

Almost Global Convergence to p -Dominant Equilibrium*

William H. Sandholm
Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706
whs@ssc.wisc.edu
www.ssc.wisc.edu/~whs

November 7, 2000

Abstract

A population of players repeatedly plays an n strategy symmetric game. Players update their strategies by sampling the behavior of k opponents and playing a best response to the distribution of strategies in the sample. Suppose the game possesses a $\frac{1}{k}$ -dominant strategy which is initially played by a positive fraction of the population. Then if the population size is large enough, play converges to the $\frac{1}{k}$ -dominant equilibrium with arbitrarily high probability.

Key words: evolutionary game theory, coordination games

* I thank Bart Lipman, Larry Samuelson, two anonymous referees, and an anonymous Associate Editor for their comments.

1. Introduction

Multiplicity of equilibria creates fundamental difficulties for predicting behavior in non-cooperative games. These difficulties are most conspicuous in coordination games. Since strict equilibria satisfy virtually all equilibrium refinements proposed in the literature, games with multiple strict equilibria do not admit a single, obviously correct prediction of play.

In this paper, we address this issue using a simple evolutionary model. In our model, a fixed population of players repeatedly plays an n strategy symmetric game, G . Players myopically adjust their strategy choices in response to their opponents' current behavior.

Many models of evolution assume that players know the current population state when deciding how to act. However, one might expect precise information about the population's behavior to be difficult or costly to obtain. For this reason, it seems more consistent with the assumption of myopia to have players base their decisions on limited information about opponents' play. To capture this notion as simply as possible, we assume that when a player receives an opportunity to update his behavior, he draws a sample of $k \geq 2$ players from the population. He then chooses a best response to the distribution of behavior within his sample, viewing it as representative of the behavior of the population as a whole.

Surprisingly, this model can generate unique predictions of play, even in games with multiple strict equilibria. Suppose that G has a $\frac{1}{k}$ -dominant strategy which is initially played by a positive fraction of the population. We establish that if the population size is large enough, play converges to the $\frac{1}{k}$ -dominant equilibrium with arbitrarily high probability. Thus, in some coordination games, when players have limited information about opponents' behavior, almost global convergence to a single equilibrium is virtually guaranteed.

Our model does not offer a unique prediction of play in every coordination game: in generic 2×2 games, a unique prediction is guaranteed only if the sample size is two. Still, it seems natural to assume that the sample size k is small; while this does not ensure a unique prediction in every game, it does create unique predictions for many games. Moreover, our conclusions for games which do possess a $\frac{1}{k}$ -dominant equilibrium are quite strong: populations learn to play the predicted equilibrium from nearly all initial conditions, even when other strict equilibria are available.

To shed more light on our conclusions, we find it useful to contrast them with those that have been obtained using models of stochastic stability (e.g., Kandori, Mailath, and Rob (1993) and Young (1993a)). These models introduce rare mutations to some underlying evolutionary process in order to establish the existence of a unique stochastically stable equilibrium. By definition, such an equilibrium is very likely to be played by the population in the sufficiently distant future.

The uniqueness of the stochastically stable prediction relies on the ergodicity of the perturbed evolutionary process. Unfortunately, ergodicity brings with it the undesirable consequence that the predicted equilibrium is played and then abandoned an infinite number of times. In addition, if a strict equilibrium which is not stochastically stable is reached first, this equilibrium should be expected to persist; in some specifications, the expected amount of time before the equilibrium is departed grows exponentially in the population size.¹

In our model, convergence occurs with high probability from nearly all initial conditions; becoming stuck at the wrong equilibrium is much less of an issue. Convergence is also permanent: once the population reaches the $\frac{1}{k}$ -dominant equilibrium, it never departs. Finally, while the large population sizes which are natural to consider in evolutionary contexts can create tensions within stochastic stability models, they make our results easier to prove. However, we shall see that our conclusions are not dependent on having a very large population: the predictions of our model still have force when the population is moderate in size.

