Experimentation, Imitation, and Stochastic Stability

Douglas Gale,*
Department of Economics, New York University, and
Robert W. Rosenthal,*
Department of Economics, Boston University

Revised June 1998

Abstract

Do boundedly rational agents repeatedly playing a symmetric game with a unique symmetric equilibrium learn over time to play it? In this paper we model the dynamic interaction of two types of such agents, experimenters and imitators, whose behavior is characterized by simple rules of thumb. We find that the stochastic process describing their play is stable in the large: it converges globally and with probability one to a compact neighborhood of the equilibrium. However, its local behavior near the equilibrium depends in interesting ways on the details of the model.

1 Introduction

In a changing world, economic agents are constantly adapting their behavior. Adaptation can take a number of forms. In this paper, we study a model of

*We are indebted to Yossi Feinberg, Daniel Goroff, J. Michael Harrison, Glenn Loury, Ramon Marimon, Michael Manove, Debraj Ray, Peter Sorensen, Charles Wilson, referees, an associate editor, and participants at many seminars for helpful comments; to Bhaskar Chattaraj, Xianghai Ma, and Sumon Majumdar for invaluable research assistance; and to the National Science Foundation for financial support.
boundedly rational, adaptive behavior in a dynamic, strategic context. The interest of the model lies in the complex dynamics that are produced by the interaction of two apparently simple behavioral rules, experimentation and imitation.

Agents are assumed to have little direct knowledge of the environment in which they interact, but they can learn about it both directly from their own experiences and indirectly by observing the behavior of other agents who are in the same situation. When agents experiment, they obtain information directly. When they imitate, they are implicitly hoping that the agents they copy are already well adapted. While a sensible agent might engage in both experimentation and imitation, we simplify here by assuming that each agent specializes in exactly one of these activities. This could be justified perhaps by assuming fixed costs of switching between activities. We simplify further by assuming that there is only one experimenter and that the rest of the agents are imitators.

At each date \( t = 1, 2, \ldots \) the agents play a symmetric normal-form game with the payoff function for agent \( i \)

\[-(q_i - B\bar{q}_{-i})^2,\]

where \( q_i \in \mathbb{R}^1 \) is the action of agent \( i \) and \( \bar{q}_{-i} \) is the average action of the other agents. \( B \) is the slope of each player’s best-response line. When \( B < 0 \), agents’ actions are strategic substitutes, as in the standard Cournot model; when \( B > 0 \), actions are strategic complements. When \( B \neq 1 \), the game has a unique symmetric equilibrium at which all agents choose 0.\(^1\)

In each period, the agents observe their own actions and payoffs and the actions, but not the payoffs, of the other agents. The imitators adjust their actions a constant fraction \( \lambda \) of the distance between their previous actions and the average action of the other agents. Since it is only the average of the imitators’ actions that matters and their decision rule is linear, there is no loss of generality in replacing them with a representative agent.

The experimenter searches randomly for a better strategy. If agent \( i \) is the experimenter and he chose a strategy \( q_{it-1} \) last period, then he “tests” a new strategy that is uniformly distributed on the interval \([q_{it-1} - 1, q_{it-1} + 1]\). Note that this interval is centered on last period’s action and that its size does not change over time. If the randomly drawn strategy falls in the better-response

\(^1\)When \( B = -1 \) and the number of agents is even, there is also an asymmetric equilibrium at which half the agents choose one action and the other half choose its negative.
set
\[ \{q \mid (q - B\bar{q}_{-it})^2 \leq (q_{it-1} - B\bar{q}_{-it})^2 \}, \]
then he adopts it as \( q_{it} \). Otherwise, he puts \( q_{it} = q_{it-1} \). (We are assuming for simplicity that the experimenter can test against the imitators’ current actions and does not infect their future actions in the event of failure. Our results go through under alternative specifications.)

Denote by \( X_t \) the experimenter’s action at \( t \) and by \( Y_t \) the average of the imitators’ actions at \( t \). Then, under the behavioral rules specified above, for any initial condition \((x_0, y_0)\) these behavioral rules define a Markov chain \( \{(X_t, Y_t)\} \) having state space \( \mathbb{R}^2 \).

This model has a number of interesting dynamic properties which are explored in Section 2.

- First, it is stable in the large: from any initial state \((x_0, y_0)\) the chain converges with probability one to a compact neighborhood of the origin. (For this result we need to assume only that \( B < 1 \); when \( B > 1 \), it is easy to see that imitation and experimentation feed off each other in an explosive way. For the case \( B = 1 \), there is a continuum of symmetric equilibria, each being a kind of weak local attractor of the chain.)

- Secondly, for \( B < 0 \) and sufficiently small, the symmetric equilibrium is unstable in the small: for any sufficiently small neighborhood of the origin and any initial condition \((x_0, y_0)\) in that neighborhood \((x_0 \neq 0)\), the chain leaves the neighborhood with probability one.

- Finally, however, we show that when the symmetric equilibrium is unstable in the small, it is not too unstable: for any neighborhood of the equilibrium, however small, the probability of the chain being in the given neighborhood at date \( t \) converges to one as \( t \) approaches \( \infty \).

At the macro level, stability in the large tells us that these adaptive rules do work, by bringing the agents’ actions to a compact neighborhood of the equilibrium where they are approximately optimal. On the other hand, at the micro level, instability in the small tells us that the agents can never learn the equilibrium strategies exactly. Their actions fluctuate permanently around the equilibrium levels. How large these fluctuations are depends on the parameters \( \lambda \) and \( B \) and the size of the search window (here normalized to two).
The third result is puzzling, since it appears to contradict the second. The two results can be reconciled, but the explanation depends on another feature of the model, namely, that the chain \( \{(X_t, Y_t)\} \) is null recurrent. Although the chain leaves sufficiently small neighborhoods of the origin with probability one, the expected time this takes is infinite. The reason is that for states \((x, y)\) very close to the origin the better-response set is very small. It is hard to find a better response and, as a result, the chain changes very slowly. If an econometrician looked at the cross-sectional frequency distribution of a set of sample paths, he might conclude that the process was converging. This is a result of the fact that almost every path occasionally comes close to the origin and then takes a very long time to get away. If the same observer looked at a single sample path, he might come to a very different conclusion.

In certain cases, we can characterize the behavior of a Markov chain over a long period of time. In the case of an irreducible, aperiodic, positive-recurrent chain, the unique invariant probability measure provides such a characterization: for almost every sample path, the fraction of time the chain spends in a set in the long run is equal to the invariant measure of that set. The problem with null recurrent chains is that the unique invariant measure is not a probability measure and does not provide enough information about the behavior of a typical sample path over a finite period of time. To avoid this problem, in Section 3 we introduce a sequence of small random shocks to the model. The payoff function of agent \( i \) in the perturbed game is

\[
-(q_{it} - B\overline{q}_{-it} - \varepsilon_t)^2
\]

for every \( t \), where \( \{\varepsilon_t\} \) is an i.i.d. sequence that takes the values \( \varepsilon \) and \(-\varepsilon\) with probability \( 1/2 \) each. The shock \( \varepsilon_t \) simply shifts the best response function up or down by a small amount. For small values of \( \varepsilon \) the behavior of the resulting chain is similar to that of the unperturbed chain. Stability in the large and instability in the small continue to hold in the perturbed model, but the form of the not-too-unstable result changes. The chain is now positive recurrent and hence ergodic: expected departure times are not infinite, the long-run behavior of the chain is described by a non-degenerate invariant probability measure, and the distribution over states at date \( t \) converges to the invariant distribution as \( t \to \infty \). But as \( \varepsilon \) becomes small, all the probability mass of the invariant measure piles up around the origin, so the “not too unstable” property of the original model remains.

