Discussion Paper #1158

June 20, 1996

# "Simple and Clever Decision Rules in Single Population Evolutionary Models"

William Sandholm
Northwestern University

www.kellogg.nwu.edu/research/math

CMS-EMS
The Center for
Mathematical Studies
in Economics &
Management Sciences

Northwestern University
2001 Sheridan Road 580 Leverone Hall Evanston, IL 60208-2014 USA

Discussion Paper No. 1158

# Simple and Clever Decision Rules in Single Population Evolutionary Models†

by

William H. Sandholm§

First Version: June 7, 1995
This Version: June 20, 1996

† I would like to thank Toshimasa Maruta, Zvika Neeman, Ady Pauzner, and Nicola Persico, as well as participants in a Northwestern University seminar for their comments. I would especially like to thank Eddie Dekel and Jeroen Swinkels for their guidance. Financial support from a National Science Foundation Graduate Fellowship is gratefully acknowledged.

§ MEDS-KGSM, Northwestern University, Evanston, IL, 60208. E-mail: whs@nwu.edu

# Abstract

This paper compares two decision rules available to myopic players who are repeatedly randomly matched to play a 2 x 2 symmetric game. Players using the *simple decision rule* evaluate the strategies by comparing their current payoffs to those of an opponent currently playing the other strategy, while those following the *clever decision rule* assess the strategies under the assumption that opponents' actions are fixed. It is shown that while populations of simple players can fail to learn to play Nash equilibria or even dominant strategy equilibria, populations of clever players always learn to play approximate Nash equilibria.

## 1. Introduction

Evolutionary game theory models the development of rational group behavior in populations of myopic individuals. What distinguishes this literature from traditional game theory is the assumption that players' strategy adjustments only depend on the performance of available strategies in response to opponents' current behavior. Neither the history of the interaction nor considerations about future play affect the adjustment process.

Recent work in stochastic evolutionary game theory, most notably the papers of Foster and Young (1990), Kandori, Mailath, and Rob (1993), and Young (1993), have used evolutionary ideas in defense of Nash equilibrium and its refinements. It is well known that strong information and rationality requirements are needed for players in a one shot game to play Nash equilibrium or any refinement thereof. In contrast, work on stochastic evolution has shown that when populations of players are repeatedly randomly matched, their individually myopic choices can lead to rational outcomes on a societal level. The rationality assumptions needed in these models are quite weak, as are assumptions concerning players' knowledge about their opponents. The main result of this literature demonstrates that in coordination games, when players independently switch strategies arbitrarily with small probabilities, in the long run society will nearly always be coordinating on the risk dominant equilibrium.

We study the behavior of populations all of whose members follow one of two myopic decision rules. In order to decide whether to switch strategies, players using the *simple decision rule* compare their payoffs to those of opponents currently selecting the other strategy. Under certain assumptions about the situation modeled, a player can implement the simple decision rule without any knowledge of the game being played or of the population playing it. He must simply ask a player who is playing the other strategy how well he fared, and switch if the opponent's strategy outperformed his own.

Players using the *clever decision rule* players are more astute. A clever player realizes that in a single population random matching environment, the payoff received by an opponent currently playing another strategy is not the same as the payoff he himself should expect from switching to that strategy. This difference arises because no player is ever randomly matched against himself. For example, suppose a player currently playing $s_1$ were to switch to $s_2$. Compared to that which a current $s_2$ player encounters, the $s_1$ player would face a population which contains one less $s_1$ player (himself) and one more $s_2$ player (his opponent). Unlike a simple player, a clever player assumes that the play of the *rest of the population* is fixed when evaluating his options. Doing so requires that a clever player know the payoffs of the game, so that he can carry out the counterfactual reasoning required to calculate his payoffs from switching. However, as players no longer must rely on realized payoffs to make their decisions, sporadic matching is sufficient to drive the model.

We show that when players use the simple decision rule, the play of the population may converge to distributions arbitrarily far from Nash equilibria, and may not converge at all. In particular, we give an example of a dominant strategy game in which the dominant strategy equilibrium payoffs are the Pareto dominant payoffs of the game, but in which members of a population of simple players eventually all play the dominated strategy. This is a counterexample to Theorem 2 of Kandori, Mailath, and Rob (1993).[1] For any fixed dominant strategy game, there is a population size large enough that the dominant strategy equilibrium is selected; however, for any fixed population size, a counterexample can be found. Furthermore, we show that generically, the play of a population of simple players will not converge to distributions in which both strategies are used. Thus, mixed strategies cannot be learned by such a population.

When players are clever, the results are much brighter. We show that in all 2 x 2 symmetric games, and for all population sizes, the play of a population of clever players converges to a distribution of strategies which approximates a Nash equilibrium of the underlying game. In particular, in games in which there is a unique symmetric Nash equilibrium which is in mixed strategies, play always converges to the distribution which best represents the Nash equilibrium.[2]

The interpretation of these results depends in turn on one's interpretation of stochastic evolutionary game theory. One may view stochastic evolution as a defense of Nash equilibrium and its refinements. Given some game, one would then want to find a model with the weakest possible assumptions concerning the players' capabilities in which play converges to a Nash equilibrium. We show that for fixed population sizes, the simple decision rule, and hence the assumptions

---

1 I have learned through correspondence with one of the authors that the first such counterexample was discovered by Peter McCabe.

2 We do not apply any notion of stochastic stability in this paper. The purpose of the paper is to contrast the behavior of populations following the simple and clever decision rules; adding the machinery of stochastic stability would only obscure the main point. Of course, the results which would arise from a stochastic stability analysis are easily derived.

which underlie it, are insufficient to generate convergence to Nash equilibrium. Allowing arbitrarily large population sizes may temper this conclusion somewhat, but it should be borne in mind that generically, simple players never learn to play a mixed strategy equilibrium. In contrast, if one is willing to allow the knowledge assumptions required for clever players, convergence to Nash equilibrium is guaranteed.

Rather than viewing stochastic evolutionary game theory as an attempt to defend Nash equilibrium, it may also be understood as a modeling tool for situations in which players base their decisions on heuristics. To act in a fully rational manner in an environment involving repeated interactions requires that a player perform complex computations based on the history of play and beliefs about opponents' entire repeated game strategies. In many situations, in particular those in which the interaction is only of minor consequence to the players, a less sophisticated analysis may be warranted; everyday problems are not solved with dynamic programs.