2. The Theorem

Let $G = \left\{ \{s_i\}_{i=1}^n, \{\pi_i\}_{i=1}^n \right\}$ be an n strategy symmetric game. If Δ denotes the simplex in \mathbf{R}^n , then each $\pi_i: \Delta \rightarrow \mathbf{R}$ represents the payoffs to strategy s_i as a function of the proportion of players choosing each strategy. Following Morris, Rob, and Shin (1995), we say that strategy s_i is *p-dominant* if $\pi_i(x) > \pi_j(x)$ for all $j \neq i$ whenever $x \in \Delta$ satisfies $x_i \geq p$. Thus, a strategy is *p-dominant* if it is the unique best response when it is played by at least proportion p of the population. Observe that increasing p makes *p-dominance* less demanding: if $p < q$, any *p-dominant* strategy is also *q-dominant*.

¹ See, for example, Ellison (1993).

The standard evolutionary framework considers populations of players who are repeatedly randomly matched to play a normal form game. In the current context, this corresponds to the case in which each π_i is linear: $\pi_i(\mathbf{x}) \equiv (\Pi\mathbf{x})_i$ for some payoff matrix $\Pi \in \mathbf{R}^{n \times n}$.² We call a normal form game an $n \times n$ *coordination game* if each of the n strategies constitutes a strict equilibrium of the game: that is, if $\Pi_{ii} > \Pi_{ji}$ for all i and all $j \neq i$. A $\frac{1}{2}$ -dominant strategy of a 2×2 coordination game is called *strictly risk dominant*; such a strategy is a strict best response against an opponent who is equally likely to play either strategy.

Our results will concern games with $\frac{1}{k}$ -dominant strategies, where $k \geq 2$ is an integer. For examples of these games, consider the class of $n \times n$ *pure coordination games*, which satisfy $\Pi_{ii} > 0$ for all i and $\Pi_{ij} = 0$ for all pairs ij with $i \neq j$. Suppose that $\Pi_{11} = c$ and that $\Pi_{ii} \in (0, 1]$ for all $i \neq 1$. Then it is easily verified that strategy s_1 is $\frac{1}{k}$ -dominant whenever $c > k - 1$. If in addition $\Pi_{ii} = 1$ for some $i \neq 1$, the converse statement also holds.

We model the evolution of play as a Markov process. The game G is played repeatedly by a population of N players. The random vector $X_t^N = (X_t^{N,1}, \dots, X_t^{N,n})$ represents the number of players choosing strategies s_1 through s_n at times $t = 0, 1, \dots$; by definition, $\sum_i X_t^{N,i} = N$ at all times t . The initial condition $X_0^N \in \mathbf{N}_0^n$ is given. During each period, one player is chosen at random from the population and given the opportunity to revise his strategy choice.³

In many models of evolution (e.g., Kandori, Mailath, and Rob (1993)), it is assumed that the player granted the revision opportunity learns the exact distribution of strategies in the population and plays a best response to this distribution. Under this specification, there is an absorbing state with a non-negligible basin of attraction corresponding to each strict equilibrium of the game.

In many settings where evolutionary models are appropriate, it may be more natural to assume that players have limited information about opponents' behavior. To capture this, we suppose that the player who receives the revision opportunity randomly samples $k \geq 2$ players from the population and learns their

² In making this comparison with normal form games we implicitly assume that players can be randomly matched against themselves. When the population size is large, the effect of forbidding self-matching become negligible; hence, ruling out self-matching would not alter our main result.

³ All of our results would continue to hold if players' revision opportunities instead arrived via independent Poisson processes.

behavior. He then plays a best response to the distribution of strategies in the sample.⁴

Clearly, all strict equilibria of G correspond to absorbing states of this Markov process. But because of the randomness in the sampling procedure, the population can converge to any of these absorbing states from any interior initial condition. Nevertheless, we are able to show that if one equilibrium is $\frac{1}{k}$ -dominant, a large population is nearly certain to converge to it from nearly all initial conditions.