How do we explain these results? When the chain is close to the equilibrium, the experimenter rarely finds a better strategy than the one that he is
currently using. When the chain is far from the equilibrium, he finds a better strategy relatively frequently. Thus, adaptation is faster or slower for the experimenter depending on the degree of disequilibrium in the system. The imitators, on the other hand, are constantly adjusting their actions toward the average action (hence, on average, toward the experimenter’s action) at a constant proportionate rate. This means that the relative speeds of adjustment for experimenter and imitators vary, depending on the distance from equilibrium; and that explains how the model can be stable in the large but unstable in the small. It also accounts for the fact that the system can spend long periods close to the equilibrium, then cycle away in an increasing orbit for a long period, and then approach the equilibrium again.

Two properties of the model are crucial for these results: the presence of strategic substitutes and the non-vanishing size of the search window. (Recall that the game exhibits strategic substitutes as long as \( B < 0 \).) In Section 4.1 we study the (unperturbed) model under the assumption that the game exhibits strategic complements. More precisely, we assume that \( 0 < B < 1 \); that is, the strategic complements are “not too strong.” Under this assumption the symmetric equilibrium is shown to be stable in a strong sense: for any initial condition \((x_0, y_0)\) the chain converges to the origin with probability one.

To examine the role of non-vanishing window sizes, we extend the base-case model in Section 4.2 to allow for exogenously shrinking search windows. Formally, we assume that the experimenter chooses a strategy randomly from the interval \([x_{t-1} - d_t, x_{t-1} + d_t]\), where \( d_t > 0 \), for every \( t \). Here again we find that the equilibrium is stable in the same strong sense when the size of the search window converges to 0 as long as it does not converge too fast. More precisely, if

\[
\lim_{t \to \infty} d_t = 0 \quad \text{and} \quad \lim_{T \to \infty} \sum_{t=1}^{\infty} d_t = \infty
\]

then for any initial condition \((x_0, y_0)\) the chain converges to 0 with probability one. This result holds whenever \( B < 1 \).

Our interest in the case of non-vanishing window sizes comes from the observation that we live in a non-stationary environment. When the environment is constantly changing, one can never assume that one is close to the equilibrium or that experimentation can stop. Models of fictitious play, Bayesian learning, and adaptive learning assume both a stationary environment and that agents place less and less weight on recent experience.
as time passes. Such increasing inertia is essential to guarantee convergence. It makes sense in a stationary environment, where individual behavior can in principle converge to an optimum or equilibrium and beliefs can converge to the truth. In a world that is constantly changing, agents have no reason to assume that they have reached a permanent state of equilibrium. Consequently, they do have reason to continue to experiment and to give significant weight to recent experience. Although we study a stationary environment in order to make our results as clean and transparent as possible, the model is motivated by the assumption that agents always have something to learn.

What have we learned from these exercises?

• The first lesson is that the interaction of two simple types of behavioral adaptation can produce endogenous cycles. This may turn out to be a useful way of looking at certain kinds of macroeconomic fluctuations.

• A second lesson concerns the roles of strategic complements and substitutes. There has been a lot of interest in using models with strategic complements to explain the severity of macroeconomic fluctuations. In such models if individual activity levels are strategic complements, each agent’s best response is an increasing function of the activity levels of the others. If an agent increases his activity because of an exogenous shock, the others will increase their actions too. In this way, strategic complementarity magnifies the effect of the initial shock to the economy. One of the interesting features of our model is that either strategic substitutes or strong strategic complements are necessary for local instability, whereas with weak strategic complements the model is very stable.

• A third lesson is that the cycles in the models with strategic substitutes have a highly structured complexity that does not appear in other models in the literature. For example, in simulations we find that the amplitude of these cycles varies over time, sometimes being very damped and then growing again; but these variations are regular in the sense that the average amplitude of successive cycles is positively correlated.

• A fourth lesson concerns the role of experimentation. The randomness in the experimenter’s behavior is essential to generate the changing
relative rates of adaptation (between experimenters and imitators) that drive the dynamics.

This paper is related to several branches of the literature on learning and games. A discussion of the literature is postponed until the concluding Section 5, which also contains a discussion of some of the results, extensions of the model, and the related literature.

In the remainder of this section, we provide a formal definition of the model. If $x$ is the experimenter’s action and $y$ is the representative imitator’s action, the experimenter’s payoff is

$$- (x - By - \varepsilon_t)^2,$$

where $B < 1$ and $\{\varepsilon_t\}$ is an i.i.d. Bernoulli sequence

$$\varepsilon_t = \begin{cases} \varepsilon & \text{w. pr. } 1/2 \\ -\varepsilon & \text{w. pr. } 1/2. \end{cases}$$

The best response is given by $x = By + \varepsilon_t$, and the symmetric equilibrium is $x_t^* = y_t^* = \varepsilon_t/(1 + B)$. The better response set is

$$B(x, y, \varepsilon_t) \equiv \{x' \mid (x' - By - \varepsilon_t)^2 \leq (x - By - \varepsilon_t)^2\}.$$

The Markov chain $\Phi = \{\Phi_t\} = \{(X_t, Y_t)\}_{t=1}^\infty$ is defined by the initial condition $(x_0, y_0)$ and the laws of motion

$$X_t = \begin{cases} X_{t-1} + \omega_t & \text{if } X_{t-1} + \omega_t \in B(X_{t-1}, Y_t, \varepsilon_t) \\ X_{t-1} & \text{otherwise} \end{cases}$$

and

$$Y_t = \lambda X_{t-1} + (1 - \lambda)Y_{t-1},$$

where $\omega \equiv \{\omega_t\}$ is an independent sequence of random variables, $\omega_t$ is distributed uniformly on the interval $[-d_t, d_t]$, and $\{\omega_t\}$ and $\{\varepsilon_t\}$ are mutually independent. Of course, we assume that $0 < \lambda < 1$.

In Section 2, where the base-case model is studied, $\varepsilon = 0$ and $d_t = 1$, $\forall t$. In Section 3, $\varepsilon > 0$ and $d_t = 1$, $\forall t$. In Section 4.1, $0 < B < 1, \varepsilon = 0$, and $d_t = 1$, $\forall t$; and in Section 4.2, $\varepsilon = 0$, and $\{d_t\} \searrow 0$. 

7
2 The Base-Case Model

This section derives the three results for the base-case model; i.e., when \( \varepsilon = 0 \) and \( d_t = 1 \), \( \forall t \).

2.1 Stability in the Large

Our first result shows that the simple learning rules defined in the introduction are sensible: for any initial condition, the players’ strategies converge to a compact neighborhood of the unique symmetric equilibrium of the one-shot game. We call this property “stability in the large.” Note that this result holds whether the players’ strategies are strategic substitutes or strategic complements (as long as complementarities are not too large).

**Theorem 1** Suppose that \( B < 1 \), \( \varepsilon = 0 \), and \( d_t = 1 \), \( \forall t \). There exists a compact set \( K \) such that for any initial condition \((x,y)\), \( \Phi \) reaches \( K \) with probability one. Furthermore, for every sample path, once \( \Phi \) enters \( K \) it remains there forever.

A detailed proof is postponed until Section 4.1, since the fixed size of the search window is not necessary for this result; but the intuition behind it can be easily explained.

In Figure 1, the \( x \)-coordinate represents the action of the experimenter and the \( y \)-coordinate represents the action of the representative imitator. (The case depicted in the Figure is \( B < 0 \).) Suppose that the state of the chain is above the broken line corresponding to the equation \( y = x + \lambda^{-1} \). Here downward movement in the \( y \)-coordinate is always larger than the largest possible move in the \( x \)-coordinate, so the chain must eventually move into the corridor between the two broken lines. An exactly similar argument leads to the same conclusion if we start below the two broken lines.

