (1992) point toward the stochastic stability of the RDE in social adaptive play.

\footnote{If $J = \{1, \ldots, N\}$, we take $(\prod_{j \in J} (1 - \lambda_j \varepsilon))$ as equal to 1.}
To analyze the stochastic stability of social adaptive play formally, note first that for any \( \varepsilon > 0 \), \( P^\varepsilon \) is irreducible because with a sufficient number of mistakes (\( mN \), at most) any population history \( \tilde{h} \) can be reached from any \( h \) at the latest in \( m \) periods. It is also aperiodic because the transition from \( h \) to \( \tilde{h} \) can be carried out exactly in \( m \) periods as well as exactly in \( m + 1 \) periods. Hence \( P^\varepsilon \) has a unique stationary distribution \( \mu^\varepsilon \), \( \mu^\varepsilon P^\varepsilon = \mu^\varepsilon \), and for any \( h \in \mathcal{H} \), \( \mu^\varepsilon_h \) is a relative frequency with which \( h \) will be observed if social adaptive play goes on for long enough time.

The following definition is from Young (1993a).

**STOCHASTICALLY STABLE HISTORY:** A population history \( h \in \mathcal{H} \) is stochastically stable if \( \mu^* = \lim_{\varepsilon \to 0^+} \mu^\varepsilon_h \) exists and is positive. It is strongly stochastically stable if \( \mu^* \) equals one.

To characterize stochastically stable histories we shall, as customary in the recent evolutionary literature, take advantage of the concept of directed graphs (\( h \)-trees) introduced by Friedlin and Wentzell (1984).

**\( h \)-TREE:** \( h \)-tree \( \omega_h \) is a collection of ordered pairs of elements of \( \mathcal{H} \)—"arrows"—\((a \to b)\), such that from every \( \tilde{h} \in \mathcal{H} \), \( \tilde{h} \neq h \), there exists a unique directed "path" \( d_{\tilde{h}, h} \) to \( h \), \( d_{\tilde{h}, h} = (\tilde{h} \to h_1)(h_1 \to h_2) \ldots (h_k \to h) \), and there are no cycles.

Denote the set of \( h \)-trees as \( \Omega_h \).

**RESISTANCE:** For any pair of (distinct) population histories \( a, b \in \mathcal{H} \) define the resistance \( r(a, b) \) as the minimum number of players’ mistakes necessary for transition from \( a \) to \( b \). If not even \( N \) mistakes suffice, put \( r(a, b) = +\infty \). For any directed path \( d_{a, b} = (a \to h_1)(h_1 \to h_2) \ldots (h_k \to b) \) define the resistance \( r(d_{a, b}) \) as the sum of the resistances of constituent “arrows.” For any \( h \)-tree \( \omega_h \) we define the resistance \( r(\omega_h) \) analogously.

Note that resistance is independent of \( \lambda_i \)'s as well as of particular distributions of mistakes \( q_i \)'s.

**STOCHASTIC POTENTIAL:** Stochastic potential of the population history \( h \) is the resistance of the least-resistance \( h \)-tree:

\[
v(h) = \min_{\omega \in \Omega_h} r(\omega).
\]

\[\text{Formally, } r(a, b) = j \text{ if there exist } J \subseteq \{1, \ldots, N\}, |J| = j, \text{ such that } b \text{ is a } J\text{-follower of } a \text{ and if for all } K \subset \{1, \ldots, N\}, |K| < j, \text{ b is not a } K\text{-follower of a.}\]
From formula (6) we know that for each pair \( h, \tilde{h} \in H \), \( P_{h,\tilde{h}}^{\varepsilon} \) is a polynomial in \( \varepsilon \), and that \( \lim_{\varepsilon \to 0^+} P_{h,\tilde{h}}^{\varepsilon} = P_{h,\tilde{h}}^{0} \). Thus \( P^{\varepsilon} \) is a regular perturbation of \( P^{0} \) in the sense of Young (1993a). Since by Theorem 1 \( P^{0} \) has only two recurrent sets (absorbing states), \( C \) and \( D \), Theorem 4 by Young (1993a, p. 78) yields

**Theorem 2**: \( \mu^* = \lim_{\varepsilon \to 0^+} \mu^{\varepsilon} \) is well-defined (it exists and is unique) and it is a stationary distribution of \( P^{0} \). The only candidates for stochastically stable histories are absorbing states \( C \) and \( D \). If \( v(C) < v(D) \), \( C \) is strongly stochastically stable. If \( v(C) > v(D) \), \( D \) is strongly stochastically stable. Finally, if \( v(C) = v(D) \), both \( C \) and \( D \) are stochastically stable. Whatever the case the property of being stochastically stable is independent of \( \lambda_i \)'s and \( q_i \)'s.

We shall now investigate which of \( v(C) \) and \( v(D) \) is smaller. Since \( C \) is an absorbing state of \( P^{0} \), every directed path from \( D \) to \( C \) can be extended to a full \( C \)-tree by adding arrows of zero resistance. Besides, every \( C \)-tree must, by definition, contain some path from \( D \) to \( C \). Thus, the least-resistance \( C \)-tree must have exactly the same resistance as the least-resistance path from \( D \) to \( C \), the analogue being true for the least-resistance \( D \)-tree. Formally, this is stated in the following Lemma.