Consider, for example, a situation in which players deem reputation effects unimportant, as would be the case in a random matching environment, and in which players believe that changes in the play of the population as a whole develop slowly. In this setting, rather than undertake the immense effort of calculating a fully rational strategy, players might simply choose best responses to the current distribution of strategies in the population, in the hope that this simple choice might perform reasonably well in the short run. Even if he does not actually think that his opponents will not switch actions, a player following the clever decision rule uses this conjecture to inform his decision. If all players do so, convergence to an approximate Nash equilibrium is guaranteed.

We observe that the differences between the simple and clever decision rules only exist in single population models of evolution. If instead there is a distinct population for each role in the base game (as in, for example, Samuelson (1994)), then players need only concern themselves with the behavior of populations other than their own when deciding whether to switch strategies. Hence, the features which distinguish the simple and clever decision rules are of no consequence in this setting.

In independent work, Rhode and Stegeman (1996) analyze the behavior of the simple decision rule in the model of Kandori, Mailath, and Rob (1993). They consider an example similar to the one which we study in the following section, and perform an analysis which corresponds to our analysis of the simple decision rule in

Section 4.1. However, in contrast to our own work, their analysis does not consider any alternative decision rules.

We proceed with an example which shows the counterintuitive results which can arise under the simple decision rule, as well as the manner in which the clever decision rule avoids these difficulties. The model and our results can be found in Sections 3 and 4, respectively, and Section 5 concludes.

## 2. An Example

Consider a group of eight people who work in an office. During the course of a typical day, each pair of officemates meets once. Upon meeting an officemate, a person has two options. He may be helpful, (by, for example, providing some information that his officemate would find valuable), or he may not be helpful (by not doing so). Being helpful is a dominant strategy, because it is costless, and it makes the person feel good about himself. Naturally, a person would rather meet an officemate who is helpful than one who is not. Everyone feels that the tangible benefits of being matched with a helpful officemate far outweigh the psychological benefits of being helpful oneself.

Such a situation is modeled in the game in Figure 1.

| | $H$ | $NH$ |
|---|---|---|
| $H$ | 9, 9 | 1, 8 |
| $NH$ | 8, 1 | 0, 0 |

Figure 1

Notice that $H$ is a dominant strategy, and that (9, 9) is the unique Pareto efficient payoff.

In order to analyze this game, we need to specify a decision rule for each player and a mechanism which indicates which players will be allowed to switch strategies. For the sake of simplicity, we shall assume that exactly one player each period is allowed to switch. In the sequel, this arrangement will be called *1-selection*. In addition, to simplify our exposition, we call a state *unanimous* if at that state all players use the same strategy.

Suppose that the players use the simple decision rule. Under this rule, a player decides whether to switch strategies by comparing his payoff to that of another player who selected the other strategy. Surprisingly, at all nonunanimous states, simple players who are currently being helpful would prefer to stop being helpful: letting $\pi_s(z)$ denote the payoffs to a player using strategy $s$ if $z$ players are currently playing strategy $H$, it is easily checked that $\pi_H(z) = \frac{1}{7}(8z - 1) < \frac{1}{7}(8z) = \pi_{NH}(z)$, where the state $z$ is the number of players currently being helpful. At states 0 and 8, the players rules must be somewhat arbitrary, since there are no opponents using the unchosen strategy to consult. However, regardless of what is assumed, play from all states other than 8 leads to states 0 and 1. If players have any probability of switching at state 8, states 0 and 1 are always reached and never exited. Thus, not being helpful, the dominated strategy, becomes the norm.

Rather than using individual selection to generate a stochastic dynamic, one may, as in Kandori, Mailath, and Rob (1993, henceforth KMR), define a deterministic dynamic which is based on the simple decision rule. The slowest such dynamic is given by $b$, where $b(z) = z - 1$ at all states $z \in \{1, \ldots, 7\}$. KMR focuses on the *long run equilibria* of the game: states which occur with non-vanishing probability when players have small probabilities of choosing their strategies arbitrarily. One can show that if $b(0) \leq 6$, state 8 is not a long run equilibrium. Thus, all players selecting the dominant strategy need not be a long run equilibrium, contradicting Theorem 2 of KMR.

What goes wrong? In random matching environments, it is assumed that players are never randomly matched against themselves. When the population is finite, this fact is made explicit in the payoff functions, which as a result differ from the standard von Neumann-Morgenstern expected utility functions. In the game above, when the state is $z$, a player selecting $H$ plays $z$ - 1 helpful players. Using the simple decision rule, he compares his payoffs to those of a player currently playing $NH$. She plays $z$ helpful players. The player selecting $H$ fails to take into account that he is one of the helpful players with whom his opponent is matched. Consequently, when he switches to $NH$, this $z$th helpful opponent no longer exists, so he only gets to play the $z - 1$ helpful players who remain. This, of course, is the same number of helpful players he was facing originally. However, he is now is being unhelpful, and so does worse against each of his opponents. Thus, the player is worse off after switching.

Now consider a group of clever players. To apply the clever decision rule, we need to compare the payoffs to $H$ at each state in $\{1, 2, \ldots, N\}$ to the payoffs to $s_2$ in the state with one fewer $H$ player. For any $z \in \{1, 2, \dots, N\}$, $\pi_H(z) = \frac{1}{7}(8z - 1) > \frac{1}{7}(8z - 8) = \pi_{NH}(z - 1)$. So, under the clever decision rule, helpful players never want to switch when given the opportunity; players who are not helpful always want to switch. Thus, the unique stable distribution has all players choosing $H$, the dominant strategy.[3]

## 3. The Model

Consider a population of $N$ players which repeatedly plays the symmetric 2 x 2 game in Figure 2.

| | $s_1$ | $s_2$ |
|---|---|---|
| $s_1$ | $a, a$ | $b, c$ |
| $s_2$ | $c, b$ | $d, d$ |

Figure 2

We identify 2 x 2 symmetric games with vectors $G = (a, b, c, d) \in \Re^4$. Let $I = \{1, 2, \dots, N\}$ denote the set of players. In each period $t = 0, 1, 2, \dots$ , $\zeta_t$ represents the number of members of the population playing strategy $s_1$ in period $t$. Thus $\zeta_t$ takes values in $Z = \{0, 1, \dots, N\}$, the set of *states of the world* or *distributions.*

Players are repeatedly randomly matched to play the game in Figure 2. The payoffs to players selecting $s_1$or $s_2$ when the current state of the population is $z$ are given by:

$$\pi_1(z) = \frac{(z-1)}{(N-1)}a + \frac{(N-z)}{(N-1)}b \qquad \text{for } z \in \{1, 2, \dots, N\},$$

$$\pi_2(z) = \frac{z}{(N-1)}c + \frac{(N-z-1)}{(N-1)}d \qquad \text{for } z \in \{0, 1, \dots, N-1\}.$$

3 It may be interesting to note that when we define payoffs to include the possibility of being matched against oneself, a population of simple players will move towards playing the dominant strategy (in a 2 x 2 symmetric dominant strategy game). However, clever players still do better: for example, in many "prisoners' dilemma" games, clever players eventually all play the dominated strategy. This result is analogous to Proposition 2(*iii*), below.

