Simple and Clever Decision Rules for a Model of Evolution

William H. Sandholm*
MEDS – KGSM
Northwestern University
Evanston, IL 60208-2001
e-mail: whs@nwu.edu

June 30, 1998
Forthcoming in *Economics Letters*.

Abstract

Under the decision rule specified by Kandori, Mailath, and Rob (1993), myopic adjustment can lead to surprising results, including coordination on strictly dominated strategies. We show that under an alternative decision rule, convergence to Nash equilibrium is guaranteed. Moreover, if rare mutations are introduced, risk dominant equilibria always correspond to long run equilibria.

Journal of Economic Literature Classification Numbers: C72, C73. Keywords: evolution, learning, decision rules, Markov chains, equilibrium selection.

* This paper is a revised version of Chapter 3 of my doctoral dissertation (Sandholm (1998)). I thank Eddie Dekel, Michihiro Kandori, George Mailath, Paul Rhode, Mark Stegeman, Jeroen Swinkels, and an anonymous referee for their comments, and gratefully acknowledge financial support from a National Science Foundation Graduate Fellowship and a State Farm Companies Foundation Dissertation Fellowship.

In their model of stochastic evolution, Kandori, Mailath, and Rob (1993) (henceforth KMR) show how myopia and rare mutations can lead to the selection of risk dominant equilibria in 2 x 2 symmetric coordination games. More recently, Rhode and Stegeman (1996) have shown that under the myopic decision rule specified by KMR, even in the absence of mutations, the evolutionary process can fail to converge to Nash equilibria or even dominant strategy equilibria. In this paper, we show how an alternative decision rule can eliminate these difficulties. Under this new rule, which seems at least as plausible in economic contexts as that specified by KMR, convergence to an approximate Nash equilibrium is guaranteed. Moreover, while in KMR risk dominant equilibria of coordination games sometimes fail to be selected when the population is small, we show that under the new decision rule risk dominant equilibria are selected regardless of the population size.

A population of N players, $N \geq 2$, is repeatedly randomly matched to play the 2 x 2 symmetric game displayed in Figure 1.

	s_1	s_2
s_1	a, a	b, c
s_2	c, b	d, d

Figure 1: A 2 x 2 Symmetric Game

We identify 2 x 2 symmetric games with vectors $G = (a, b, c, d) \in \mathbf{R}^4$. For generic games without a dominant strategy,

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

is the mass placed on strategy s_1 in the symmetric mixed strategy Nash equilibrium. If G is a coordination game ($a > c$, $b < d$), strategy s_1 is *risk dominant* if $\alpha^* \leq \frac{1}{2}$, while strategy s_2 is risk dominant if $\alpha^* \geq \frac{1}{2}$.

Evolution is modeled via a Markov process. The state of the population, $z \in Z = \{0, 1, \ldots, N\}$ is the number of players currently selecting strategy s_1. The expected payoffs that a player will receive in a random match when the state is z are given by

$$\pi_1(z) = \frac{(z-1)}{(N-1)}a + \frac{(N-z)}{(N-1)}b \qquad \text{for } z \in \{1, 2, \dots, N\},$$
$$\pi_2(z) = \frac{z}{(N-1)}c + \frac{(N-z-1)}{(N-1)}d \qquad \text{for } z \in \{0, 1, \dots, N-1\}.$$

These payoffs reflect that a player is never matched against himself.

Over the course of the interactions, players occasionally switch to strategies which perform well given the current state of the population. This updating may be split into two steps. First, it is determined whether a player gets an opportunity to update his strategy. We call the procedure which makes this determination a *selection mechanism*; the opportunity itself is called a *learning draw*. If the player receives the learning draw, he chooses whether to update his strategy according to a myopic evaluation criterion which we term a *decision rule*.