Once the system enters the corridor, a simple calculation shows that it can never leave. Within the corridor, movement in the \( x \)-coordinate is always toward the best-response line. Hence, with probability 1, \((X_t,Y_t)\) must eventually pass into the compact section of the corridor illustrated in the figure.
2.2 Instability in the Small

Stability in the large tells us that from a macroscopic point of view the adaptive behavior prescribed for the experimenter and the imitators is successful in leading to approximately equilibrium behavior. From a microscopic point of view, however, the behavior of the system is very different. Adaptative behavior almost never converges to the equilibrium strategies. In fact, the equilibrium is “unstable in the small,” as the next result shows. For this result it is crucial that strategies be strategic substitutes.

**Theorem 2** Suppose that $\varepsilon = 0$ and $d_t = 1 \forall t$. For $B (< 0)$ sufficiently small, for some $\eta > 0$, and for any initial condition $(x, y) \in N(\eta) \equiv (-\eta, \eta) \times (-2\eta, 2\eta)$ satisfying $x \neq 0$, $\Phi$ leaves $N(\eta)$ with probability one.

The theorem says that any sufficiently small neighborhood of the equilibrium will be left with probability one in finite time. Taken together, stability in the large and instability in the small imply that the system will be subject to perpetual but bounded fluctuations about the equilibrium. Note that we are using the terms “large” and “small” only in a relative sense here. The size of the fluctuations will depend on the parameters of the model.

The proof of Theorem 2 is quite lengthy; we present its outline here, leaving details in lemmas which are proved in the appendix. We begin by focusing on the embedded Markov chain which only observes $(X_t, Y_t)$ in periods when the experimenter has been successful in finding a better strategy. Define a sequence of Markov stopping times $\{\tau_n\}$ recursively by:

$$\tau_1 = \min\{t > 0 | X_t \neq X_0\}$$

and

$$\tau_n = \min\{t > \tau_{n-1} | X_t \neq X_{t-1}\}, \forall n > 1.$$  

With a slight abuse of notation we denote the embedded chain by $\{(X_n, Y_n)\}$, where

$$(X_n, Y_n) \equiv (X_{\tau_n}, Y_{\tau_n}), n = 1, 2, ....$$

In other words, $(X_n, Y_n)$ is the position of the system just after the $n$-th success occurs.

As long as $x_n \neq 0$, it is clear that $\tau_{n+1}$ is finite with probability one. (As can be verified, this is not the case when $x_n = 0$.) Therefore, if $x_0 \neq 0$, an infinite sequence of successes occurs with probability one, and the chain
\{(X_n,Y_n)\} is well-defined if its state space is taken to be \((\mathbb{R}^1\setminus\{0\}) \times \mathbb{R}^1\) and the transition probabilities are derived appropriately from those of \{\{(X_t,Y_t)\}\}.

Now consider the statistic \(\xi_n \equiv |X_{n+1}|/|X_n|\), which is a measure of the rate at which the embedded chain is moving inward or outward. When the underlying chain \(\Phi\) is at some point \((x,y)\) very close to the origin, the better response set \(B(x,y,0)\) is entirely contained in the search window \([x-1,x+1]\); so the movement of the chain is unconstrained by the maximum step size of one unit that is imposed on the experimenter’s search. Suppose the position of the underlying chain just before the experimenter finds a better action is \((x,y)\). The better action \(x'\) is uniformly distributed on the better response set \(B(x,y,0)\). This means that \(|x'|/|x|\) is distributed according to a piecewise-linear distribution function which we denote by \(G(z;x,y)\), where the functional form \(G\) is independent of \(n\).\(^2\) In the special case where \(|x|\) is small and \(x = y\), it is easy to see that \(G(z;x,x)\) is independent of \((x,x)\) as well. In that case, when \(B < -1\), we can set

\[
G(z) \equiv G(z;x,x) = \begin{cases} 
0 & \text{if } z < 0 \\
\frac{2z}{2-2B} & \text{if } 0 \leq z < 1 \\
\frac{1}{2-2B} & \text{if } 1 \leq z < 1 - 2B \\
1 & \text{if } z \geq 1 - 2B.
\end{cases}
\]  

(1)

(Note the “two-step uniform” form).

Now consider the evolution of the underlying chain after \((X_n,Y_n) = (x,y)\), and denote by \(F(z;x,y)\) the cumulative distribution function of \(\xi_{n+1}\) given \((X_n,Y_n) = (x,y)\). If \((x,y)\) is near the origin, the probability of success on the first several experiments is small and the underlying chain will then move vertically toward the 45-degree line; so it seems intuitively plausible that \(F(z;x,y)\) is approximated by \(G(z)\). That intuition is confirmed by the following result.

**Lemma 3** For any \(\delta > 0\), \(\exists \eta > 0\) such that for all \((x,y) \in N(\eta)\), \(F(z;x,y) \leq (1 + \delta)G(z) \ \forall z \in \mathbb{R}.

**Proof:** See Appendix.

Next, if \(B\) is small enough, the expectation of a random variable having distribution function \(G\) above, and therefore of one having distribution function \(F\), is greater than 1. The precise bound we need is given in the next result.

\(^2\)The corresponding density has either two levels or one depending on whether the interior of the better response set includes a point on the \(y\)-axis or not.
Lemma 4 Let $\bar{F}$ be the distribution function defined by setting
\[
\bar{F}(z) = \begin{cases} 
0 & \text{if } z < 0 \\
\min \{1, (1 + \delta)G(z)\} & \text{otherwise.}
\end{cases}
\]
If $\delta > 0$ is sufficiently small and $B < 0$ is sufficiently small,
\[
\int_0^\infty \ln zd\bar{F} > 0.
\]

Proof: See Appendix.

Since $\bar{F}$ is an upper bound for $F$, this lemma tells us that $|X_n|$ is growing, on average, when $(X_n, Y_n) \in N(\eta)$. To make use of that fact, first consider an i.i.d. sequence $\{\zeta_j\}$ of random variables having distribution function $\bar{F}$ and define the process $\{Z_n\}$ by setting $Z_n = \prod_{j=1}^n \zeta_j$, for each $n = 1, 2, \ldots$. By the strong law of large numbers,
\[
\frac{1}{n}E[\ln Z_n] = \frac{1}{n}E[\sum_{j=1}^n \ln \zeta_j] 
\to E[\ln \zeta_1]
\]
with probability 1. Since $E[\ln \zeta_1] > 0$ by Lemma 4, it follows immediately that $Z_n \to \infty$ with probability 1.

We now want to exploit the fact that $|X_n| \equiv x_0 \prod_{j=1}^n \xi_j$ and use a similar argument to prove that $|X_n|$ diverges with probability 1, but there are two problems to be overcome: first, the random variables $\{\xi_j\}$ are not independent; and second, $\zeta_j$ only provides a bound for $\xi_j$ when $(X_{j-1}, Y_{j-1}) \in N(\eta)$. The first problem is overcome by the next result, which shows that the distribution of $|X_n|$ first-order stochastically dominates the distribution of $Z_n$. Hence, we can use the divergence of $\{Z_n\}$ to prove the divergence of $\{|X_n|\}$.

Lemma 5 Let $\{\zeta_j\}$ be an i.i.d. sequence with $E(\ln \zeta_1) > 0$ and let $\{\xi_j\}$ be another process, independent of $\{\zeta_j\}$, such that $\xi_j > 0$ a.s. and, conditional on $(\xi_1, \ldots, \xi_{j-1})$, $\xi_j$ dominates $\zeta_j$ in the sense of first-order stochastic dominance. Then $\prod_{j=1}^n \xi_j$ dominates $\prod_{j=1}^n \zeta_j$ in the sense of first-order stochastic dominance.