The payoffs to strategy $s_1$ in state 0 are undefined because no player is selecting $s_1$ in that state. $\pi_2$ is undefined in state $N$ for the analogous reason. The equations take into account that in the process of random matching, a player will not be matched against himself. Hence, the payoffs differ from those that would be generated by facing a mixed strategy with weight $(z/N)$ on $s_1$. A precise interpretation of the payoffs will be deferred until after the formal introduction of the decision rules.

To model the actions of the myopic individuals playing this game, we need to define a function which specifies their (possibly randomly selected) action given their previous action and the state of the world. Before doing this, we must develop notation which will allow us to define random variables. Let $(\Omega, \mathcal{F}, P)$ denote the probability space on which all random variables in the model are defined. Given some set $A$, $\Delta A \equiv \{X : \Omega \to A\}$ is the set of all random variables whose range is the set $A$. When defining random variables, we abuse notation and identify a random variable with its distribution; all of the basic random variables are assumed to be independent. Let $\delta_a$ represent a distribution degenerate at $a$, and let $(a_1, p_1; \dots ; a_k, p_k)$ denote the distribution of a random variable $X$ for which $P(X = a_i) = p_i$, $i = 1, \dots, k$.

We define a *decision rule* $r = (r_1, r_2)$ to be a pair of functions $r_1$: $\{1, \dots, N\} \to \Delta\{S, D\}$ and $r_2 : \{0, \dots, N-1\} \to \Delta\{S, D\}$, where $S$ and $D$ stand for "Switch" and "Don't switch" respectively. A player using decision rule $r$ makes his decision about his action in the next period according to $r_i$ when his current strategy is $s_i$. The argument of the function is the current state; thus $r_1$ cannot take 0 as an argument since in state 0 no one is playing $s_1$; similarly, $r_2$ cannot take $N$ as an argument. The output of the function is the player's choice of action.

Ideally, evolutionary economic models make minimal assumptions on the knowledge and computational abilities of the players. To construct a model with weak knowledge assumptions, suppose that in each period, exactly one match occurs between each pair of players. In this case, $\pi_1$ and $\pi_2$ are the payoffs per match that the players actually receive.[4] As a consequence, one can define a decision rule which is based only on realized payoffs, requiring no knowledge about either the game or the number of other players. In particular, an $s_1$ player who is considering switching to $s_2$ can simply ask an opponent who played $s_2$ last period how she fared, and then switch strategies if her payoffs were higher than his own.

[4] Alternatively, as noted in KMR, for this interpretation one could specify that in each period, an infinite sequence of matches occurs, with each match being selected independently with equal probability. Then the strong law of large numbers guarantees that the realized per match payoffs are given by $\pi_1$ and $\pi_2$.

This is the *simple decision rule, e*. Players who use this rule will be called *simple players*. The rule is defined formally by

$$e_1(z) = \begin{cases} \delta_S & \text{if } z \neq N \text{ and } \pi_2(z) > \pi_1(z), \\ \delta_D & \text{if } z \neq N \text{ and } \pi_1(z) \geq \pi_2(z), \\ (S, \alpha; D, 1-\alpha) & \text{if } z = N, \end{cases}$$

$$e_2(z) = \begin{cases} \delta_S & \text{if } z \neq 0 \text{ and } \pi_1(z) > \pi_2(z), \\ \delta_D & \text{if } z \neq 0 \text{ and } \pi_2(z) \geq \pi_1(z), \\ (S, \beta; D, 1-\beta) & \text{if } z = 0, \end{cases}$$

where $\alpha, \beta \in [0, 1]$ are constants. At states besides 0 and $N$, the rule simply expresses formally the procedure described above. At states 0 and $N$ (the *unanimous states*), a player cannot follow this procedure because all other players have chosen the same strategy as he. We assume that in this situation, a player will choose his actions randomly according to some fixed probabilities and that each player make this choice independently of the others. In the sequel we will make more specific assumptions about what the players do at these states; however, the exact assumptions made are not important.

Traditionally, game theory has players choose best responses to some conjecture about opponents' play. For myopic players involved in repeated interactions, the simplest such conjecture holds that opponents will continue to play in the coming period the strategies they chose in the previous period. The best responses to such a conjecture define the *clever decision rule, c*. Formally, the clever decision rule is defined as follows.

$$c_1(z) = \begin{cases} \delta_S & \text{if } \pi_2(z-1) > \pi_1(z), \\ \delta_D & \text{if } \pi_1(z) \geq \pi_2(z-1), \end{cases}$$

$$c_2(z) = \begin{cases} \delta_S & \text{if } \pi_1(z+1) > \pi_2(z), \\ \delta_D & \text{if } \pi_2(z) \geq \pi_1(z+1). \end{cases}$$

Players using this rule will be called *clever players*. By using this decision rule, a player essentially takes the actions of the other players as the "state of the world" (in the commonly used sense of the parameters which are outside his control) and makes his myopic best response under the assumption that this "state" is stationary. This requires counterfactual reasoning: a player must be able to calculate what his payoffs would be in the unrealized event that he were to change his strategy, holding the choices of the others fixed. In order to make this calculation, a player needs to know the state, his own previous action, and the base game payoffs; no actual matchings need occur. Therefore, for clever players, the payoff functions $\pi_1$ and $\pi_2$ can be interpreted as expected per match payoffs under the assumption that all matches are equally likely.