KMR specify deterministic dynamics whose limit behavior is essentially identical to that generated when players follow the *simple decision rule*.[1] According to this rule, players evaluate the strategies by comparing their current payoffs to those of an opponent currently playing the other strategy. If the opponent's payoffs are higher, the player switches. The rule is defined as follows:

When playing s_1 and $z \neq N$, switch if $\pi_1(z) < \pi_2(z)$.
When playing s_2 and $z \neq 0$, switch if $\pi_2(z) < \pi_1(z)$.[2]

Suppose that the current state is z, and consider a player following strategy s_2. He receives an expected payoff of $\pi_2(z)$. If he uses the simple decision rule, he decides whether to switch by comparing this expected payoff to $\pi_1(z)$. However, if the player switches to strategy s_1, while all of his opponents continue to play their old strategies, the state then becomes $z + 1$, and so his expected payoff becomes not $\pi_1(z)$, but $\pi_1(z + 1)$.

We therefore consider the *clever decision rule*. Under this criterion, a player evaluates strategies by comparing his current expected payoffs to those he would

1 KMR do not explicitly present a decision rule and a selection mechanism. Rather, they assume that play evolves according to a deterministic dynamic, b: $Z \to Z$, which is *Darwinian*: $sgn(\pi_1(z) - \pi_2(z)) = sgn(b(z) - z)$. Evolution under a Darwinian dynamic and evolution based on the simple decision rule lead to nearly equivalent limit behavior.

2 In states 0 and N, in which all players are following the same strategy, the comparison suggested above cannot be made. In these states the simple decision rule must be defined according to some other criterion. For two different approaches to this issue see Rhode and Stegeman (1996) and Sandholm (1996).

obtain if he and no one else were to switch strategies.[3] Formally, the clever decision rule is specified as follows:

When playing s_1, switch if $\pi_1(z) < \pi_2(z-1)$.
When playing s_2, switch if $\pi_2(z) < \pi_1(z+1)$.

The clever decision rule is a basic optimizing rule; to implement it, a player must know the payoff matrix, the population size, and the current state, and must be able to calculate the effect of a strategy change. In contrast, the simple decision rule is an imitative rule, requiring only that players be capable of recognizing opponents' success. While the simple decision rule is less stringent in both its informational and computational demands, its extreme simplicity is not without cost: for example, it may recommend the play of dominated strategies. Our results establish that a dividend of the added complexity of the clever decision rule is a guarantee of approximate Nash equilibrium play.

We consider the evolution of play under *Bernoulli selection*: in each round, each player independently with some fixed probability $\theta \in (0, 1]$ receives the *learning draw*. If $\theta = 1$, all players receive the learning draw in each period; this is known as the *best response dynamics*. If $\theta < 1$, then in each period, every subset of the players receives the learning draw with positive probability.

The selection mechanism and the decision rule define a Markov chain on the state space Z. We call this the *unperturbed process*, and identify it with its transition matrix $M \in Z \times Z$. A *recurrent class* of the unperturbed process is a minimal set of states $A \subseteq Z$ which, once entered, is never departed.[4] A *recurrent state* is a member of a recurrent class. The lone element of a singleton recurrent class is called a *stable state*; a stable state z satisfies $M(z, z) = 1$.

KMR complete their model by incorporating rare mutations. In each period, after myopic strategy adjustments have been made, each player independently with some fixed probability ε switches strategies. This generates the *perturbed process*, which we identify with its transition matrix, M^ε. Since mutations make transitions between any pair of states possible, the entire state space Z is the unique recurrent

3 A variant on the clever decision rule was first noted by KMR (footnote 20). Kandori and Rob (1995) apply this decision rule in their analysis of evolution in symmetric n x n games, but their selection results are proved under the assumption that the population size N is "large enough". Our results hold regardless of the population size.

4 Formally, a recurrent class is a maximal set of states A with the property that for all $x, y \in A$, there exists a sequence of states $x = z_0, z_1 \dots, z_k = y$ such that $M(z_{i-1}, z_i) > 0$ for all $i = 1, \dots, k$.

class of the perturbed process. The process therefore has a unique stationary distribution, μ^{ε}. The *long run equilibria* are the states which receive positive weight in the limiting stationary distribution: $LRE = \{z \in Z\colon \lim_{\varepsilon \to 0} \mu^{\varepsilon}(z) > 0\}$. When the rate of mutation is small, the population spends the vast majority of time visiting long run equilibria.