**Lemma 2**: The stochastic potential \( v(C) \) can be determined as the minimum number of mistakes needed for the transition from \( D \) to \( C \). The stochastic potential \( v(D) \) can be determined as the minimum number of mistakes needed for the transition from \( C \) to \( D \).

Theorems 3 and 4 below contain the gist of the paper. They specify values of \( v(C) \) and \( v(D) \) given that \( n_c > n_d \).

The following terminology is used throughout the proofs of the theorems and auxiliary lemmata (some of which are in the Appendix). A player is called **contagious**, if he can meet a different-strategy player without changing his own strategy.\(^{14}\) A player makes an **error** if he plays a strategy different from the best-response to his history. Note that each error is a mistake but not vice-versa because a player may hit the best response even by mistake.

**Theorem 3**: Suppose that \( N > 2, m > 1 \) and \( n_c = n_d + k, k \geq 1 \). Then, it holds that

\[
v(D) = n_d. \tag{7}\]

\(^{14}\)This meaning of contagious may seem quite far away from its everyday meaning, i.e. “catching,” or “exciting similar emotions or conduct in others.” The point is that a contagious player is capable of propagating his own strategy without becoming ‘infected by’ (‘converted to’ in our terminology) his opponent’s strategy.
**Proof:** By Lemma 2 we have to show that the minimum number of mistakes necessary for a transition from $C$ to $D$ is $n_d$. We first show that as soon as the first two players are “converted away” from playing $C$, no more mistakes are needed to “convert” other players. Label the first converts-to-be as Cyril and Methodius.\(^{15}\) Suppose that Cyril and Methodius are first paired with each other for $m$ periods and then each with a $C$-player for $n_d$ periods. After the first $m$ periods Cyril’s and Methodius’s contagiousness increases since their histories become net of $C$. After the next $n_d$ periods Cyril and Methodius will still play $D$ since only $m - n_d < n_c$ $C$s have entered their histories. Also Cyril’s and Methodius’s opponents will play $D$ since $n_d$ $D$s have entered their histories. It follows that if Cyril and Methodius meet often with each other (to reinforce their strategy choice) and only sometimes with other $C$-players (to proselytize their strategy choice), the number of $D$-players shall increase, reaching eventually the whole population (see also proof of Theorem 1 for a more detailed discussion on “infectiousness” of strategy $D$).

Label now the first two players to be converted away from playing $C$ as 1 and 2. The minimum number of mistakes players 1 and 2 must make to change their strategy is the same as the minimum number of errors required for such a change: a mistake that results in playing the best response to a player’s history might well not have been made and nothing would change. Besides, players 1 and 2 benefit from their errors most if they meet only with each other: meetings with other players would only introduce unnecessary $C$s into their histories.\(^{16}\)

Set $t = 1$ and consider

$$h^1(1) = h^2(1) = (C, \ldots, C).$$

Obviously, fewer than $n_d$ errors, even if they are committed by the same player (say player 1), do not suffice to convert player 2 to strategy $D$. Suppose, therefore, that player 1 makes an error throughout periods 1, $\ldots$, $n_d$. Then, at

\(^{15}\)Cyril (or Constantine, 827–869 A.D.) and Methodius (826–885 A.D.) were two Greek missionaries, brothers, called “apostles to the Slavs” and “fathers of Slavonic literature.” After having spread Christianity in what is today Ukraine, they traveled in 863 A.D. to Moravia, a part of the present-day Czech Republic. Their mission in Moravia was a response to the invitation of Prince Rastislav, who sought missionaries able to preach in the Slavonic vernacular and thereby check German influence in Moravia. Cyril is considered the creator of the Cyrillic alphabet, still used in Russia, Bulgaria, Yugoslavia, Macedonia, Bosnia, Belorus and Ukraine.

\(^{16}\)Similar logic will be used in ensuing proofs.
\[ t = n_d + 1 \] players’ histories will be

\[
h^1(n_d + 1) = (C, \ldots, C, C, \ldots, C),
\]

\[
h^2(n_d + 1) = (C, \ldots, C, D, \ldots, D),
\]

Hence from \( t = n_d + 1 \) on, player 2 will play strategy \( D \). After the next \( n_d \) periods the players’ histories will become

\[
h^1(2n_d + 1) = (C, \ldots, C, C, \ldots, C),
\]

\[
h^2(2n_d + 1) = (C, \ldots, C, D, \ldots, D, C, \ldots, C).\]

Hence from \( t = 2n_d + 1 \) on, both players will play \( D \).

Computing \( v(C) \)—i.e., determining a number of mistakes necessary to move the whole population \( D \) to \( C \)—is a more complicated task. A player who plays \( D \) will not change his strategy easily as a result of his meetings with \( C \)-players, because the condition \( n_c > n_d \) makes strategy \( C \) less “infectious.”

Instead of proving the analogue to Theorem 3 directly, we shall proceed in steps. We shall examine how many mistakes are necessary for \( D \)-to-\( C \) conversion of the first two players, the next two players, another pair of players, and so forth.\(^{18}\) The following lemmata correspond to these steps, respectively.