Proof: See Appendix.
The second problem is easily overcome by adapting the definition of the chain \( \{\lvert X_n \rvert \} \) outside the neighborhood \( N(\eta) \). Consider the process which is \( \lvert X_n \rvert \) until \( (X_n, Y_n) \) leaves \( N(\eta) \) for the first time and is incremented multiplicatively by the \( \{\zeta_j\} \) process thereafter. By the above lemmas, this process dominates \( \{Z_n\} = \{\prod_{j=1}^{n} \zeta_j\} \) in the sense of first-order stochastic dominance. Since \( \{Z_n\} \to \infty \) with probability 1, \( \{(X_n, Y_n)\} \) leaves \( N(\eta) \) with probability 1 in finite time. This completes the proof of Theorem 2.

2.3 Long-Run Behavior

Despite the instability of the equilibrium, it turns out that it is “not too unstable.” What we mean by “not too unstable” is described more precisely by the following result.

**Theorem 6** Suppose that \( \varepsilon = 0 \) and \( d_t = 1 \) \( \forall t \). For \( B (< 0) \) sufficiently small, for any \( \eta > 0 \), and any initial condition \( (x, y) \), there exists a number \( T < \infty \) such that \( \Pr[\Phi_t \in N(\eta)] > 1 - \eta \) if \( t > T \).

Thus \( \{\Phi_t\} \) converges weakly to the equilibrium.

Most of the rest of this section is devoted to establishing Theorem 6. To do so, we shall need to make heavy use of Markov chain theory. An excellent source, and the one from which we take all the definitions in the sequel, is Meyn and Tweedie (1993), hereafter [MT]. To begin, recall that the origin is an absorbing state of our chain, so the theorem is trivially true from that initial state. From other initial states on the vertical axis, either the chain stays on the axis, in which case the theorem is again true, or it leaves the axis and returns only with probability zero. It is sufficient, therefore, and more convenient to work from now on with the probabilistically equivalent chain having state space \( S \equiv (\mathbb{R}^1 \setminus \{0\}) \times \mathbb{R}^1 \) and \( (x_0, y_0) \in S \). We use the same notation \( \{(X_t, Y_t)\} \) for this chain, when there is no ambiguity. We take as our topology \( S \) the collection of open sets of \( \mathbb{R}^2 \) with points on the vertical axis deleted. Note that in this topology, every compact set must exclude some open neighborhood of the vertical axis. In particular, the set \( K \setminus \{(0) \times \mathbb{R}^1\} \), though bounded, is not compact. Denoting by \( \mathcal{A} \) the \( \sigma \)-algebra generated by \( S \), for all \( A \in \mathcal{A} \) and for all \( s \in S \), define

\[
P(s, A) = \Pr((X_{t+1}, Y_{t+1}) \in A \mid (X_t, Y_t) = s) \quad \text{and} \quad P^n(s, A) = \Pr((X_{t+n}, Y_{t+n}) \in A \mid (X_t, Y_t) = s) \text{ for } n > 1.
\]

The chief tool in establishing Theorem 6 is the following.
**Proposition 7** \( \{(X_t, Y_t)\} \) is a \( \psi \)-irreducible, aperiodic, Harris, null-recurrent, Feller chain.

**Proof:** See Appendix.

The definitions of terms used in the proposition can also be found in the appendix. The intuition for Theorem 6 follows from Proposition 7 using the simpler language of countable-state theory: in the long run an irreducible, aperiodic, null-recurrent chain wanders among its recurrent states in such a way that the fraction of time it spends in any one state converges to zero as time goes to infinity. The counterpart of this in the uncountable-state-space setting is that the fraction of time our chain spends in any compact set converges to zero as time goes to infinity, even though it visits individual compact sets infinitely often. The way this happens is that the chain returns to every neighborhood of the vertical axis infinitely often, and the closer it is to the axis, the longer it takes to get away on average.

To prove Theorem 6 from Proposition 7, observe that the chain can admit no invariant probability measure (this is the definition of a null chain) but must have a unique invariant measure \( \pi \) ([MT, Theorem 10.4.4]), with \( \pi(K\backslash\{(0) \times \mathbb{R}^1\}) = \infty \), which is finite on compact sets ([MT, Theorem 12.3.2] using Theorem 2 and Proposition 7 above to satisfy the hypotheses\(^3\)). Now take \( Q \) to be the open vertical strip of width \( q \) centered on the \( y \)-axis; then \( K\backslash Q \) is compact and so \( \pi(K\backslash Q) < \infty \). If \( q \) is sufficiently small, from any point in \( (K \cap Q)\backslash N(\eta) \), the chain must leave \( Q \) in one step with some probability \( \xi > 0 \). Hence \( \pi(K\backslash Q) \geq \xi \pi((K \cap Q)\backslash N(\eta)) \).

Thus \( \pi(K\backslash N(\eta)) = \pi((K \cap Q)\backslash N(\eta)) + \pi((K \backslash Q)\backslash N(\eta)) \leq (1 + \xi^{-1})\pi(K\backslash Q) < \infty \).

But, from [MT, Theorem 18.1.3], for every \( \delta > 0 \),

\[
\lim_{n \to \infty} \frac{P^n((x_0, y_0), (K\backslash N(\eta)))}{\pi(K\backslash N(\eta)) + \delta} = 0.
\]

The result then follows from the finiteness of \( \pi(K\backslash N(\eta)) \).

From the proof of Theorem 6, it also follows that the expected time the chain takes to leave any sufficiently small square centered at the origin is infinite. To this end, define \( \tau(\eta) \equiv \min\{t : \Phi_t \notin N(\eta)\} \).

---

\(^3\)Theorem 2 states that \( \Phi \) leaves \( N(\eta) \) with probability 1. To do this it must move to a point for which \( x \notin (-\eta, \eta) \) and hence enter a compact set, which is a hypothesis of [MT, Theorem 12.3.2].
**Corollary A**  Under the hypotheses of Theorem 6, if $\eta$ is sufficiently small, $E(\tau(\eta) \mid \Phi_0 \in N(\eta)) = \infty$.

**Proof:** See Appendix.

## 3  The Perturbed Model

In this section we look at a slightly perturbed version of the model in which $\varepsilon > 0$ and $d_t = 1$ for all $t$. The introduction of $\varepsilon$ ensures that the experimenter always has a chance that is bounded away from zero of finding a better response, even near the equilibrium. The resulting chain is positive recurrent.

### 3.1 Stability in the Large

As one might expect, stability in the large is not affected by small perturbations.

**Theorem 8**  Suppose that $B < 1$, $\varepsilon > 0$, and $d_t = 1$ for all $t$. There exists a compact set $K_\varepsilon$ such that for any initial condition $(x, y)$, $\Phi$ reaches $K_\varepsilon$ with probability one. Furthermore, for every sample path, once $\Phi$ enters $K_\varepsilon$ it remains there forever.

The arguments are the same as for Theorem 1; again the proof is postponed until Section 4.1.

### 3.2 Instability in the Small

A version of instability in the small also continues to hold for the perturbed model, but its interpretation requires some care. Since the symmetric equilibrium of the one-shot game moves, we cannot hope for asymptotic convergence. However, the result is stronger than this negative statement. It says that for any fixed, sufficiently small neighborhood of the origin and for arbitrarily small perturbations $\varepsilon$, $\Phi$ will leave the neighborhood with probability one. In other words, we can choose the perturbations as small as we like relative to the fixed neighborhood. So it is not the perturbations alone that are leading to instability; the equilibrium is inherently unstable (in the small).
Theorem 9 Suppose that $B < 0$, $\varepsilon > 0$, and $d_t = 1$ for all $t$. For some $\eta > 0$ and any initial state $(x_0, y_0) \in N(\eta)$, $\Phi$ leaves $N(\eta)$ with probability one. Furthermore, $\eta$ can be chosen independent of $\varepsilon$.