We call strategy s_i *asymptotically almost globally stable* if for all $\delta > 0$,

$$\lim_{N \rightarrow \infty} P\left(\lim_{t \rightarrow \infty} X_t^{N,i} = N \mid X_0^{N,i} \geq N\delta\right) = 1.$$

In words: s_i is asymptotically almost globally stable if for any positive δ and ε , there exists an $N = N(\delta, \varepsilon)$ with the following property: if the population size exceeds N and at least proportion δ of the population initially plays s_i , then the probability that play converges to the state in which all players choose s_i exceeds $1 - \varepsilon$.

We are now able to state our main result.

Theorem: *If s_i is $\frac{1}{k}$ -dominant, it is asymptotically almost globally stable.*

Corollary 1: *If $k = 2$, G is a 2×2 coordination game, and s_i is strictly risk dominant, then s_i is asymptotically almost globally stable.*

We first offer intuition for the case in which G has exactly two strategies: s_1 , which is $\frac{1}{k}$ -dominant, and s_2 , which is also a strict equilibrium. Suppose that the current proportion of s_1 players is x . Let $D(x)$ be the *draw rate* for strategy s_1 : the probability that the player who is given the opportunity to switch strategies is an s_1 player. Clearly, $D(x) = x$. Similarly, let $C(x)$ be the *choice rate* for strategy s_1 : the probability that the player given the chance to switch chooses strategy s_1 . Since s_1 is $\frac{1}{k}$ -dominant, and since the sample size is k , strategy s_2 will only be chosen when all

⁴ For simplicity, we assume that samples are drawn with replacement; allowing sampling without replacement would not alter our main result. In the event that a player has multiple best responses to a particular sample, his behavior can be specified arbitrarily.

Models of evolution with incomplete sampling have also been considered by Young (1993a, 1993b), Hurkens (1995), Kaniowski and Young (1995), and Sáez-Martí and Weibull (1999). In these models, players choose best responses to an incomplete memory of past play rather than an incomplete sample of current behavior.

k players queried choose s_2 . Therefore, the probability that s_1 is chosen is $C(x) = 1 - (1 - x)^k$. We can compute the expected change in the number of s_1 players as

$$(1 - D(x))C(x) - D(x)(1 - C(x)) = C(x) - D(x) = (1 - x) - (1 - x)^k,$$

which is strictly positive whenever $x \in (0, 1)$.⁵ Therefore, whenever the population is not at an absorbing state, the expected change in the number of s_1 players is positive. But when the population size is sufficiently large, the course of evolution is almost completely governed by the expected direction of motion. We are therefore able to show that for any positive initial proportion of s_1 players, if the population size is large enough, convergence to the state in which all players choose s_1 is virtually guaranteed.

Now suppose there are more than two strategies, but that s_1 is still $\frac{1}{k}$ -dominant. While the draw rate is unchanged, the choice rate can now depend on the distribution of behavior among strategies besides s_1 . However, since s_1 is $\frac{1}{k}$ -dominant, the choice rate must be at least $1 - (1 - x)^k$. Thus, the expected change in the number of s_1 players remains positive. We can therefore still establish asymptotically almost global convergence to the $\frac{1}{k}$ -dominant equilibrium.⁶

Our definition of stability contains a limit in the population size N . This allows us to consider arbitrarily large populations when establishing convergence probabilities close to 1. How large a population do we actually need? To address this question, we let s_1 be a $\frac{1}{k}$ -dominant strategy, and define the population bound \underline{N} by

$$\underline{N}(\delta, \varepsilon) = \min\{M : \forall N \geq M, P\left(\lim_{t \rightarrow \infty} X_t^{N,i} = N \mid X_0^{N,i} \geq N\delta\right) > 1 - \varepsilon\}.$$

Suppose that at the onset of play, at least proportion δ of the population plays strategy s_1 . Then if the population size is at least $\underline{N}(\delta, \varepsilon)$, convergence to the $\frac{1}{k}$ -dominant equilibrium will occur with a probability of at least $1 - \varepsilon$.

We now show that this bound grows slowly as δ and ε become small.

⁵ That $C(x) \geq D(x)$ is observed by Chu (1993) in an evolutionary model of law enforcement.