**Lemma 4:** Suppose that \( N > 2, m > 1 \) and \( n_c = n_d + k, k \geq 1 \). Let at some time \( t \geq 1 \) the population history be \( D \). Then, at least \( n_c + k + 1 \) errors are necessary to convert the first two players (say players 1 and 2) to strategy \( C \).

**Proof:** Note first that \( n_c = n_d + k \) implies that \( m = 2n_c - k - 1 \). If player 1 makes \( n_c \) errors in a row, then from \( t = n_c + 1 \) on player 2 will play \( C \). At

\(^{17}\) Note that \( m - 2n_d \geq 0 \) because \( 2n_d < n_d + n_c = m + 1 \).

\(^{18}\) Note that proceeding by pairs is generally the “best” we can do (it surely requires fewer mistakes than other ways). If no player has yet been converted to \( C \), then the best for the first convert-to-be is to be paired repeatedly with the same player. Hence the conversion of the first two players goes hand in hand. Now suppose there is an even group of converts (\( C \) players) in a heretic world of \( D \)-players. It is surely better if the converts concentrate in their proselytizing on as few “heretics” as possible; the danger of “falling away” from strategy \( C \) is acute. On the other hand, they cannot really concentrate on a single heretic only, for there is an even number of them and the interaction takes place in pairs. See the proofs of Lemmata 4-6 for details.
If nobody makes an error at $t = m + 1$, then player 2 will get converted back to strategy $D$ at $t = m + 2$. To prevent this, we need player 2 to make errors from $t = m + 1$ on and to make them enough to convert player 1 to $C$ too, i.e., at least $k + 1$. If this happens, then at $t = m + k + 2$ the players’ histories will be

\[
h^1(m + k + 2) = (D, \ldots, D, C, \ldots, C) \rightarrow C,
\]
\[
h^2(m + k + 2) = (C, \ldots, C, D, \ldots, D) \rightarrow C.
\]

Hence from $t = m + k + 2$ on, both players will play $C$. This completes the proof. It is worth noting that the obvious way of “conversion” whereby players make $n_c$ errors throughout $t = 1, \ldots, n_c$ requires more errors than the described way because $2n_c \geq n_c + k + 1$ (this follows from $n_c \geq 1$).

**Lemma 5:** Suppose that $N > 2$, $m > 1$ and $n_c = n_d + k$, $k \geq 1$. Let that at some time $t \geq 1$ the histories of players 1 and 2 contain only $C$s, while the histories of all other players contain only $D$s. Then, at least $(k+1) + 1 \left( \frac{3}{2}(k + 1) > n_c \right) [3(k + 1) - 2n_c]$ errors are necessary to convert two more players (say players 3 and 4) to strategy $C$.

**Lemma 6:** Suppose that $N > 2$, $m > 1$ and $n_c = n_d + k$, $k \geq 1$. Let $s \in \mathbb{N}$, $\frac{N}{2} > s \geq 2$ be given, and let at some time $t \geq 1$ the histories of players 1, \ldots, $2s$ contain only $C$s, while the histories of all other players contain only $D$s. Then, at least $v(s)$ errors are necessary to convert two more players (say players $s + 1$ and $s + 2$) to strategy $C$, where $v(s)$ is given as

\[
v(s) = \mathbf{1} \left( \frac{s + 1}{2s} \right)^k \left( k + 1 \right > n_c \right) \left\{ \left( s(k + 1) - (s - 1)n_c \right) + \right.\]
\[
+ \left. \mathbf{1} \left( \frac{2s + 1}{2s} \right) \left( k + 1 > n_c \right) \left( (2s + 1)(k + 1) - 2sn_c \right) \right\}.
\]

As the proofs of both lemmata are very similar to that of Lemma 4, we relegated them to the Appendix.

With Lemmata 4, 5 and 6 at hand we are ready to compute the value of $v(C)$.  

---

19 The arrows and letters that follow the histories indicate what strategy each player chooses on the basis of a particular history, if he does not make a mistake.

20 Here and hereafter $\mathbf{1}(\cdot)$ stands for the indicator function.
THEOREM 4: Suppose that \( N > 2 \), \( m > 1 \) and \( n_c = n_d + k \), \( k \geq 1 \). Set \( n = \frac{N}{2} \). Then, it holds that

\[
v(C) = (n_c + k + 1) + (k + 1) + 1 \left( \frac{3}{2}(k + 1) > n_c \right) [3(k + 1) - 2n_c] + \\
\sum_{s=2}^{n-1} 1 \left( \frac{s}{s-1}(k + 1) > n_c \right) \left\{ (s(k + 1) - (s-1)n_c) + \\
+ 1 \left( \frac{2s+1}{2s} (k + 1) > n_c \right) [(2s+1)(k + 1) - 2sn_c] \right\}.
\]

By Theorems 3 and 4 \( v(D) \) and \( v(C) \) depend only on \( n_c \) and \( n_d \); if \( n_c > n_d \) then \( v(C) > v(D) \). By Theorem 2, \( v(C) > v(D) \) implies the strong stochastic stability of \( D \). Since the inequality \( n_c > n_d \) follows directly from the risk-dominance of \((D, D)\), we have the following:

COROLLARY 1: Suppose that \( N > 2 \) and \( m \) is large enough. Then, the unique (strongly) stochastically stable population history is the history in which all pairs play the RDE of game \( \Gamma \).\(^{21}\)

6 Discussion

The (strong) stochastic stability of the RDE,\(^{22}\) concluded in Corollary 1, is qualitatively the same result as obtained by KMR (1993) and Young (1993a). But is it quantitatively the same?