Long run equilibria are always recurrent states of the unperturbed Markov chain (Young (1993, Theorem 4)). Call a state $z \in Z$ an *approximate Nash equilibrium* if there is a symmetric Nash equilibrium of G which puts mass β on strategy s_1 and which satisfies $|z - N\beta| \leq 1$. If none of the recurrent states are approximate Nash equilibria, then neither are the long run equilibria; if all recurrent states are approximate Nash equilibria, then so are all long run equilibria.

A Nash equilibrium is defined by the property that no unilateral deviation is profitable. Under the simple decision rule, players may choose to switch strategies despite the absence of a profitable unilateral deviation; this observation underlies Rhode and Stegeman's (1996) examples. That a clever player wants to switch strategies precisely when a profitable unilateral deviation exists drives our first result.

Theorem 1: *Under Bernoulli selection with* $\theta < 1$ *and the clever decision rule, all recurrent states are approximate Nash equilibria.*

Proof: We prove this result for generic games; the proofs for the non-generic cases are similar. The requirement that $\theta < 1$ is only needed for games with no symmetric Nash equilibrium in pure strategies.[5]

First suppose that G possesses a strictly dominant strategy, and without loss of generality assume that it is s_1. In this case, $a > c$ and $b > d$. Then $\pi_1(z) > \pi_2(z - 1)$ for all $z \in Z - \{0\}$, so s_1 players never want to switch and s_2 players always want to switch. Hence, state N, in which all players choose s_1, is the only stable state.

Now suppose that G is a coordination game: $a > c$ and $b < d$. Define

[5] Conditions on a deterministic dynamic b: $Z \to Z$ which capture the clever decision rule are: $\pi_1(z) \geq \pi_2(z-1)$ & $\pi_2(z) \geq \pi_1(z+1) \Rightarrow b(z) = z$; $\pi_1(z) \geq \pi_2(z-1)$ & $\pi_2(z) < \pi_1(z+1)$ & $z \neq N \Rightarrow b(z) > z$; $\pi_1(z) < \pi_2(z-1)$ & $\pi_2(z) \geq \pi_1(z+1)$ & $z \neq 0 \Rightarrow b(z) < z$; $\pi_1(N) \geq \pi_2(N-1) \Leftrightarrow b(N) = N$; $\pi_2(0) \geq \pi_1(1) \Leftrightarrow b(0) = 0$. These conditions guarantee convergence to approximate Nash equilibria in games with strictly dominant equilibria and in coordination games. In games without a symmetric pure strategy equilibrium, in order to obtain results using a deterministic dynamic a condition which guarantees that the population does not perpetually overshoot the equilibrium (in the spirit of KMR's contraction condition (p. 48)) is also required.

$$x_1^* = \frac{N(d-b)+(a-c)}{(d-b)+(a-c)} = (N-1)\alpha^* + 1,$$
$$x_2^* = \frac{(N-1)(d-b)}{(d-b)+(a-c)} = (N-1)\alpha^*.$$

Observe that $x_1^* = x_2^* + 1$ and that $N\alpha^* \in (x_2^*, x_1^*)$. A calculation shows that s_1 players want to switch whenever $z < x_1^*$ and that s_2 players want to switch whenever $z > x_2^*$. Therefore, if play starts at $z \le x_2^*$ it must settle at state 0, and if play starts at $z \ge x_1^*$ it must settle at state N. Lastly, since $N\alpha^* \in (x_2^*, x_1^*)$, any state $\hat{z} \in (x_2^*, x_1^*)$, is an approximate Nash equilibrium. (As an aside, we observe that a state besides 0 or N can only be stable if $\theta = 1$, N is even, and $\frac{N}{2} \in (x_2^*, x_1^*)$, in which case state $\frac{N}{2}$ is stable.)