⁶ It is worth noting that when $k = 1$, evolution is essentially random. For example, when G is a coordination game, the draw and choice rates for each strategy are always equal, and each component of the evolutionary process is a martingale.

Corollary 2: *The population bound $\underline{N}(\delta, \varepsilon)$ satisfies*

$$\underline{N}(\delta, \varepsilon) \leq \max \left\{ \frac{2 \ln \left(\frac{4}{\delta \varepsilon} \right)}{\delta \ln k}, \frac{8}{\delta} \right\} + 1.$$

Thus, for each fixed ε , $\underline{N}(\delta, \varepsilon) \in O(\delta^{-1} \ln \delta^{-1})$, and for each fixed δ , $\underline{N}(\delta, \varepsilon) \in O(\ln \varepsilon^{-1})$.

Corollary 2 shows that the population sizes needed for our predictions to be relevant are not large, even if few players initially choose the $\frac{1}{k}$ -dominant strategy, and even if we demand a probability of convergence very close to 1. In particular, as we lower the probability ε of a failure to converge, the bound \underline{N} only grows as the logarithm of ε^{-1} : small populations can learn to play the $\frac{1}{k}$ -dominant equilibrium with very high probabilities.

The proof of these results implicitly defines an algorithm for computing convergence probabilities, the running time of which is linear in the population size. Using this algorithm, we can determine exact population bounds $\underline{N}(\delta, \varepsilon)$ for specific choices of δ and ε . In Table 1, we present bounds for the case in which $k = 2$; these bounds are valid when the game G has a $\frac{1}{2}$ -dominant equilibrium and the sample size is 2. The population sizes needed to ensure convergence to the $\frac{1}{2}$ -dominant equilibrium are quite small. For example, if at least 5% of the population initially chooses the $\frac{1}{2}$ -dominant strategy, the probability of a failure to converge to the $\frac{1}{2}$ -dominant equilibrium is less than one in a million whenever the population size is at least 381.

Table 2 contains population bounds for the case in which $k = 5$, which are valid when the game G has a $\frac{1}{5}$ -dominant equilibrium and the sample size is 5. Fewer games have $\frac{1}{5}$ -dominant equilibria than have $\frac{1}{2}$ -dominant equilibria. But in those which do, convergence to the equilibrium is robust to larger sample sizes. Moreover, as the tables illustrate, the larger samples allow us to guarantee convergence to the equilibrium in smaller populations.⁷

⁷ That most of the numbers in the tables have last digit 1 is a consequence of the discreteness of the state space. For example, if $N = 100$, there is a state at which exactly 10% of the population plays strategy s_1 , while if $N = 101$, all states at which at least 10% play s_1 actually have at least $11/101 \approx 10.89\%$ playing s_1 . For this reason, when $\delta = .10$ the probabilities of failures to converge drop discretely after each multiple of 10, leading the population bounds to occur at these points.

		ε				
		10^{-1}	10^{-2}	10^{-4}	10^{-6}	10^{-8}
δ	.25	9	21	45	69	93
	.10	31	61	121	191	251
	.05	61	121	241	381	521
	.01	301	601	1301	1901	2601

Table 1: Population bounds $\underline{N}(\delta, \varepsilon)$ when $k = 2$.

		ε				
		10^{-1}	10^{-2}	10^{-4}	10^{-6}	10^{-8}
δ	.25	5	9	17	29	33
	.10	11	21	51	71	101
	.05	21	41	101	161	221
	.01	101	201	501	801	1101

Table 2: Population bounds $\underline{N}(\delta, \varepsilon)$ when $k = 5$.

3. The Proof

As above, we begin with the case in which G has exactly two strategies: a $\frac{1}{k}$ -dominant strategy s_1 , and a strict equilibrium strategy s_2 . We find it convenient to speak in terms of the number of players who are *not* playing the risk dominant strategy: define $Y_t^N \equiv X_t^{N,2}$ to be the number of players choosing strategy s_2 . Let $Z^N = \{0, 1, \dots, N\}$. Since s_1 and s_2 are both strict best responses to themselves, $\{Y_t^N\}_{t=0}^\infty$ is a Markov chain on Z^N whose only absorbing states are 0 and N . We want to show that Y_t^N is absorbed at state 0 with high probability.