Throughout the paper, we assume that all members of each population being considered use the same decision rule. Doing so will emphasize the differences between the performance of the simple and clever rules; namely, that populations of players playing the clever rule always learn to play Nash equilibria, while populations of simple players do not.[5]

In order to generate a Markov chain representing the evolution of the play of the population, we need to explain how in each period the set of players who may change strategies is determined. Formally, a *selection mechanism* is a random variable $m \in \Delta\{0, 1, \ldots, N\}$ which indicates the number of players who may switch their strategies in each period. The exact subset selected is determined by choosing randomly from all subsets of the population of the appropriate cardinality. That is, the selection mechanism is anonymous. Furthermore, we assume non-triviality: $P(m = 0) < 1$.

A game $G$, a decision rule $r$, and a selection mechanism $m$ determine a $|Z|\times|Z|$ transition matrix $Q$ which gives the probability transitions between any two states. Fixing the initial distribution over states $\mu \in \Delta Z$ allows us to define the Markov chain $\{\zeta_t\}_{t=0}^{\infty}$ which describes the evolution of play.

For concreteness, we mention two examples of selection mechanisms. When $m = \delta_k$, with $k \in \{1, \ldots, N\}$, the mechanism is called *k-selection*; exactly $k$ players are selected each period, with each subset of cardinality $k$ occurring with equal probability. If $m$ is given by the binomial distribution,

$$m = (0, (1-\theta)^N; \ldots ; i, \binom{N}{i}\theta^i(1-\theta)^{N-i}; \ldots ; N, \theta^N),$$

for some $\theta \in (0, 1)$, the mechanism is called *θ-selection*. This mechanism results when each player is selected independently with an identical selection probability of $\theta$.

What does it mean for a player to be "selected"? A natural interpretation is that in each period, because of some restriction inherent in the game environment, only

5 See the Conclusion for comments on modeling non-homogenous populations.

certain players are able to switch actions; others simply may not or cannot switch. However, when evolution is used to model situations in which players use heuristics to make decisions, selection need not be interpreted as a constraint on the players' flexibility. Rather, selection may reflect that the players do not constantly monitor the efficacy of their actions. One may think of a "selected" player as one who has opted to put forth the effort to consider changing his strategy in the given period; whether he actually switches strategies depends on the application of his decision rule. Selection is thus interpreted as self-selection; that players are only selected occasionally is a consequence of their disinterest.

## 4. Convergence to Stable Distributions

A base game $G$, a population $I$, a decision rule $r$, a selection mechanism $m$, and an initial distribution $\mu \in \Delta Z$, one can determine a Markov chain $\{\zeta_t\}_{t=0}^{\infty}$ describing the evolution of the play of the population. We study the limiting behavior of this process. Given a decision rule $r$ and a population size $N$, we define a *stable distribution* to be any $z \in Z$ such that: $r_i(z) = \delta_D$ for all $r_i$ that are defined at $z$, where $i \in \{1, 2\}$. Once a stable distribution is reached, no player switches strategies ever again. Thus, stable distributions are simply absorbing states of the Markov chain $\{\zeta_t\}_{t=0}^{\infty}$. The remainder of this section considers whether populations of simple or clever players eventually arrive at stable distributions.

### 4.1 Convergence with simple players

When players use the simple decision rule, each compares how his strategy is currently performing to how the other strategy is currently performing. Such a rule is well defined at all non-unanimous states. Furthermore, at these states, it can be completely characterized by a single linear function. One may define the (*simple*) *payoff difference* $\Delta$: $\{1, \ldots, N-1\} \rightarrow \Re$ by

$$\Delta(z) = \pi_1(z) - \pi_2(z).$$

At a nonunanimous state $z$, players currently playing $s_1$ would like to switch if and only if $\Delta(z)$ is strictly negative, and an $s_2$ player will want to switch exactly when $\Delta(z)$ is strictly positive. Stable interior distributions are those for which $\Delta(z)$ equals zero. It will be convenient to extend $\Delta$ to a linear function on the real line.

Substituting in the payoff functions, we define the *(extended) payoff difference* $\Delta : \Re \rightarrow \Re$.

$$\Delta(x) = \frac{1}{N-1}\{[(a-c)+(d-b)]x + N(b-d)+(d-a)\}.$$

One may characterize simple population dynamics by using the sign of the first derivative of the payoff difference function, which is constant for any fixed game. In Table I, we partition the set of 2 x 2 symmetric games, represented by $\Re^4$, into six varieties according to their best response properties. In the final column, we note the signs of the derivative of the payoff difference, $\Delta'$, which may arise in each variety of game. Let $V = \{v_{DS}, v_{WD1}, v_{WD2}, v_C, v_{MS}, v_{OC}\}$ be the partition of $\Re^4$ into the different varieties of 2 x 2 symmetric games listed in Table I.

| Variety of Game | Payoffs | | BR to $s_1$ | BR to $s_2$ | Sign of $\Delta'$ |
|---|---|---|---|---|---|
| Dominant Strategy | $a > c$ | $b > d$ | $s_1$ | $s_1$ | +, 0, or – |
| | $a < c$ | $b < d$ | $s_2$ | $s_2$ | +, 0, or – |
| Weakly Dominant (1) | $a = c$ | $b > d$ | both | $s_1$ | – |
| | $a < c$ | $b = d$ | $s_2$ | both | – |
| Weakly Dominant (2) | $a > c$ | $b = d$ | $s_1$ | both | + |
| | $a = c$ | $b < d$ | both | $s_2$ | + |
| Coordination | $a > c$ | $b < d$ | $s_1$ | $s_2$ | + |
| Mixed Strategy | $a < c$ | $b > d$ | $s_2$ | $s_1$ | – |
| Opponent's Choice | $a = c$ | $b = d$ | both | both | 0 |

Table I: Varieties of 2 x 2 Symmetric Games

For payoff differences with non-zero derivatives, define

$$x^* \equiv \Delta^{-1}(0) = \frac{N(d-b)+(a-d)}{(d-b)+(a-c)}.$$

When $\Delta'$ is positive, the population moves away from $x^*$. Thus, it would appear that the distribution tends to coordinate at one of the unanimous states. However, this is only true when $x^* \in [0, N]$. When $x^*$ is outside of this range, motion from any internal state will always tend toward one of the two unanimous states: $s_1$ if $x^* < 0$, or $s_2$ if $x^* > N$. When $\Delta'$ is negative, the population moves towards $x^*$. Thus, when $x^* \in (0, N)$, the law of motion satisfies our intuitive notion of how the population should behave in mixed strategy games. But once again, when $x^*$ is outside this interval, play converges to a single unanimous state: towards $s_1$ if $x^* \geq$

$N$, and $s_2$ if $x^* \leq 0$. When $\Delta'$ equals zero, $\Delta$ is a constant. The direction that the population moves is the same in every nonunanimous state: towards $N$ if $\Delta > 0$, or towards 0 if $\Delta < 0$. When $\Delta \equiv 0$, every state is stable.