A natural way to assess the “strength” of stochastic stability of the RDE is to look at the difference between the number of mistakes required for the transition away from and toward the RDE. In our notation this measure reads as \( v(C) - v(D) \). In KMR’s and Young’s models, the number of mistakes required for the transition toward each equilibrium depended linearly on the loss of deviating (LD) from that equilibrium. The difference between the stochastic potentials of both equilibria was therefore pro rata related to the difference between their LDs; in our notation, \( v(C) - v(D) \approx k \). In our model, as follows from Theorems 3 and 4, \( v(C) - v(D) > k \). Besides, \( v(C) - v(D) \) generally increases with the population size since \( v(C) \)—the number of mistakes required for the transition toward the PDE—increases as the population grows while \( v(D) \)—the number

\(^{21}\)The need for large \( m \) arises from the rounding-off in (3). If the difference between LD from \((C, C)\) and LD from \((D, D)\) is negligible, small \( m \) may not suffice to have this difference manifested in the difference between \( n_c \) and \( n_d \).

\(^{22}\)Speaking about stochastic stability of an equilibrium is imprecise; precision dictates speaking about stochastic stability of a population history in which all players play a certain equilibrium. Since chances of confusion are small (and since the precise phrasing is unbearably long-winded), we allow ourselves here and henceforth to be imprecise.
of mistakes required for the transition toward the RDE—stays constant. Hence in our model stochastic stability of the RDE is stronger than in KMR’s and Young’s models. What is the intuition behind this observation?

In KMR’s model, the whole population moved from the RDE to the PDE as soon as at least one player found such a switch in strategy advantageous. If one player switched, so did the others because all players conditioned their strategy choice on the same information set (a population strategy profile). Likewise, in Young’s model all players could move from the RDE to the PDE as soon as the first pair of players coordinated on the PDE. Once this happened, there was a positive probability that the next pair of players would sample (almost) the same history as their immediate predecessors and hence follow in their footsteps.

In our model the situation is dramatically different. Even if one pair of players moves from the RDE to the PDE as a result of either player’s mistakes, more mistakes are necessary for the same conversion of other players (see proofs of Lemmata 5 and 6). A pair of players that plays the PDE strategy in the midst of RDE-strategy players is constantly in danger of being “converted back” to the RDE strategy. The larger the population, the more acute this danger. Thus, the number of mistakes required for the transition from the RDE toward the PDE grows as the population increases.23 In contrast to this, the number of mistakes required for the reverse transition does not depend on the population size. The reason is the “infectiousness” of the RDE-strategy: given enough luck in random pairing, a single pair of RDE-strategy players is capable of propagating this strategy further, eventually ‘infecting’ a population of any size.

The qualitative accord between our result and that of Young (1993a) suggests that stochastic stability of the RDE is robust to the details of learning from past experience. It is very likely that RDE will be stochastically stable even in a model where players learn from both their own experience and their fellows’ experience. In fact, the way we proved Theorems 3 and 4 points to an even more general conclusion: if a stage game is a $2 \times 2$ coordination game, then RDE will be strongly stochastically stable whenever (i) each player’s decision can be modelled as his best-response to a finite, perhaps idiosyncratic history, and (ii) all players’ histories reflect in some way the current or recent population

---

23This number is not, however, an unbounded function of the population size. Whatever the difference between the LDs of both equilibria, there exists a population size beyond which this number stays constant [this size equals $2s$, where $s$ is the smallest integer such that $\frac{k + 1}{\pi} \leq n_0$; see formula (8)]. The reason is that however much risk-dominated the PDE strategy is, there always exist a minimum number of players who, if they adopt this strategy, can as a group spread this strategy further in a population of any size.
strategy profile (and hence are updated as the population evolves).\footnote{Implicit in this claim is the assumption that players make mistakes with constant probability, independent of the current population strategy profile. See Bergin and Lipman (1996) for an analysis of what happens if this is not the case. The claim may also fail if other factors, such as for example group selection, enter the model.}

It follows from the proofs of Theorems 3 and 4 that the easiest way to upset a coordinated-on equilibrium is not by simultaneous mistakes of many players, but by sequential mistakes of one or a few players. This is a direct consequence of learning from past experience; the same is true about Young’s model but not about KMR’s model. To interpret a sequence of mistakes made by one player as a sequence of purely unintentional erroneous acts is unconvincing. A question arises whether a player cannot “make mistakes” on purpose. Is it not possible that what myopic players perceive a mistake—since it is at odds with their decision rule—is an intentional deed of a far more for-sighted player who, perhaps, sees the benefits of coordinating on a different equilibrium and tries to manipulate the population toward it? The problem is whether such a hyper-rational player can find the manipulation beneficial and thus indeed make “a mistake.” Ellison (1997) addresses this problem in detail. He concludes that a single hyper-rational player in a population of myopic players can only find it profitable to (attempt to) manipulate the population toward the RDE. The expected time for the reverse transition is so long that the manipulation does not pay, even if the risk-dominated equilibrium is the PDE. This is in perfect accord with our finding that the transition toward the RDE requires far fewer mistakes than the reverse one.