For the number of players choosing strategy s_2 to increase, the player who is given the opportunity to switch strategies must initially be playing s_1 , and he must draw a sample which consists solely of s_2 players. Hence, for $z \in \{0, \frac{1}{N}, \dots, 1\}$,

$$P(Y_{t+1}^N = Y_t^N + 1 \mid Y_t^N = Nz) = (1 - z) z^k \equiv p(z).$$

Similarly, for Y_t^N to fall, the player given the chance to switch must be playing s_2 , and his sample must contain at least one player choosing s_1 . Therefore,

$$P(Y_{t+1}^N = Y_t^N - 1 \mid Y_t^N = Nz) = z(1 - z^k) \equiv q(z).$$

We now define three functions which will prove useful. Define $\rho: (0, 1) \rightarrow \mathbf{R}$ by

$$\rho(z) = \frac{q(z)}{p(z)} = \frac{z(1 - z^k)}{(1 - z)z^k} = z^{1-k} \sum_{l=0}^{k-1} z^l = \sum_{l=0}^{k-1} z^{-l}.$$

Since $k \geq 2$, $\rho(\cdot)$ is decreasing and $\lim_{z \uparrow 1} \rho(z) = k$. Next, for $m \in \{1, \dots, N - 1\}$ define

$$\pi^N(m) = \prod_{j=1}^m \rho\left(\frac{j}{N}\right).$$

Clearly, $\pi^N(\cdot)$ is increasing in m . Finally, for $y \in Z^N$ define

$$\phi^N(y) = \begin{cases} 0 & \text{if } y = 0, \\ 1 & \text{if } y = 1, \\ 1 + \sum_{m=1}^{y-1} \pi^N(m) & \text{if } y > 1. \end{cases}$$

Notice that $\phi^N(\cdot)$ is strictly increasing.

A standard result on birth and death chains (Durrett (1991, Theorem 5.3.7)) tells us that if $Y_0^N = y$, the probability that Y_t^N converges to N is equal to $\phi^N(y)/\phi^N(N)$. It will prove useful later on to see why this is so. The function $\phi^N(\cdot)$ was constructed in such a way that $\phi^N(Y_t^N)$ is a martingale: 0 and N are absorbing states of Y_t^N , and it is easily checked that for all $y \in \{1, \dots, N - 1\}$,

$$\begin{aligned} \phi^N(y) &= p\left(\frac{y}{N}\right)\phi^N(y+1) + \left(1 - p\left(\frac{y}{N}\right) - q\left(\frac{y}{N}\right)\right)\phi^N(y) + q\left(\frac{y}{N}\right)\phi^N(y-1) \\ &= E\left(\phi^N(Y_{t+1}^N) \mid Y_t^N = y\right) \\ &= E\left(\phi^N(Y_{t+1}^N) \mid \phi^N(Y_t^N) = \phi^N(y)\right). \end{aligned} \tag{1}$$

Let T_0 and T_N be the (random) times that Y_t^N hits 0 and N respectively, and let $T = \min(T_0, T_N)$. Since $\phi^N(Y_t^N)$ is bounded and T is almost surely finite, the Optional Stopping Theorem (Durrett (1991, Theorem 4.7.4)) tells us that

$$\phi^N(y) = E\left(\phi^N(Y_T^N) \mid Y_0^N = y\right) = \phi^N(N) P(T_N < T_0).$$

Rearranging this equation yields the desired result.