At this point, we make assumptions about what the simple decision rule suggests at the unanimous states. Given the lack of information that players have about their alternative, one might want to always have the players experiment occasionally at these states. However, for most games, doing so would prevent any state from being stable. We want to define the simple rule at the unanimous states in a way that balances these two concerns. We proceed under the following somewhat arbitrary assumptions: $e(N) = \delta_0$ if and only if $\Delta(N) \geq 0$. $e(0) = \delta_0$ if and only if $\Delta(0) \leq 0$. Under other circumstances, we assume that players will choose to switch strategies with some probability strictly between zero and one. While the results are easier to state under these assumptions, the particular assumptions made do not alter the main conclusion.

Under the assumption stated above about play at unanimous states, we can characterize the states that are stable under the simple decision rule. Which states are stable depends only on the sign of $\Delta'$ and on the value of $x^*$ (or, if $\Delta' = 0$, on the unique value of $\Delta$). In Table II, we completely characterize the possibilities for stability. Items included in parentheses describe the non-generic cases, which occur when $Nx^* \in Z$.

| $\Delta' > 0$ | | $\Delta' < 0$ | | $\Delta' = 0$ | |
|---|---|---|---|---|---|
| Condition | Stable States | Condition | Stable States | Condition | Stable States |
| $x^* < 0$ | $\{N\}$ | $x^* \leq 0$ | $\{0\}$ | $\Delta > 0$ | $\{N\}$ |
| $x^* \in [0, N]$ | $\{0, (x^*,) N\}$ | $x^* \in (0, N)$ | $\varnothing$ (or $\{x^*\}$) | $\Delta = 0$ | $Z$ |
| $x^* > N$ | $\{0\}$ | $x^* \geq N$ | $\{N\}$ | $\Delta < 0$ | $\{0\}$ |

Table II: Stability classes under simple dynamics

Since the values of $\Delta'$ and $x^*$ (and, when $\Delta' = 0$, $\Delta$) only depend on the payoffs of the game, we can partition $\Re^4$, representing the set of payoffs to 2 x 2 symmetric games, according to the stability properties under simple dynamics. In particular, let $\{c^{+1}, c^{-1}, c^0\}$ partition $\Re^4$ into regions where $\Delta'$ is positive, negative, and zero, respectively. For each $c^i$, $i \in \{+1, -1, 0\}$, let $\{c_1^i, c_2^i, c_3^i\}$ partition $c^i$ into the three conditions in Table II. For example, $c_1^{+1}$ is the set of games such that $\Delta' > 0$ and $x^* < 0$, so that the unique stable state is $N$.

We can now state our main proposition concerning the simple decision rule.

**Proposition 1:** *Fix a variety $v \in V$. If $v \cap c^i \neq \emptyset$, $i \in \{+1, -1, 0\}$, then $v \cap c^i_j \neq \emptyset$ for $j$ = 1, 2, 3.*

Proof: In order to prove this Proposition, we must show that for each of the six varieties of 2 x 2 symmetric games listed in Table I, and for each sign of $\Delta'$ that is possible for that given variety, each of the conditions listed in Table II under the value of $\Delta'$ is possible. For what follows, we assume $N \geq 2$ to be fixed.

Dominant Strategy ($a > c, b > d$): The different values of $\Delta'$ are generated by varying $(a - c)$ and $(b - d)$. For $\Delta' > 0$, let $b = 1, d = 0$, and let $e = a - c = 2$. Then $x^* = a - N$, so picking $a$ appropriately generates any desired $x^*$. For $\Delta' < 0$, choose $a = 0, c = -1, f = b - d = 2$ so that $x^* = 2N + d$, where $d$ is a free variable. For $\Delta' = 0$, let $b = 1, d = 0$, and $e = a - c = 1$. Then $\Delta = \frac{1}{N-1}(N - a)$, so the appropriate choice of $a$ can make the payoff difference positive, negative or zero.

Weakly Dominant (1) ($a = c, b > d \Rightarrow \Delta' < 0$): Let $a = c = 0, b = 1$. Then $x^* = N - \frac{d}{d-1}$, so since $d < b = 1$, any $x^* > N - 1$ is possible. Choosing $b = -2$, $x^* = N - \frac{d}{d+2}$, so any $x^* < N + 2$ can be generated.

Weakly Dominant (2) ($a > c, b = d \Rightarrow \Delta' > 0$): Let $b = d = 0$. Letting $c = 1$, $x^* = \frac{a}{a-1}$, so since $a > c = 1$, $a$ can be chosen to generate any $x^* > 1$. Similarly, letting $c = -2$, $x^* = \frac{a}{a+2}$, so by choosing $a > c = -2$ appropriately, one can generate any $x^* < 2$.

Coordination ($a > c, b < d \Rightarrow \Delta' > 0$): Let $b = -1, d = 0, e = a - c = 1$. Then $x^* = \frac{1}{2}(N + a)$.

Mixed Strategy ($a < c, b > d \Rightarrow \Delta' < 0$): Let $b = 1, d = 0, e = c - a = 1$. Then $x^* = \frac{1}{2}(N - a)$.

Opponent's Choice ($a = c, b = d \Rightarrow \Delta' = 0$): Here, $\Delta = \frac{1}{N-1}(d - a)$. ■

This result shows that when players use the simple decision rule, play can converge to states that are arbitrarily different from what one might expect, and often may not converge at all. For a nonunanimous state to be stable, it must be the case that the payoffs to $s_1$ and $s_2$ are identical at that state; generically, this does not occur. This is particularly relevant in mixed strategy games, in which the population always adjusts towards $x^*$. These observations imply the following result.