7 Conclusion

We developed a large-population learning model in the spirit of KMR (1993) and Young (1993a) based on learning from personal experience. Our aim was, first, to complement existing evolutionary studies of other types of learning based on the exchange of information among players, and, second, to see what learning from personal experience has to say about selection among equilibria.

We considered a large population of players that are paired repeatedly at random to play a $2 \times 2$ symmetric coordination game. Each player chooses his strategy as the best response to a collection of his own recent plays (a history). Every so often, players make mistakes.

We showed that the unperturbed learning dynamic converges to a state where all players coordinate on the same Nash equilibrium. From these equilibria the perturbed learning dynamics selects the RDE; in other words, a strongly
stochastically stable state is the one in which all players coordinate on the RDE. This is true for any extent of the risk-dominance of the RDE as long as players’ memory is sufficiently long. This accords with findings of KMR and Young who also establish stochastic stability of the RDE. In our model, however, stochastic stability of the RDE is more pronounced because learning from personal experience makes the transition from the RDE to the PDE far more difficult than other modes of learning.

Since the RDE is stochastically stable under both learning-from-personal-experience dynamics (our model) and learning-from-common-experience dynamics (Young, 1993a), there is good reason to believe it will be so under learning dynamics between the two, i.e., dynamics based on a combination of one’s personal experience and one’s fellows’ experience. In fact, the nature of our proofs points toward a more general conclusion: the RDE will be stochastically stable in any model where players react to finite samples if these reflect in some way the current or recent population strategy profile.

The more pronounced stochastic stability of the RDE in our model (as compared to KMR, 1993, and Young, 1993a) grants our model apparently greater predictive power. The number of mistakes necessary for the transition from the RDE toward the PDE is so high—and the pattern of these mistakes and players’ encounters so specific—that this transition is almost impossible. This is true even if the probability of a mistake is far from negligible. One can therefore be tempted to argue that as soon as the population of “Doubting Thomases” coordinates on the RDE, this equilibrium will become almost inescapable. In contrast, coordination on the PDE seems quite fragile.

Some caution is, however, necessary. While the transition toward the PDE is very improbable, the transition toward the RDE is not very probable either. Unless the difference between the security levels of both equilibria is great and players’ memory short, one cannot expect that swings between equilibria will happen frequently. Even more than in KMR’s model it is true that players will converge to the RDE only in the very long run.25

Perhaps the most general implication of our model is that learning from personal experience has great “inertia”. Once the population of “doubting Thomases” coordinates on an equilibrium, this equilibrium will be very difficult to upset (cf. Ellison, 1997). In the intermediate run, different equilibria can be in vogue concurrently.

25Introducing local interaction into the model may speed up the convergence considerably (see Ellison, 1993).
Appendix

Proof of Lemma 1: If at \( t = 1 \) both players play the same strategy, they will continue doing so in the future and an absorbing state will be reached in at most \( m \) periods. Suppose, therefore, that at \( t = 1 \) each player plays a different strategy: say player 1 strategy \( D \) and player 2 strategy \( C \). Examine players' histories after the first \( m \) periods. One of two things must have happened: either at some period \( t \leq m \) one player changed his strategy while the other did not or no such \( t \) exists. In the former case both players started to play the same strategy at \( t + 1 \) and thus an absorbing state will be reached in at most \( t + m < 3m \) periods.

Examine now the latter case. Note first that some strategy change must have occurred during the first \( m \) periods. Why? Suppose that neither player has changed his strategy during the first \( m - 1 \) periods. Then player 2 will change his strategy in the \( m \)-th period because at \( t = m - 1 \) his history contains \( m - 1 \) \( D \)s and \( m - 1 \geq n_d \). Since we have ruled out the case when only one player changed his strategy at some \( t \leq m \), it must be that at any \( t \leq m \) either both players change their strategies or neither of them does. Hence, at \( t = m + 1 \) the histories of both players must be “complementary”: for all \( \tau \leq m \), \( s_1(\tau) = D \) if and only if \( s_2(\tau) = C \). The following diagram displays an example of such histories for \( m = 5 \):

\[
\begin{align*}
h_1(5 + 1) &= (C, C, D, C, D) \\
h_2(5 + 1) &= (D, D, C, D, C).
\end{align*}
\]

By assumption, at \( t = m + 1 \) each player still plays a different strategy (but possibly other than the strategy he played at \( t = 1 \)), e.g., player 1 strategy \( C \), player 2 strategy \( D \). For any \( t \geq m + 1 \), denote by \( d(t) \) the number of \( D \)s in \( h_1(t) \) and by \( c(t) \) the number of \( C \)s in \( h_2(t) \). Clearly, \( n_d > d(m + 1) = c(m + 1) \).