Finally, suppose that G is a game with no symmetric pure strategy equilibrium: $a < c$ and $b > d$. A calculation shows that s_1 players want to switch whenever $z > x_1^*$ and that s_2 players want to switch whenever $z < x_2^*$. Hence, all states in $[x_2^*, x_1^*]$ are stable; since $x_1^* = x_2^* + 1$, at least one such state exists. If $z < x_2^*$, only s_2 players want to switch, and since $\theta < 1$, single period transitions are possible to states z, $z + 1$, ..., and N. This list includes any states in the interval $[x_2^*, x_1^*]$. Similar reasoning for $z > x_1^*$ allows us to conclude that the states in $[x_2^*, x_1^*]$, are the only recurrent states. Since $N\alpha^* \in (x_2^*, x_1^*)$, all states in $[x_2^*, x_1^*]$ are approximate Nash equilibria. ■

The main result of KMR shows that in coordination games, the state in which all players choose the risk dominant strategy is a long run equilibrium. However, because of the difficulties with the simple decision rule described above, their result is only true for population sizes which are large enough given the underlying payoffs. Under the clever decision rule, this restriction is unnecessary, as our final result shows.

Theorem 2: *Fix a coordination game G and a population size N. Under Bernoulli selection and the clever decision rule, any state in which all players select the risk dominant strategy is a long run equilibrium.*

Proof: Let B_N be the basin of attraction of state N: the set of states (including state N) starting from which unperturbed play must return to state N. Similarly, let B_0 be the basin of attraction of state 0. (If $\theta = 1$, N is even, and $\frac{N}{2} \in (x_2^*, x_1^*)$, $\frac{N}{2}$ is also a stable state, but its basin of attraction is just $\{\frac{N}{2}\}$.) Applying Theorem 1 of KMR, it is easily verified that state $z \in \{0, N\}$ is a long run equilibrium if and only if $|B_z|$ = max

$\{|B_0|, |B_N|\}$: states with the biggest basins of attractions are long run equilibria.[6] Recalling the proof of Theorem 1 above, we see that $|B_N|$ = [N – x_1*] + 1 = [(N – 1)(1 – α*)] + 1 and $|B_0|$ = [x_2*] + 1 = [(N – 1)α*] + 1, where [·] represents the greatest integer function. Hence, if s_1 is risk dominant, $\alpha^* \leq \frac{1}{2}$ and state N is a long run equilibrium, and if s_2 is risk dominant, $\alpha^* \geq \frac{1}{2}$ and state 0 is a long run equilibrium. ■

If N is odd, the converse of Theorem 2 also holds: any long run equilibrium entails coordination on a risk dominant strategy. In contrast, if N is even, both unanimous states are long run equilibria whenever $\left|\alpha^* - \frac{1}{2}\right| < \frac{1}{2(N-1)}$: in this case, if the mixed strategy equilibrium is close enough to $\frac{1}{2}$, both unanimous states have basins of attraction with cardinality $\frac{N}{2}$ and are therefore long run equilibria. Finally, if N = 2 and θ = 1 all three states are long run equilibria.

References

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

Kandori, M., and Rob, R. 1995. "Evolution of Equilibria in the Long Run: A General Theory and Applications." Journal of Economic Theory 65, 383-414.

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

Sandholm, W. H. 1996. "Simple and Clever Decision Rules in Single Population Evolutionary Models," CMS-EMS Discussion Paper #1158, Northwestern University.

Sandholm, W. H. 1998. Evolutionary Justification of Nash Equilibrium. Ph.D. dissertation, Northwestern University.

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

---

[6] In the terminology of KMR, the costs of states 0 and N are $|B_N|$ and $|B_0|$ respectively if no stable interior state exists; if there is a stable interior state, then its cost is N, and the costs of states 0 and N are $|B_N| + 1 \leq N$ and $|B_0| + 1 \leq N$. Since the long run equilibria are the cost minimizing states, the claim follows.