Proof: Choose $\eta$ and $\varepsilon$ so that, for any $(x, y) \in N(\eta)$, the better response set is contained in the search window; that is, $B(x, y, \varepsilon_t) \subset [x - 1, x + 1]$.

Note that $\varepsilon$ can be chosen arbitrarily small without affecting the value of $\eta$.

Let $(x, y)$ be a fixed but arbitrary point in $N(\eta)$ and put $y_1 = \lambda x + (1 - \lambda)y$ and $y_{n+1} = \lambda x + (1 - \lambda)y_n$ for $n = 1, 2, ...$

The proof proceeds in a number of steps.

Step 1. For any $\gamma > 0$ there exists $\bar{n}$ such that for any $(x, y) \in N(\eta)$, $|y_n - x| < \gamma$ for any $n > \bar{n}$. (Proof obvious.)

Step 2. For some constant $\delta > 0$, some number $n$ and any initial condition $(x, y) \in N(\eta)$,

$$P^n((x, y), \{(x', y') : |x'| \geq |x| + \delta\}) \geq \delta.$$  

Proof: Without loss of generality, we can assume that $x \leq 0$. (The other case is exactly symmetrical.) Choose $\gamma$ so that $B\gamma + \varepsilon > 0$, and let $\bar{n}$ satisfy the conditions of Step 1. Then suppose that for the first $\bar{n}$ periods, the experimenter is not successful in finding a better strategy. Since the probability of failure in each period is at least $1/2$, the probability of this event is at least $1/2^{\bar{n}}$. In the next period, $\varepsilon_{\bar{n}+1} = \varepsilon$ with probability $1/2$. The largest value in the better response set is $2(By_{\bar{n}+1} + \varepsilon) - x$, so the elements in the better response set with absolute value greater than $|x|$ comprise the interval $[-x, 2(By_{\bar{n}+1} + \varepsilon) - x]$. The length of this interval is $2(By_{\bar{n}+1} + \varepsilon)$ so the midpoint is $(By_{\bar{n}+1} + \varepsilon) - x$ and the probability that the experimenter’s choice falls in the right half of the interval is $\frac{1}{2}(By_{\bar{n}+1} + \varepsilon)$. Note that the probability $\frac{1}{2}(By_{\bar{n}+1} + \varepsilon)$ is bounded away from zero:

$$By_{\bar{n}+1} + \varepsilon \geq B(x + \gamma) + \varepsilon \geq B\gamma + \varepsilon,$$

since $Bx \geq 0$. So, if $x'$ is the next strategy chosen, then with probability $\frac{1}{2}(B\gamma + \varepsilon)$ it has absolute value of at least $|x| + \frac{1}{2}(B\gamma + \varepsilon)$. Putting $\delta = (1/2^{\bar{n}+2})(B\gamma + \varepsilon) \leq \frac{1}{2}(B\gamma + \varepsilon)$, we have the desired result. ■
Step 3. For some $\delta > 0$ and any initial condition $(x, y) \in N(\eta)$, there exists a fixed integer $n$ such that $P^n((x, y), \{(x', y'): (x', y') \notin N(\eta)\}) \geq \delta$.

**Proof:** Apply Step 2 repeatedly, since $\pi$ is independent of $(x, y)$. $\blacksquare$

Step 4. For any initial condition $(x, y) \in N(\eta)$, the Markov chain $(X_t, Y_t)$ leaves $N(\eta)$ in finite time with probability one.

**Proof:** From Step 3,

$$\Pr[(X_t, Y_t) \in N(\eta), \forall t] \leq \Pr[|X_t| \leq k, t = 1, ..., nm] \leq (1 - \delta)^m$$

for any finite $m$. Letting $m \to \infty$ gives the desired result. $\blacksquare$

This completes the proof of Theorem 9. $\blacksquare$

Note that this result holds for arbitrarily small values of $\varepsilon$, but for fixed $\eta$. Thus the degree of instability is “large” relative to the vanishingly small perturbations.

Although Theorem 9 tells us that for small $\eta$ the chain leaves $N(\eta)$ with probability one, it leaves open the possibility that the expected departure time is infinite, as in Corollary A. However, we have

**Corollary B** Under the hypotheses of Theorem 9, for sufficiently small $\eta$ and $\varepsilon$ there is a uniform bound $M(\eta, \varepsilon)$ such that $E[\tau(\eta)| \Phi_0 \in N(\eta)] < M(\eta, \varepsilon)$.

**Proof:** From the proof of Theorem 9,

$$E[\tau(\eta)| \Phi_0 \in N(\eta)] \leq \sum_{m=1}^{\infty} nm(1 - \delta)^m < \infty. \blacksquare$$

Note that in the proof of Theorem 9, $\delta$ depends on $\varepsilon$ and goes to zero as $\varepsilon$ does. So the bound on the expected first departure time goes to infinity as the size of the perturbations goes to zero.

### 3.3 Long-run Behavior

Long-run properties of the perturbed chain can be studied using the same tools as for the unperturbed chain, but the arguments are more straightforward since the perturbation eliminates the odd behavior of the base-case model in the neighborhood of the origin. In particular, we can adopt here the state space $\mathbb{R}^2$ and the usual topology.
Lemma 10  Suppose that $B < 0$, $\varepsilon > 0$, and $d_t = 1$ for all $t$. Then $\Phi$ is an aperiodic, Feller, $\psi$-irreducible, positive, Harris recurrent chain.

Proof: The proof of aperiodicity, the Feller property and $\psi$-irreducibility follow from the same arguments as were used on the unperturbed model. Similarly, it can be shown that the support of $\psi$ has a non-empty interior, so Theorem 6.2.8 of [MT] implies that all compact sets are “petite.” Theorem 9.1.7 of [MT] tells us that $\Phi$ is Harris recurrent if there exists a petite set $C$ such that the probability is one that the chain eventually hits $C$ from any initial condition $(x, y)$. The set $K$ from Theorem 8 satisfies this condition, so $\Phi$ is Harris recurrent. Then Theorem 12.3.2 of [MT] implies that there exists an invariant measure $\pi$ that is finite on compact sets. Since the support of $\pi$ is a compact set, this means that $\pi$ is (proportional to) a probability measure. In other words, $\Phi$ is positive recurrent.

Theorem 11  Suppose that $B < 0$, $\varepsilon > 0$, and $d_t = 1$ for all $t$. Then for every initial condition $(x_0, y_0)$,

$$\|P^n((x_0, y_0), \cdot) - \pi\| \to 0, \text{ as } n \to \infty,$$

where $\pi$ is the unique invariant probability measure of the chain and $\| \cdot \|$ denotes the total variation norm ([MT], p.311).

Proof: This result is just a restatement of Theorem 13.3.3 of [MT] once we note that $\Phi$ is an aperiodic and positive Harris chain.

What this result tells us is that as long as there is nonvanishing movement in the position of the equilibrium, the long-run behavior of the chain is well approximated by its invariant probability distribution. The support of the invariant distribution must include points in $K \setminus N(\eta)$, so the fluctuations about the equilibrium are persistent and the average volatility is bounded away from zero.

However, as $\varepsilon \searrow 0$ so does average volatility.