Examine now, period by period, how the values of \( d(t) \) and \( c(t) \) change as players continue playing. If \( s_1(1) \) is \( D \), then \( d(m + 2) = d(m + 1) \) because \( h_1(m + 2) \) will be made by dropping the left-most \( D \) from \( h_1(m + 1) \) and adjoining a new \( D \) to its right end. If, however, \( s_1(1) \) is \( C \), then \( d(m + 1) = d(m) + 1 \). Due to the “complementariness” of \( h_1(m + 1) \) and \( h_2(m + 1) \), \( d(m + 1) = c(m + 1) \) in both cases. It follows that \( d(t) \) and \( c(t) \) are identical non-decreasing functions of \( t \) as long as both players keep playing the same strategies as they did at \( t = m + 1 \).

Now note that player 2 will keep playing strategy \( D \) at least as long as player 1 strategy \( C \), because the number of \( C \)s in the history of player 2 grows exactly as fast as the number of \( D \)s in the history of player 1, and because \( n_d < n_c \). At the same time, player 1 cannot keep playing \( C \) until \( t = 2m \) since then

\[
d(2m + 1) = d(m + 1) + \{ \text{number of } C \text{s in } h_1(m + 1) \} \geq d(m + 1) + n_c > n_d,
\]
which is a contradiction. Hence, player 1 must change his strategy from $C$ to $D$ at some point $t^*$, $t^* \leq 2m$; this implies that $d(t^*) = n_d$. Since, as argued, player 2 cannot change his strategy before player 1, it must be that $c(t^*) = d(t^*) = n_d < n_c$. Hence both players will play $D$ from $t^*$ on and state $D$ will be reached at most in $t^* + m \leq 3m$ periods.

Proof of Lemma 5: As before, $m = 2n_c - k - 1$. Consider first the case when players 3 and 4 make errors only when paired with each other. Reset time so that $t = 1$. Let at $t = 1, \ldots, n_c - k - 1$ player 1 meet with player 3, and player 2 with player 4. Then, at $t = n_c - k$ players’ histories will be

$$
\begin{align*}
    &h^1(n_c - k) = h^2(n_c - k) = (C, \ldots, C, D, \ldots, D) \rightarrow C, \\
    &h^3(n_c - k) = h^4(n_c - k) = (D, \ldots, D, C, \ldots, C) \rightarrow D.
\end{align*}
$$

Players 1 and 2 are no longer contagious at $t = n_c - k$ and therefore cannot be paired with players 3 and 4 until (i) they become contagious again by repeated meetings with each other or (ii) both players 3 and 4 play strategy $C$.\(^{26}\) The former requires that players 1 and 2 meet with each other until $t = m$, thus becoming contagious at $t = m + 1$. If the latter is to happen before $t = m + 1$, players 3 and 4 would have to make at least $(k + 1) + [(k + 1) - (n_c - k - 1)] = 2k + 2 - n_c$ errors; this, as we are going to demonstrate, is not the least number of errors required for the conversion of players 3 and 4.

If players 1 and 2 meet with each other until $t = m$, then also players 3 and 4 must meet with each other until $t = m$.\(^{27}\) Suppose first that players 3 and 4 make so few errors between $t = n_c - k$ and $t = m$ that at $t = m + 1$ they still play strategy $D$. Then, players 1 and 2 will only be able to meet with players 3 and 4 from $t = m + 1$ to $t = m + (n_c - k - 1)$ because a period later they will be contagious no longer. At the same time, these meetings will not help in converting players 3 and 4 to strategy $C$ because for every $C$ adjoined to the right end of $h^3$ or $h^4$ one $C$ will be deleted from the left end. Hence players 3 and 4 will both still play $D$ at $t = m + (n_c - k)$. Further, $h^3$ and $h^4$ will at $t = m + (n_c - k)$ look almost the same as they did at $t = n_c - k$, except, perhaps, for a few $Cs$ among the $n_c$ leftmost strategies (these $Cs$ reflect the errors made from $t = n_c - k$ to $t = m$). As for $h^1$ and $h^2$ at $t = m + (n_c - k)$, they will look exactly the same as they did at $t = n_c - k$. History thus will, in a way, repeat itself: players 1 and 2 will be no more contagious at $t = m + (n_c - k)$ and will

\(^{26}\) If only, say, player 3 plays strategy $C$ and players 1 and 2 are not contagious, then whichever of them is paired with player 4 will be converted back to strategy $D$.

\(^{27}\) See the argument in the main text preceding footnote 16.
need to meet with each other until \( t = 2m \) or until both players 3 and 4 play strategy \( C \). At the same time, errors that players 3 and 4 will make between \( t = n_c - k \) and \( t = m \) because all \( C \)s that entered \( h^3 \) and \( h^4 \) before \( t = m \) will disappear before \( t = 2m \). It follows that whatever result can be achieved by errors that players 3 and 4 make between \( t = m + (n_c - k) \) and \( t = 2m \) could have been achieved with fewer errors if these were made between \( t = n_c - k \) and \( t = m \).