Fix $\delta, \varepsilon > 0$. It is enough to show that if $y \leq N(1 - \delta)$ and

$$N \geq \max \left\{ \frac{2 \ln \left(\frac{4}{\delta \varepsilon} \right)}{\delta \ln k}, \frac{8}{\delta} \right\},$$

then $\phi^N(y) / \phi^N(N) < \varepsilon$. Define

$$\begin{aligned} y_a^N &= \max\{y \in Z^N: y \leq N(1 - \delta)\}; \\ y_b^N &= \min\{y \in Z^N: y \geq N(1 - \frac{\delta}{2})\}. \end{aligned}$$

Observe that $y_b^N < N(1 - \frac{\delta}{2}) + 1 < N(1 - \frac{\delta}{4})$. Hence,

$$\frac{y_a^N}{N - y_b^N - 1} \leq \frac{N}{N - (N(1 - \frac{\delta}{2}) + 1) - 1} = \frac{1}{\frac{\delta}{2} - \frac{2}{N}} = \frac{2}{\delta - \frac{4}{N}} \leq \frac{4}{\delta}.$$

Furthermore, since $\rho(\cdot)$ is decreasing and $\lim_{z \uparrow 1} \rho(z) = k$, we see that

$$\frac{\pi^N(y_b^N)}{\pi^N(y_a^N)} = \prod_{j=y_a^N+1}^{y_b^N} \rho\left(\frac{j}{N}\right) \geq \rho\left(1 - \frac{\delta}{4}\right)^{y_b^N - y_a^N} \geq \rho\left(1 - \frac{\delta}{4}\right)^{\frac{N\delta}{2}} > k^{\frac{N\delta}{2}} \geq \frac{4}{\delta \varepsilon}.$$

Therefore, since $\phi^N(\cdot)$ is increasing and positive and since $\pi^N(\cdot)$ is increasing,

$$\begin{aligned} \frac{\phi^N(y)}{\phi^N(N)} &< \frac{\phi^N(y_a^N)}{\phi^N(N) - \phi^N(y_b^N)} \\ &= \frac{1 + \sum_{m=1}^{y_a^N-1} \pi^N(m)}{\sum_{m=y_b^N}^{N-1} \pi^N(m)} \\ &\leq \frac{y_a^N \pi^N(y_a^N)}{(N - y_b^N - 1) \pi^N(y_b^N)} \\ &\leq \left(\frac{4}{\delta}\right) \left(\frac{\delta \varepsilon}{4}\right) = \varepsilon. \end{aligned}$$

This completes the proof of the case in which G is a two strategy coordination game.

We now consider the general case. Suppose that strategy s_1 is $\frac{1}{k}$ -dominant, and let $\hat{Y}_t^N \equiv N - X_t^{N,1}$ be the number of players who are playing strategies other than s_1

at time t . Since s_1 is a strict equilibrium, it is clear that 0 is an absorbing state of \hat{Y}_t^N , and that no state in $\{1, \dots, N-1\}$ is absorbing. We need to show that \hat{Y}_t^N is absorbed at 0 with high probability.

It is worth noting as an aside that \hat{Y}_t^N need not be a Markov chain. The probability that a player's sample causes him to switch to s_1 can depend on the distribution of the population's behavior among strategies besides s_1 . Past values of \hat{Y}_t^N can contain information about this distribution for which the current value of \hat{Y}_t^N is not a sufficient statistic. Therefore, transition probabilities for \hat{Y}_t^N that condition only on the value of \hat{Y}_t^N in the current period will typically differ from transition probabilities which also condition on the values in past periods.

To prove the result, we will establish that $\phi^N(\hat{Y}_t^N)$ is a supermartingale. We can then define \hat{T}_0 and \hat{T}_N to be the times that \hat{Y}_t^N hits 0 and N and let $\hat{T} = \min(\hat{T}_0, \hat{T}_N)$. Then applying the Optional Stopping Theorem yields

$$\phi^N(y) \geq E\left(\phi^N(\hat{Y}_{\hat{T}}^N) \mid \hat{Y}_0^N = y\right) = \phi^N(N) P(\hat{T}_N < \hat{T}_0).$$

Thus, $P(\hat{T}_N < \hat{T}_0) \leq \phi^N(y) / \phi^N(N)$, and the result follows from the analysis above.