**Corollary:** *Fix a population of N players and any selection mechanism. In mixed strategy games for which play does not converge to a unanimous state, it is generically true that play does not converge at all.*

Under 1-selection, play in most mixed strategy games will eventually cycle between two adjacent states. However, under more general selection rules, the recurrent set will be much larger. For example, under $\theta$-selection for any $\theta \in (0, 1)$, the recurrent set encompasses the entire state space.[6]

To emphasize particular oddities which may arise under the simple decision rule, we present the following results.

**Proposition 2:** *Fix a population of N players using the simple decision rule, and assume 1-selection.*

*(i) Fix a state* $z \in Z$, *and a constant* $\beta \in (0, 1)$. *Then there exists a mixed strategy game G whose unique mixed strategy equilibrium is* $(\beta, 1 - \beta)$ *and whose unique stable state is z.*

*(ii) There are coordination games with a single strategy that is both Pareto dominant and risk dominant such that play from any initial state converges to all players selecting the other strategy.*

*(iii) If play in a dominant strategy game does not converge to all players choosing the dominant strategy, then the dominant strategy equilibrium payoffs Pareto dominate the payoffs when two players select the dominated strategy. Furthermore, if from any initial state play converges to all players selecting the dominated strategy, the payoffs to the dominant strategy are the unique Pareto efficient payoffs.*

Proof: (*i*) Recall that a game $G = (a, b, c, d)$ is a mixed strategy game if $a < c$ and $b > d$, that $s_1$ is played in the mixed strategy equilibrium with probability $\frac{(d-b)}{(d-b)+(a-c)}$, and that the unique stable state, if it exists, is given $\frac{N(d-b)+(a-d)}{(d-b)+(a-c)}$. So, if $z > N\beta$, chose $a = 1$, $d = 0$, $b = \frac{\beta}{\beta N - z}$, and $c = 1 + \frac{1-\beta}{\beta N - z}$. A similar choice can be made if $z < N\beta$. If $z = N\beta$, chose $a = d = 0$, $b = \beta$, and $c = 1 - \beta$.

(*ii*) For any fixed $N$, choose $a = 1$, $b = -N - 1$, $c = 0$, and $d = -N$.

(*iii*) Without loss of generality, assume that $s_1$ is dominant, so that $a > c$ and $b > d$. $\pi_2(z) > \pi_1(z)$ if and only if

6 One might argue that the maximum number of players that can be selected in a single period can be assumed to be small if we think of the period length as very short. However, such an assumption is not without loss of generality, since the timing of the payoffs, information aggregation, and opportunities for reconsideration in the environment being modeled should determine the selection rate. For instance, in the example from Section 2, if the officemates only reassess their behavior at the end of the day, restricting the rate of selection is not sensible.

(*) $z((a-c)+(d-b)) < N(d-b)+(a-d)$.

There are three cases to consider, depending on the sign of $((a-c)+(d-b))$.

If $((a-c)+(d-b)) = 0$, then (*) reduces to $a > N(b-d)+d$. It follows that $a > b + (N-1)(b-d) > b$, so $(a, a)$ is the unique Pareto efficient payoff.

Now suppose that $((a-c)+(d-b)) > 0$. In this case, (*) reduces to $z < \frac{N(d-b)+(a-d)}{(a-c)+(b-d)}$. Thus, if play does not move towards state $N$ from all initial states, it must be that $\frac{N(d-b)+(a-d)}{(a-c)+(b-d)} > 0$. This reduces to $a > N(b-d)+d > d$, so the payoffs to $(s_1, s_1)$ are greater than the payoffs to $(s_2, s_2)$. If play always moves towards state 0, then $\frac{N(d-b)+(a-d)}{(a-c)+(b-d)} > N - 1$. This implies that $a > (N-1)(a-c)+b > b$.

Finally, suppose that $((a-c)+(d-b)) > 0$, so that (*) reduces to $z > \frac{N(d-b)+(a-d)}{(a-c)+(b-d)}$. If the dominant strategy is unstable, it is necessary that $\frac{N(d-b)+(a-d)}{(a-c)+(b-d)} < N$, which in turn implies that $a > N(a-c)+d > d$. Now, in addition suppose that play converges to state 0 from any initial state. This implies that $\frac{N(d-b)+(a-d)}{(a-c)+(b-d)} < 1$. Rearranging terms, the previous equation becomes $a > (N-2)(b-d)+(a-c)+b > b$, completing the proof. ■

What drives these odd results is the combination of the form of payoffs in random matching environments with the finitude of the population. As long as each player has a positive mass in the population, games can be constructed that make the simple decision rule look foolish. On the other hand, if the game is held fixed and the population size made large, each player's mass relative to that of the population becomes small. Hence, one might expect that the oddities described above might cease to exist. The following result makes this argument precise.

**Proposition 3:** (*i*) *Fix a dominant strategy game G. For N large enough, play under the simple decision rule converges to the dominant strategy.*

(*ii*) *Fix a coordination game or mixed strategy game G'. Let $x^*(N)$ be the division point of the simple dynamic when the population size is N, and let $(\alpha^*, 1-\alpha^*)$ be the mixed strategy equilibrium of G'. Then $N\alpha^* - x^*(N) = \frac{d-a}{(d-b)+(a-c)}$ for all N, so $(x^*/N) \to \alpha^*$ as $N \to \infty$.*

Proof (*i*): Assume without loss of generality that $s_1$ is the dominant strategy, so that $a > c$ and $b > d$. Recall that $\Delta(x) = \frac{1}{N-1}\{[(a-c)+(d-b)]x + N(b-d)+(d-a)\}$. If $a - c \geq b - d$, then for $x \in [0, N]$, $\Delta(x) \geq \frac{1}{N-1}(N(b-d)+(d-a))$, which is positive for $N$ large enough. If $a - c < b - d$, then for $x \in [0, N]$, $\Delta(x) \geq \Delta(N) = \frac{1}{N-1}(N(a-c)+(d-a))$, which is positive for $N$ large enough.

(*ii*) The result follows immediately from the observation that $x^*(N) = \frac{N(d-b)+(a-d)}{(d-b)+(a-c)}$ and $\alpha^* = \frac{(d-b)}{(d-b)+(a-c)}$. ■

The interpretation of part (*i*) is clear. Part (*ii*) says that if the state space is superimposed on to mixed strategy space, represented by the unit interval, the division point of the simple dynamic approaches the mixed strategy equilibrium as $N$ approaches infinity. Hence, viewed from this light, the simple dynamic works properly once the population is large enough. Note, however, that the Corollary above still applies: it is generically true that play will not converge in mixed strategy games, since this can only happen if $x^*(N)$ is an integer.