Suppose now that player 4 makes \( k + 1 \) errors between \( t = n_c - k \) and \( t = m \) so that player 3 starts to play strategy \( C \) before \( t = m + 1 \). Suppose first that these errors are made throughout \( t = n_c - k, \ldots, m \). Then, players’ histories at \( t = n_c + 1 \) will be

\[
\begin{align*}
  h^1(n_c + 1) &= (C, \ldots, C, D, \ldots, D, C, \ldots, C) \to C; \\
  h^2(n_c + 1) &= (D, \ldots, D, C, \ldots, C, C, \ldots, C) \to C; \\
  h^3(n_c + 1) &= (D, \ldots, D, C, \ldots, D, C, \ldots, D) \to D.
\end{align*}
\]

At \( t = n_c + 1 \) only player 3 is contagious and will stay so until \( t = m \). If throughout \( t = n_c + 1, \ldots, m \) the pairing pattern is \( \{1, 2\}, \{3, 4\} \), then at \( t = m + 1 \) players’ histories will be

\[
\begin{align*}
  h^1(m + 1) &= (D, \ldots, D, C, \ldots, C) \to C; \\
  h^2(m + 1) &= (D, \ldots, D, C, \ldots, D) \to C; \\
  h^3(m + 1) &= (C, \ldots, C, D, \ldots, D, C, \ldots, C) \to 0.
\end{align*}
\]

If \( 2(n_c - k - 1) \geq n_c \) (or, equivalently, if \( n_c \geq 2(k + 1) \)), player 4 plays \( C \) at \( m + 1 \) and the proof is finished. Suppose that \( n_c < 2(k + 1) \). At \( t = m + 1 \) both players 1 and 2 are contagious and will stay so until \( t = m + (n_c - k - 1) \). If throughout \( t = m_c + 1, \ldots, m + (n_c - k - 1) \) the pairing pattern is \( \{1, 3\}, \{2, 4\} \),
then at \( t = m + (n_c - k - 1) + 1 \) players’ histories will be

\[
\begin{align*}
    h^1(m + (n_c - k - 1) + 1) &= (C, \ldots, C, C, \ldots, C) \to C; \\
    h^2(m + (n_c - k - 1) + 1) &= (C, \ldots, C, D, \ldots, D) \to C; \\
    h^3(m + (n_c - k - 1) + 1) &= (C, \ldots, C, D, \ldots, D, C, \ldots, C) \to C; \\
    h^4(m + (n_c - k - 1) + 1) &= (D, \ldots, D, C, \ldots, C) \to D.
\end{align*}
\]

At \( t = m + (n_c - k - 1) + 1 \) only player 1 is contagious and will stay so until \( t = m + 2(n_c - k - 1) \). If throughout \( t = m + (n_c - k - 1) + 1, \ldots, m + 2(n_c - k - 1) \) the pairing pattern is \( \{1, 4\}, \{2, 3\} \), then at \( t = m + 2(n_c - k - 1) + 1 \) players’ histories will be

\[
\begin{align*}
    h^1(m + 2(n_c - k - 1) + 1) &= (C, \ldots, C, D, \ldots, D) \to C; \\
    h^2(m + 2(n_c - k - 1) + 1) &= (C, \ldots, C, D, \ldots, D, C, \ldots, C) \to C; \\
    h^3(m + 2(n_c - k - 1) + 1) &= (C, \ldots, C, D, \ldots, D, C, \ldots, C) \to C; \\
    h^4(m + 2(n_c - k - 1) + 1) &= (D, \ldots, D, C, \ldots, C) \to ?
\end{align*}
\]

If \( 3(n_c - k - 1) \geq n_c \) [or, equivalently \( n_c \geq \frac{3}{2}(k + 1) \)], player 4 plays \( C \) at \( t = m + 2(n_c - k - 1) + 1 \) and the proof is finished. If \( \frac{3}{2}(k + 1) > n_c \), however, player 4 must make additional errors—play \( C \) against whomever he is paired with—because none of the other three players is contagious anymore. The number of errors required is the lower of the following two: (i) number of additional \( D \)s that must enter \( h^4 \) for player 4 to start playing \( C \), or (ii) the minimum number of periods that must elapse before player 1, 2 or 4 becomes contagious again. A short examination shows that the former number is smaller and thus player 4 must make \( n_c - 3(n_c - k - 1) = 3(k + 1) - 2n_c \) errors.

It is possible but lengthy to show that the same or a worse result would be achieved even if player 4 did not make his \( k + 1 \) errors throughout \( t = n_c - k, \ldots, n_c \), but in some \( k + 1 \) periods between \( t = n_c - k \) and \( t = m \). What remains to be shown is that a better result cannot be achieved if players 3 and 4 make errors even when paired with players 1 and 2. Think again of a situation at \( t = n_c - k \). If the pairing pattern is \( \{1, 3\}, \{2, 4\} \) at \( t = n_c - k \), then both players 3 and 4 will need to make an error or else player 1 or 2 will get converted to strategy \( D \). The same will apply to periods \( t = n_c - k + 1, \ldots, n_c \), the result being
2(k + 1) errors. Since 2(k + 1) > (k + 1) + 1 \left( \frac{3}{2}(k + 1) > n_c \right) [3(k + 1) - 2n_c], the proof is finished.