To show that $\phi^N(\hat{Y}_t^N)$ is a supermartingale, we first observe that since strategy s_1 is $\frac{1}{k}$ -dominant, a player given the chance to switch strategies will choose to play s_1 if at least one member of his sample plays s_1 . (He may also choose s_1 even if no one in his sample chooses s_1 ; whether he does so depends on how the members of his sample are distributed among the other strategies.) In any case, it is clear that if $z \in \{0, \frac{1}{N}, \dots, 1\}$,

$$\hat{p}_t(z) \equiv P(\hat{Y}_{t+1}^N = \hat{Y}_t^N + 1 \mid \hat{Y}_t^N = Nz) \leq (1 - z) z^k = p(z).$$

Similarly, a player will certainly choose s_1 if his sample contains at least one player playing s_1 . Therefore,

$$\hat{q}_t(z) \equiv P(\hat{Y}_{t+1}^N = \hat{Y}_t^N - 1 \mid \hat{Y}_t^N = Nz) \geq z(1 - z^k) \equiv q(z).$$

Thus, applying equation (1) and utilizing the fact that $\phi^N(\cdot)$ is strictly increasing, we find that if $y \in \{1, \dots, N-1\}$,

$$\phi^N(y) = p\left(\frac{y}{N}\right)\phi^N(y+1) + \left(1 - p\left(\frac{y}{N}\right) - q\left(\frac{y}{N}\right)\right)\phi^N(y) + q\left(\frac{y}{N}\right)\phi^N(y-1)$$

$$\begin{aligned}
&\geq \hat{p}_t\left(\frac{y}{N}\right)\phi^N(y+1) + \left(1 - \hat{p}_t\left(\frac{y}{N}\right) - q\left(\frac{y}{N}\right)\right)\phi^N(y) + q\left(\frac{y}{N}\right)\phi^N(y-1) \\
&\geq \hat{p}_t\left(\frac{y}{N}\right)\phi^N(y+1) + \left(1 - \hat{p}_t\left(\frac{y}{N}\right) - \hat{q}_t\left(\frac{y}{N}\right)\right)\phi^N(y) + \hat{q}_t\left(\frac{y}{N}\right)\phi^N(y-1) \\
&= E\left(\phi^N\left(\hat{Y}_{t+1}^N\right) \middle| \hat{Y}_t^N = y\right) \\
&= E\left(\phi^N\left(\hat{Y}_{t+1}^N\right) \middle| \phi^N\left(\hat{Y}_t^N\right) = \phi^N(y)\right).
\end{aligned}$$

In addition, since \hat{Y}_t^N can never exceed N ,

$$E\left(\phi^N\left(\hat{Y}_{t+1}^N\right) \middle| \phi^N\left(\hat{Y}_t^N\right) = \phi^N(N)\right) \leq \phi^N(N),$$

and since 0 is an absorbing state of \hat{Y}_t^N ,

$$E\left(\phi^N\left(\hat{Y}_{t+1}^N\right) \middle| \phi^N\left(\hat{Y}_t^N\right) = \phi^N(0)\right) = \phi^N(0).$$

This establishes that $\phi^N(\hat{Y}_t^N)$ is a supermartingale, completing the proof of the theorem.

References

- Chu CYC (1993) Oscillatory vs. stationary enforcement of law. *International Review of Law and Economics* 13: 303-315.
- Durrett R (1991) *Probability: theory and examples*. Brooks/Cole, Belmont, CA.
- Ellison G (1993) Learning, local interaction, and coordination. *Econometrica* 61: 1047-1071.
- Hurkens S (1995) Learning by forgetful players. *Games and Economic Behavior* 11: 304-329.
- Kandori M, Mailath GJ, Rob R (1993) Learning, mutation, and long run equilibria in games. *Econometrica* 61: 29-56.
- Kaniovski Y, Young HP (1995) Learning dynamics in games with stochastic perturbations. *Games and Economic Behavior* 11: 330-363.
- Morris S, Rob R, Shin HS (1995) p -dominance and belief potential. *Econometrica* 63: 145-157.
- Sáez-Martí M, Weibull JW (1999) Clever agents in Young's evolutionary bargaining model. *Journal of Economic Theory* 86: 268-279.
- Young HP (1993a) The evolution of conventions. *Econometrica* 61: 57-84.
- Young HP (1993b) An evolutionary model of bargaining. *Journal of Economic Theory* 59: 145-168.