4.2 Convergence with Clever Players

Under the clever decision rule, the population dynamic is also characterized by a single function which compares how strategy $s_1$ performs in state $z$ to how strategy $s_2$ performs in state $z - 1$. However, it will be more convenient to represent the decision rule by two payoff difference functions, used by $s_1$ players and $s_2$ players respectively to determine whether they want to switch strategies. These functions, $\Delta_1(z) = \pi_1(z) - \pi_2(z - 1)$ and $\Delta_1(z) = \pi_1(z) - \pi_2(z + 1)$ are initially defined on $\{1, \dots, N\}$ and $\{0, \dots, N - 1\}$ respectively, and are nonnegative when a player using the given strategy is content. Extending these to functions on the entire real line, we see that

$$\Delta_1(x) = \tfrac{1}{N-1}[(x-1)(a-c)+(N-x)(b-d)],$$
$$\Delta_2(x) = \tfrac{1}{N-1}(x(c-a)+(N-x-1)(d-b)).$$

The remaining results are all proven using standard finite state Markov chain arguments. In what follows, the population size $N$ is any fixed integer greater than one, and "convergence" means almost sure convergence.

**Proposition 4:** *Suppose that players use the clever decision rule. Then for any selection mechanism m:*

(*i*) *In dominant strategy games, play converges to the state in which all players choose the dominant strategy.*

*(ii) In weakly dominant strategy games with one Nash equilibrium, play converges to a state in which all or all but one player select the weakly dominant strategy.*

*(iii) In weakly dominant strategy games with two Nash equilibria, both unanimous states are stable, and play from any nonunanimous state converges to the state in which everyone plays the weakly dominant strategy.*

*(iv) In opponent's choice games, all states are stable.*

Proof: *(i)* Suppose without loss of generality that $s_1$ is the dominant strategy: $a > c$ and $b > d$. Then $\Delta_1 > 0$ on $\{1, \ldots, N\}$ and $\Delta_2 < 0$ on $\{0, \ldots, N-1\}$. Thus, state $N$ is stable, and for all $z < N$, there exists a $x > z$ such that $Q_{zx} > 0$, and for all $y > z$, $Q_{zy} = 0$. Thus, play converges to state $N$.

*(ii)* Suppose without loss of generality that $s_1$ is the weakly dominant strategy. Since only $(s_1, s_1)$ is a Nash equilibrium, $a = c$ and $b > d$. Thus $\Delta_1(z) \geq 0$ on $\{1, \ldots, N\}$. $\Delta_2(z) < 0$ on $\{0, \ldots, N-2\}$, and $\Delta_2(N-1) = 0$. So, states $N$ and $N-1$ are stable. Also, for all $z < N-1$, there exists an $x > z$ such that $Q_{zx} > 0$, and for all $y > z$, $Q_{zy} = 0$. Thus, play converges to state $N$ or state $N-1$.

*(iii)* Suppose without loss of generality that $s_1$ is the weakly dominant strategy. Since $(s_2, s_2)$ is a Nash equilibrium, $a > c$ and $b = d$. So, $\Delta_1(z) > 0$ on $\{1, \ldots, N\}$. $\Delta_2(z) < 0$ on $\{1, \ldots, N-1\}$, and $\Delta_2(0) = 0$. So, states 0 and $N$ are stable. For all nonunanimous states $z$, there exists a $x > z$ such that $Q_{zx} > 0$, and for all $y > z$, $Q_{zy} = 0$. Thus, play from all nonunanimous states converges to state $N$.

*(iv)* Under these assumptions, $\Delta_1 \equiv \Delta_2 \equiv 0$. Thus, no one ever wants to switch. ■

We now consider the cases of coordination games and mixed strategy games. In both cases the payoff difference functions cross zero at

$$x_1^* \equiv \Delta_1^{-1}(0) = \frac{N(d-b)+(a-c)}{(d-b)+(a-c)},$$

$$x_2^* \equiv \Delta_2^{-1}(0) = \frac{(N-1)(d-b)}{(d-b)+(a-c)}.$$

Thus, $x_1^* - x_2^* = 1$. The probability of playing $s_1$ in the symmetric mixed strategy Nash equilibrium in coordination games and mixed strategy games is given by

$$\alpha^* = \frac{(d-b)}{(d-b)+(a-c)}.$$

Therefore, $N\alpha^*$, the representation of the mixed strategy Nash equilibrium on $[0, N]$ always lies between $x_1^*$ and $x_2^*$. The convergence result for coordination games is now easily proved.

**Proposition 5:** *Suppose players use the clever decision rule, and that the selection mechanism $m$ satisfies $P(m = N) < 1$. If $G$ is a coordination game, $0$ and $N$ are the stable states, and play converges to one of them.*

Proof: Since, $a > c$ and $b < d$, $\Delta_1' > 0$ and $\Delta_2' < 0$, so $\Delta_1(z) \geq 0$ if and only if $z \geq x_1^*$, and $\Delta_2(z) \geq 0$ if and only if $z \leq x_2^*$. Thus, states 0 and $N$ are stable, and following the steps in the proof of Proposition 4(*i*), it can be shown that play starting at any node outside $(x_2^*, x_1^*)$ converges to state 0 or $N$.

It remains to show that play starting at a state $z \in (x_2^*, x_1^*)$ will leave this set. Since $x_1^* - x_2^* = 1$, there is at most one such state. Since $\Delta_1(z)$ and $\Delta_2(z)$ are both negative, any player selected at this state will switch strategies. Thus, we only need show that it is possible at $z$ for the selection not to select the same number of $s_1$ and $s_2$ players. That $P(m = N) < 1$ guarantees this. ■

In mixed strategy games, a slightly stronger condition on the selection mechanism is needed to guarantee convergence. Clearly, if all players are selected every period, and play begins at a state which is not stable, the distribution will cycle between states 0 and $N$ forever. Moreover, it is easily checked that for $k > N/2$, if no stable state is within $k$ of 0 or $N$, then play that does not start at a stable state will cycle between states within $k$ of state 0 and states within $k$ of state $N$. However, barring this, convergence is guaranteed, as Proposition 6 shows.