Proof of Lemma 6: The intuition behind this proof is very similar to that behind the previous proof. Thus, we will only sketch the proof without going into details.

Reset time so that \( t = 1 \). Suppose that in periods \( t = 1, \ldots, n_c - k + 1 \) the pairing pattern contains pairs \{ (2s+1), 1 \} and \{ (2s+2), 2 \}, in periods \( t = n_c - k, \ldots, 2(n_c - k - 1) \) pairs \{ (2s+1), 3 \} and \{ (2s+2), 4 \}, and so forth up to periods \( t = (s - 1)(n_c - k - 1) + 1, \ldots, s(n_c - k - 1) \) when it contains pairs \{ (2s+1), (2s-1) \} and \{ (2s+2), (2s-1) \}. Let in all these periods players 1, \ldots, 2s who are not paired with (2s+1) or (2s+2) meet only among themselves. For \( i \) such that \( 1 \leq i \leq s \), \( h^{2i-1}, h^{2i} \) at \( t = s(n_c - k - 1) + 1 \) will be

\[
h^{2i-1}(s(n_c - k - 1) + 1) = \left( C, C, C, \ldots, C, D, D, D, C, \ldots, C \right) \rightarrow C;
\]

\[
h^{2i}(s(n_c - k - 1) + 1) = h^{2i-1}(s(n_c - k - 1));
\]

while \( h^{2s+1} \) and \( h^{2s+1} \) at \( t = s(n_c - k - 1) + 1 \) will be

\[
h^{2s+1}(s(n_c - k - 1) + 1) = \left( D, D, D, D, \ldots, D, C, \ldots, C \right) \rightarrow ?;
\]

\[
h^{2s+2}(s(n_c - k - 1) + 1) = h^{2s+1}(s(n_c - k - 1));
\]

If \( s(n_c - k - 1) \geq n_c \) [or, equivalently, \( n_c \geq \frac{1}{s-1}(k+1) \)], players 2s+1 and 2s+2 play \( C \) at \( t = s(n_c - k - 1) + 1 \) without having to make any errors and the proof is finished. If \( \frac{1}{s-1}(k+1) > n_c \), the best we can do is to let player (2s+2) make \( n_c - s(n_c - k - 1) = s(k-1) - (s-1)n_c \) errors in a row against player (2s+1). As a result of these meetings, at \( t = s(n_c - k - 1) - 1 + [s(k+1) - (s-1)n_c] = n_c + 1 \) the histories of players (2s+1), (2s+2) will be

\[
\begin{align*}
    h^{2s+1}(n_c + 1) &= (D, \ldots, D, C, \ldots, C, C, \ldots, C) \rightarrow C; \\
    h^{2s+2}(n_c + 1) &= (D, \ldots, D, C, \ldots, C, C, D, D, D, \ldots, D) \rightarrow D.
\end{align*}
\]

If throughout \( t = s(n_c - k - 1) + 1, \ldots, n_c + 1 \) players 1, \ldots, 2s have been only meeting with each other, the histories \( h^{2i-1} \) and \( h^{2i} \), \( 1 \leq i \leq s \), at \( t = n_c + 1 \) will be

\[
\begin{align*}
    h^{2i-1}(n_c + 1) &= \left( C, \ldots, C, D, \ldots, D, C, C, C, \ldots, C \right) \rightarrow C.
\end{align*}
\]
At $t = n_c + 1$ player $(2s+2)$ is the last one to be converted, player $(2s+1)$ is contagious and will stay so for $n_c - k - 1$ periods, and for each $i$, $1 \leq i \leq s$, players $(2i-1)$ and $2i$ will become contagious in $i(n_c - k - 1)$ periods, staying so for at least $n_c - k - 1$ periods. If player $(2s+2)$ is paired with player $(2s+1)$ first, then with player 1, player 2, etc., up to player $2s$, meeting everybody for $n_c - k - 1$ periods in a row, and if the remaining $2s - 2$ players are always paired among themselves, then at $t = n_c + 1 + (2s + 1)(n_c - k - 1)$ the history of player $2s + 2$ will be

$$i^{2s+2}(n_c + 1 + (2s + 1)(n_c - k - 1)) = \begin{array}{c} D, D, \ldots, D, C, C, \ldots, C \end{array} \to \emptyset.$$ 

If $(2s + 1)(n_c - k - 1) \geq n_c$ [or, equivalently, $n_c \geq \frac{2s+1}{2s}(k + 1)$], player 4 plays $C$ at $m + 1$ and the proof is finished. If $\frac{2s+1}{2s}(k + 1) > n_c$, however, player $(2s+1)$ must make additional $n_c - (2s+1)(n_c - k - 1) = (2s+1)(k + 1) - 2sn_c$ mistakes.

That it is not possible to convert players $(2s+1)$ and $(2s+2)$ to $C$ with fewer mistakes follows from the discussion in the proof of Lemma 5.

8 References