Theorem 12  Suppose that $B < 0$ and $d_t = 1$ for all $t$. Let $\{\varepsilon_j\} \searrow 0$, and let $\pi_j$ be the unique invariant measure for the chain perturbed by $\varepsilon = \varepsilon_j$. Then $\{\pi_j\}$ converges weakly to the Dirac probability measure at the origin.

Proof: The proof follows immediately from Theorem 12.13 in Stokey and Lucas (1989) once we notice that for the reducible version of the unperturbed
chain, i.e. having state space $\mathbb{R}^2$, the Dirac probability measure at the origin is the unique invariant probability measure. That it is invariant is obvious. If there were another invariant probability measure, its restriction to $\mathbb{R}^1 \backslash \{0\} \times \mathbb{R}^1$ would be a finite invariant measure for the chain studied in Section 2, contradicting Proposition 7. ■

It may be helpful for intuition to compare the stochastic stability of our perturbed process with that of the well-known model of Kandori, Mailath, and Rob (1993), hereafter KMR. Consider the simplest KMR model, involving a $2 \times 2$ coordination game with strategies $X$ and $Y$. Let the state space of the induced Markov chain be $\{0, 1, \ldots, N\}$, where a state denotes the number of people (out of $N$) playing strategy $X$. Let $(Y, Y)$ be the risk-dominant equilibrium, and let $\varepsilon (> 0)$ be the probability of a mutation. Then the Markov chain has a unique invariant probability distribution, the state distribution converges to this invariant distribution as time approaches infinity, and the invariant distribution converges weakly to a mass of unity on state 0 as $\varepsilon$ approaches zero. Now let $p(\varepsilon, t)$ be the probability that, starting in state zero, the process leaves state zero within the first $t$ periods, and let $T(\varepsilon)$ be the expected length of time until the process first leaves state zero. Then $T(\varepsilon)$ approaches infinity as $\varepsilon$ approaches zero and, for any integer $t$, $p(\varepsilon, t)$ approaches zero as $\varepsilon$ approaches zero. One can interpret all this as analogous to unstable, but not too unstable.

Our perturbed Markov chain also has a unique invariant probability distribution, and the state distribution converges to this invariant distribution as time approaches infinity. Also, as $\varepsilon$ approaches zero, the invariant distribution converges weakly to a mass of unity on the state $(0, 0)$. Furthermore, if we define $T(\varepsilon)$ as the expected time to escape from an appropriate neighborhood of the origin $(0, 0)$ starting from $(0, 0)$ and $p(\varepsilon, t)$ as the probability that the process first escapes from the neighborhood by time $t$, then $T(\varepsilon)$ approaches infinity as $\varepsilon$ approaches zero and, for any fixed $t$, $p(\varepsilon, t) = 0$ for all $\varepsilon$ sufficiently small. The reason is that for $\varepsilon$ sufficiently small, the maximum size of the steps taken from $(0, 0)$ become so small that it is impossible to escape from the neighborhood in $t$ steps. However, suppose that $(x, y)$ is fixed but arbitrary close to $(0, 0)$ and define $p(x, y, \varepsilon, t)$ analogously to $p(\varepsilon, t)$ but with $(x, y)$ as the starting point rather than $(0, 0)$. Then for any fixed initial condition $(x, y)$ there is a time $t$ such that the escape probability $p(x, y, \varepsilon, t)$ is bounded away from zero as $\varepsilon$ converges to zero.
In this sense then, our perturbed model is more unstable than that of KMR. But the comparison is awkward in a number of ways. First, the KMR model has a finite state space; so escaping from the state 0 in that model is not the same as escaping from a neighborhood of the origin in ours, and there is no counterpart there to starting states that are arbitrarily close to, but not the same as, the origin in our model and hence no counterpart to the condition that $t$ must grow as $(x, y)$ approaches the origin in order for $p(x, y, \varepsilon, t)$ to remain bounded away from zero. Second, a single mutation in KMR is enough to move the chain from one state to another. What is being reduced in KMR is the probability of a mutation, not its size, as $\varepsilon$ converges to zero. In our model, it is the size of a perturbation, not its probability, that is being reduced as $\varepsilon$ converges to zero. Third, in KMR, it appears that mutations play a much bigger role in driving the process away from, and hence selecting among, equilibria than perturbations do in the present model. Here the perturbations merely serve to counteract the singularity at $(0, 0)$ so that the inherent instability of the process becomes more visible.

4 Asymptotic Stability

So far our analysis has focussed on the failure of adaptive behavior to converge to equilibrium behavior. In this section, we explore two cases in which such convergence does occur. To get sharp results, we assume that $\varepsilon = 0$, so that the unique symmetric equilibrium of the game in each period is the origin.

In the first case, we allow for both strategic complements and strategic substitutes by assuming only that $B < 1$. However, we let the size of the search window shrink over time. If the window size eventually converges to zero but not too quickly, we show that $\Phi$ converges to zero with probability one. This means that the assumption of continual search ($\lim \sup \{d_t\} > 0$) is essential for persistent fluctuations.

In the second case, we assume that the strategies of the experimenter and the imitators are strategic complements, $0 < B < 1$, and show that again $\Phi$ converges to zero with probability one. This shows that the assumption of strategic substitutes, $B < 0$, is essential for persistent fluctuations.
4.1 Vanishing Search Windows

In the previous sections, the hypothesis that the search window is non-vanishing plays a crucial role in sustaining the model’s persistent dynamic fluctuations. The proof of this lies in the fact that, if the search window is allowed to shrink over time at just the right rate, it is possible to ensure that the equilibrium of the underlying game is asymptotically stable with probability one. This can be established by the same techniques used to demonstrate stability in the large, and the proofs of all these results are gathered here to make clear the close relationships among them.

We begin by establishing some lemmas for the most general model, in which we allow for random perturbations and varying size of the search window. For every $t$ the random variable $\omega_t$ is distributed uniformly on the interval $[-d_t, d_t]$, and the perturbation $\varepsilon_t$ is defined as in Section 3 with the parameter $\varepsilon \geq 0$. The Markov chain $\{(X_t, Y_t)\}_{t=1}^{\infty}$ is defined by the initial condition $(x_0, y_0)$, and the law of motion is as before.

For any fixed but arbitrary $d > 0$ let $C_d$ denote the corridor of points around the diagonal defined by

$$C_d = \{(x, y) \mid |x - y| \leq d/\lambda\}.$$

**Lemma 13** Suppose that $d_t \leq d$ for all $t > T$. For any initial condition $(x, y)$, $\Phi$ reaches $C_d$ with probability one. Furthermore, for any sample path, if $\Phi_t \in C_d$ for $t > T$ then $\Phi_s \in C_d$ for all $s > t$.

**Proof:** See appendix.

Let $x_{ed}$ denote the largest value of $x$ that belongs to the set of points in $C_d$ to the left hand of the best response line. That is,

$$x_{ed} = \sup\{x \mid (x, y) \in C_d, x \leq By + \varepsilon\}.$$

Symmetrically,

$$-x_{ed} = \inf\{x \mid (x, y) \in C_d, x \geq By - \varepsilon\}.$$

The value of $x_{ed}$ corresponds to the intersection of the best response line $x = By + \varepsilon$ with one of the boundaries of $C_d$. If $0 < B < 1$ then $x_{ed}$ corresponds to the intersection with the upper boundary $y = x + d/\lambda$, and

$$x_{ed} = (Bd/\lambda + \varepsilon)/(1 - B).$$
If $B < 0$ then $x_{ed}$ corresponds to the intersection with the lower boundary $y = x - d/\lambda$, and

$$x_{ed} = (\varepsilon - Bd/\lambda)/(B - 1).$$

Let $K_{ed} = \{(x, y) \in C_d | -x_{ed} - d \leq x \leq x_{ed} + d \}$.