**Proposition 6:** *Suppose that a population of $N$ clever players plays a mixed strategy game $G$, and that the selection mechanism $m$ satisfies $P(m > \frac{N}{2}) < 1$. Then play converges to the unique $z \in (x_2^*, x_1^*)$ if $x_1^* \notin Z$, or to either $x_1^*$ or $x_2^*$ otherwise. Furthermore, no stable distribution is unanimous.*

Proof: Since $\text{sgn}(\Delta_1(z)) = \text{sgn}(x_1^* - z)$ and $\text{sgn}(\Delta_2(z)) = \text{sgn}(z - x_2^*)$, $s_1$ players want to switch at states $z > x_1^*$, and $s_2$ players want to switch at states $z < x_2^*$. Thus, the set of stable states is given by $S = [x_2^*, x_1^*] \cap Z$. To prove convergence into $S$, we only need to show that for every $z \notin Z$, there exists a path $z = z_0, z_1, \ldots, z_j$ such that $Q_{z_0 z_1} > 0, \ldots, Q_{z_{j-1} z_j} > 0$. Suppose without loss of generality first that there is a stable state $z^*$

$\leq \frac{N}{2}$. For each $x < z^*$, $z \notin S$, there exists a $y > x$ such that $Q_{xy} > 0$, so we only need to consider states $z > z^*$. For states $z > [\frac{N}{2}]_-$, if $[\frac{N}{2}]_-$ players are selected, including all of the $N - z$ $s_2$ players, the next state will be $[\frac{N}{2}]_+$. If less than $[\frac{N}{2}]_-$ players are selected, it is possible for the next state to be in $\{[\frac{N}{2}]_+, \ldots, z - 1\}$. Thus, a path may always be constructed from $z$ to $[\frac{N}{2}]_-$. For states $z$ in $\{\max S, \ldots, [\frac{N}{2}]_+\}$, there is a positive probability of a transition to $z - 1$, since it is possible for exactly one $s_1$ player to be selected. Therefore, the path needed can always be constructed. ■

Combining Propositions 4 through 6, we can state our main result. To relate symmetric Nash equilibria of a 2 x 2 symmetric game $G$ to the stable distributions of an $N$ player random matching game, we define an *approximate Nash distribution* in an $N$ player random matching environment to be any state $z \in Z = \{0, 1, \ldots, N\}$ such that there is a symmetric Nash equilibrium of $G$, $(\alpha^*, 1 - \alpha^*)$ satisfying $|z - N\alpha^*| \leq 1$. The following result is an immediate consequence of the three previous propositions.

**Theorem:** *Let G be a 2 x 2 symmetric game. Suppose that a population of N clever players is repeatedly randomly matched according to a selection mechanism m satisfying* $P(m > \frac{N}{2}) < 1$. *Then play from any initial distribution converges almost surely to an approximate Nash distribution.*

## 5. Conclusion

This paper compares the behavior of populations whose members all use either the simple or clever decision rule. Populations of simple players need not learn to play Nash equilibria or even dominant strategy equilibria of the underlying game. While these negative results are somewhat mitigated when the size of the population is large, it is generically true that simple populations cannot learn to play mixed strategy equilibria. In contrast, populations of clever players always learn to play approximate Nash equilibria.

These results can be interpreted as a comparison of two heuristics which players might use in lieu of computing best responses to their beliefs about the play of opponents. Using such heuristics is sensible in large population games in which reputation effects are minimal. In particular, if players follow the clever decision rule, the population will eventually reach a stable distribution of strategies which is an approximate Nash equilibrium of the underlying game. Once such a distribution is reached, no player can benefit from a unilateral deviation. The clever decision rule thus has a *self-sustaining* property: if all players follow the rule, in the long run (once a stable distribution is reached), there are no benefits from deviating. Clearly, the simple decision rule does not share this property. This suggests the following question: What is the myopic decision rule with the weakest knowledge requirements which is self-sustaining and results in Nash equilibrium play?

While the simple decision rule is not self-sustaining, this does not necessarily mean that simple players would discover the inadequacy of their rule. Recalling the example from Section 2, consider a simple player who just switched from being helpful to not being helpful. One might think that after receiving a lower payoff, the player would realize that he was doing worse against both helpful and unhelpful officemates, and would conclude that he should return to being helpful. However, if the player reasoned like this, he would be well on his way to being clever, since in order to undertake this analysis he would need to know much of the information required to implement the clever decision rule. Indeed, the main advantage of the simple decision rule is that a player can implement it without knowing the distribution of strategies in the population or even of the base game being played. He may see his payoffs as a single entity, not as the sum of the results of a series of encounters. Consequently, the realization described above need not occur to him. Rather, when a simple player wishes to evaluate his choice of action, he comes equipped with a means of doing so: his simple decision rule. And at all interior states, he is satisfied with his choice: a player who has just decided not to be helpful may notice that he is worse off, but comparing his lot to that of players who remain helpful, it appears to him that his choice was the correct one. Thus, that the simple decision rule is not self-sustaining does not imply that it would never be used by a population of sufficiently uninformed players.

Nevertheless, in many situations of economic interest, it is reasonable to assume that players have the capacity to apply either decision rule. In light of our results, and because of its intuitive appeal, it seems sensible to assume that such players use the clever decision rule. However, this paper has only compared performance of populations which are homogenous in their choice of decision rule. A better comparison of decision rules might be attempted by considering evolution in a population of players who are able to use the simple and clever decision rules, as well as other easily implemented rules. Evolution would occur on two levels: The distribution of strategies within the population would evolve according to the decision rules in the population, which would themselves evolve according to their relative performance. Such a model would provide harder evidence concerning the viability of the clever decision rule, and is a topic for future research.

## References


Foster, D., and Young, H. P. (1990). "Stochastic Evolutionary Game Dynamics." *Theoretical Population Biology*, 38, 219-232.

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

Rhode, P., and M. Stegeman (1996). "A Comment on 'Learning, Mutation, and Long Run Equilibria in Games'." *Econometrica*, 64, 443-449.

Samuelson, L. (1994). "Stochastic Stability in Games with Alternative Best Replies." *Journal of Economic Theory*, 64, 35-65.

Young, H. P. (1993). "The Evolution of Conventions." *Econometrica*, 61, 57-84.