**Lemma 14** Suppose that for all $t > T$, $d_t \leq d$, and suppose $\sum_t d_t = \infty$. For any initial condition $(x_0, y_0)$, the chain reaches $K_{ed}$ with probability one. Furthermore, for any sample path, if $(X_t, Y_t) \in K_{ed}$ for $t > T$, then $(X_s, Y_s) \in K_{ed}$ for all $s > t$.

**Proof:** See appendix.

**Proof of Theorem 1:** Lemma 14 with $K = K_{01}$ and $d = 1$.

**Proof of Theorem 8:** Lemma 14 with $K_{\varepsilon} = K_{\varepsilon 1}$ and $d = 1$.

We can also prove from Lemma 14 the asymptotic stability of equilibrium when the size of the search window vanishes at the right rate.

**Theorem 15** Suppose that $B < 1$, $\varepsilon = 0$, and $\omega_t$ is uniformly distributed on the interval $[-d_t, d_t]$ for every $t$. Suppose further that $\{d_t\} \searrow 0$ and $\sum_t d_t = \infty$. Then for any initial condition $(x, y)$, $\Phi$ converges to 0 with probability one.

**Proof:** For any neighborhood of the origin, there exists a $d > 0$ such that $K_{0d}$ is contained in the neighborhood. Since $\{d_t\} \searrow 0$, for any $d > 0$ there exists a $T$ such that $d_t < d$ for all $t > T$. Since we have assumed that $\sum_t d_t = \infty$, the conditions of Lemma 14 are satisfied and imply that $\Phi$ reaches $K_{0d}$ with probability one and remains there for all $t$ sufficiently large. Hence, $\Phi$ reaches every neighborhood of the origin with probability one and remains there for all $t$ sufficiently large, so $\Phi$ converges to 0 with probability one. ■

Shrinking the size of the search window has two effects on the dynamics of the model. On the one hand, it lowers the speed of adaptation by restricting the experimenter’s ability to move toward the best response. This effect makes convergence more difficult when the chain is far from the equilibrium. On the other hand, when the chain is close to the equilibrium, shrinking the window size increases the probability that the experimenter will find a better response. Close to the equilibrium the experimenter’s current action is close to a best response and a small search window ensures a higher probability
of landing in the better response set. These two effects have to be carefully balanced in order to make the chain asymptotically stable. We have to let the window size converge to zero in order to take advantage of the second effect. But it must not converge to zero too quickly or the first effect will prevent getting close to the equilibrium at all.

4.2 Strategic Complements

In this section we return to the assumption of a fixed search window of size two: \( d_t = 1 \) for every \( t \). We also assume \( 0 < B < 1 \) (strategic complements).

For any number \( a > 0 \) let \( K_a \) denote the parallelogram with vertices \((a, 0), (a, a/B), -(a, 0)\) and \(-(a, a/B)\). The best response line is the diagonal of \( K_a \) running through \(-(a, a/B)\) and \((a, a/B)\). By construction, the diagonal is horizontally equidistant from the adjacent sides. (See Figure 2.) Since the better response set is symmetric about the best response line, \((x, y) \in K_a\) implies that \( B(x, y) \subset K_a\).

**Lemma 16** For any sample path \( \{(x_t, y_t)\}_{t=1}^{\infty} \) and any \( a > 0 \), if \((x_t, y_t) \in K_a\) then \((x_s, y_s) \in K_a\) for all \( s > t \).

**Proof**: Suppose that \((x_t, y_t) \in K_a\) for some \( a > 0 \). Since \( y_{t+1} = \lambda x_t + (1 - \lambda)y_t \), \((x_t, y_{t+1})\) lies vertically between \((x_t, y_t)\) and the 45-degree line, which means that \((x_t, y_{t+1}) \in K_a\). Then \((x_{t+1}, y_{t+1}) \in B(x_t, y_{t+1}) \subset K_a\). Repeating this argument we find that \((x_s, y_s) \in K_a\) for all \( s > t \) as required. ■

**Lemma 17** For any \( a > 0 \) and \( b = a - \gamma \), where \( \gamma = \min\{1/2, a/2, 1 - B\} \), there exists a number \( \delta > 0 \) and an integer \( n \) such that for any initial condition \((x, y) \in K_a\), \( P^n((x, y), K_b) \geq \delta \).

**Proof**: If \((x, y) \in K_b\), there is nothing to prove; so suppose that \((x, y) \in K_a \setminus K_b\). As in the proof of Theorem 9, after \( \bar{n} \) failures, which occur with probability \( 1/2^n \), the state of the system will have moved to within a vertical distance \( \gamma \) of the 45-degree line. Thus, there is no essential loss of generality in assuming that \( |x - y| < \gamma \). There are two cases to consider.

**Case 1.** Suppose that \( |x| \leq b \). A simple calculation shows the maximum value of \( |x - y| \) over the set \( K_b\) is \( \min\{1, (1 - B)/B\}b \). The assumption that \( \gamma \leq \min\{a/2, 1 - B\} \) is equivalent to \( \gamma \leq \min\{1, (1 - B)/B\}b \), so it follows immediately that \((x, y) \in K_b\).
Case 2. Suppose that \( b < |x| \leq a \). Since \( |x - y| < \gamma \) the probability that the next state \((x', y')\) lies in \( K_b \) is at least

\[
\left(\frac{1}{2}\right) \min \left\{ \frac{1}{2}, b - By' \right\} \geq \left(\frac{1}{2}\right) \min \left\{ \frac{1}{2}, b - B(a + \gamma) \right\}
\]

where \( y' = \lambda x + (1 - \lambda)y \leq \lambda a + (1 - \lambda)(a + \gamma) \leq (a + \gamma) \). This number is bounded away from zero, since

\[
b - B(a + \gamma) = a - \gamma - B(a + \gamma) = (1 - B)a - (1 + B)\gamma \\
\geq (1 - B)a - \frac{(1 - B)a}{2} = \frac{(1 - B)a}{2} > 0.
\]

Now we are ready to complete the proof of the convergence theorem.

**Theorem 18** Suppose that \( 0 < B < 1, \varepsilon = 0, \) and \( d_t = 1 \) for all \( t \). Then for any initial condition \((x, y)\), \( \Phi \) converges to 0 with probability one.

**Proof:** Let \((x, y)\) be the initial condition and suppose first that \((x, y)\) \(\in A\), where

\[ A = \{(x, y) | |x| \geq B|y|\}. \]

Now put \( a_1 = |x| \) and define

\[ a_{n+1} = a_n - \min \left\{ \frac{1}{2}, \frac{a_n}{2}, 1 - B \right\} \]

inductively for every \( n = 1, 2, ..., \). From Lemma 17 and [MT Theorem 9.1.3] it follows that

\[ \{ \Phi \in K_{a_n} \text{ i.o.} \} = \{ \Phi \in K_{a_{n+1}} \text{ i.o.} \} \]

with probability one. Since Lemma 16 implies that \( \Phi \in K_{a_1} \) i.o. for every sample path, it follows by induction that \( \Phi \in K_{a_n} \) i.o. with probability one for every \( n \). Then Lemma 16 implies that with probability one \( \Phi_t \in K_{a_n} \) for all \( t \) sufficiently large. Since \( a_n \to 0 \) as \( n \to \infty \), every neighborhood of the origin contains the set \( K_{a_n} \) for some \( n \) sufficiently large, so \( \Phi \) converges to 0 with probability one.

To complete the proof of the theorem, note first that for any initial condition \((x, y)\) with \( x \neq 0 \), \( \Phi \) reaches \( A \) with probability one. This follows from the fact that the set \( A \) is uniformly accessible from the set \( K_0 \setminus \{(0) \times \mathbb{R}^1\} \) defined in the last section. Finally, if \( x = 0 \), either \( \Phi \) reaches \( A \) or no successful experiments occur, in which case \( \Phi \) converges to 0. \( \blacksquare \)
5 Discussion

5.1 Model Variations

For the purpose of illustrating some basic properties and simplifying the formal analysis, we have focused on a simple version of the model. In this section we discuss some modeling issues and possible extensions.

Virtual experimentation

In modeling the behavior of the experimenter, we have assumed that he conducts a virtual experiment (to determine whether a new strategy belongs to the better response set) before he actually adopts the strategy. Here is a possible interpretation of this rule. Think of a period as an interval of time. At the beginning of the period, the experimenter randomly selects a strategy from a neighborhood of the previous strategy. He experiments with the new strategy for a very short part of the interval, so short that the imitators do not observe the change. If the new strategy has a higher payoff against the current imitators’ strategy than the experimenter’s previous strategy, then he uses it for the rest of the period. Otherwise, he reverts to his previous strategy. In either case, the imitators observe the strategy he is using by the end of the period and use it to update their behavior in the next period.

Without such virtual experimentation, the imitators would observe and be influenced by the random selection of an alternative strategy, whether or not it leads to an improvement in the experimenter’s performance. The “noise” introduced by random search would then prevent the adaptive process from settling into an equilibrium: even if an equilibrium were reached, the experimenter would continue to test new strategies and this testing would jolt the imitators’ behavior away from equilibrium. By filtering out this “noise,” we highlight the inherent instability of the model.

Multiple experimenters

For simplicity, we have focused in this paper on the case of a single experimenter. To see whether the assumption of a single experimenter is restrictive or not, we simulated the base-case model with different numbers of experimenters.

Dealing with multiple experimenters raises some modeling issues. At the beginning of the period, all the imitators simultaneously update their strategies based on information from the previous period. If there is a single experimenter, the better response set is a function of the imitators’ average
strategy. When there are multiple experimenters, the better response set depends as well on what the other experimenters are doing. It will make a difference whether they experiment simultaneously or sequentially and, if sequentially, the order will matter.

In our simulations we assumed simultaneous experimentation. After the experiment, some of the experimenters sometimes revert to their previous strategies. Hence the payoffs received by some experimenters during experimentation will be different from the payoffs they receive after adoption, leading to regret. So some experimenters may have adopted new strategies which give lower payoffs than their previous strategies, and some may have abandoned strategies that would have given higher payoffs. These type I and type II errors do not arise when there is a single experimenter, because the imitators’ actions are the same in virtual and actual play.

The simulations we performed gave the same qualitative results as we got with a single experimenter, but the size of the fluctuations was not monotonic in the number of experimenters. With small numbers of experimenters, the fluctuations about the symmetric equilibrium were greater than with one experimenter. With larger numbers, the fluctuations got somewhat smaller. The reason appears to be that when the number of experimenters is small, they interfere with each other, adding noise to the adaptive process. This noise is eliminated as the number of experimenters becomes large.

The interaction of several experimenters makes analysis of the model more difficult, but it raises some interesting possibilities. First, the simulations suggest that the social system may track the equilibrium of the underlying game less efficiently with multiple experimenters than with one experimenter. Secondly, interaction among experimenters may exacerbate the uncertainty that arises when the environment is also subject to shocks: Suppose that the system has been in the neighborhood of the equilibrium of the underlying game for some time and the game is then perturbed in a way that shifts the equilibrium. This will immediately produce an increase in the number of successful experiments, but it is not clear that the experimenters will immediately move in the direction of the new equilibrium because of the confounding effect mentioned above. There may thus be a period of increased uncertainty, as measured by the dispersion of the experimenters’ actions. We think that analytical study of a model like ours with multiple experimenters could well produce interesting new effects.

*Finite-state approximations*
A finite-state Markov chain cannot be null recurrent: If the chain returns to a given state repeatedly it must do so in finite expected time. This indicates that there is something essential about the infinite state space used in this paper. This in turn raises the issue of whether the simulations we have performed on a finite-state machine (computer) are good approximations to the behavior of the theoretical model.

An interesting question is whether one could approximate the behavior of the infinite-state model using a finite grid to approximate each player’s strategy set. There are some delicate questions about exactly how the grid should be chosen. For example, if the grid includes the zero point, then any random process will find the origin eventually and if the origin is a rest point of the process (a necessary condition for convergence) that will be the end of the story. On the other hand, if zero is not included, which seems natural if the players are not very smart, then the one-shot equilibrium cannot be reached, and the long-run properties of the chain could be very different.

**B regions**

For the case of strategic substitutes ($B < 0$), we have a fairly complete characterization of the dynamics of the perturbed model (Section 3). For the unperturbed model (Section 2) we have proofs of “instability in the small” and “not too unstable” only for sufficiently small values of $B$. This leaves open the possibility that these results do not hold for larger values of $B < 0$, or it may simply be that a different proof is called for.

For moderate-sized strategic complements ($0 < B < 1$), we have only analyzed the unperturbed model. In that case the convergence result in Section 4.1 tells us all we need to know. Something similar ought to hold for the perturbed model, but the statement would be messy since one could only hope for convergence to a neighborhood of the origin.

For strong strategic complements ($B > 1$), the model is clearly unstable, although we have not bothered to state and prove this as a formal result. For intuition, note that for large enough values of $(x, y)$, the chain moves to and then stays in the cone between the 45-degree line and the best-response line. In this cone, the motion is upward in both coordinates. (Similarly when $(x, y) << 0$).

For the borderline cases $B = 0$ and $B = 1$ we do not have theorems for the unperturbed model, but it appears that the former is “convergent” and the latter is “weakly stable” (after a shock the chain returns to an equilibrium but not the same one).

26
The meaning of “not too unstable”
The meaning of “not too unstable” in Sections 2.3 and 3.3 is somewhat different. In both cases, the Markov chains are shown to have an ergodic property. In Section 2.3, this takes the form of weak convergence of the state distributions to the origin as $t$ goes to infinity. In Section 3.3, it takes the form of weak convergence of the state distributions to the non-degenerate invariant probability distribution as $t$ goes to infinity.

Although these results are different, they reinforce each other: as $\varepsilon \to 0$, the invariant measure of the perturbed model converges weakly to the Dirac measure at the origin. In this sense, slight perturbations of the model lead to slight changes in the invariant distribution.

5.2 Parameters and Asymptotic Volatility

The asymptotic volatility of the model is captured by the invariant measure $\pi$. Although we have some information about $\pi$ (for example, “stability in the large” tells us that the support of $\pi$ is contained in a compact set $K$ and “instability in the small” tells us that the support of $\pi$ is a superset of the origin), these results do not tell us very much about how large the volatility is or how it changes with the parameters of the model. It would be nice to be able to measure more precisely how much volatility the model exhibits and how it responds to changes in the parameters. Unfortunately, the complexity of the stochastic process makes it hard to capture the amount of volatility and its sensitivity to the parameters. A few qualitative conclusions can be drawn, however.

For one parameter, the size of the (fixed) search window $d$, it is easy to describe the sensitivity of $\pi$ to changes in $d$. From the general specification of the model, it is clear that a change in $d$ (together with an appropriate change in the initial state) is equivalent to a change in the scale of the Markov chain. An increase in $d$ simply increases each element of the sequence $\{(X_t(\omega), Y_t(\omega))\}$ proportionately, for every realization of the state of nature $\omega$. Thus, every aspect of the Markov chain, including the invariant measure $\pi$, is reproduced at a larger scale as the size of the search window increases, and in this precise sense we can say that an increase in $d$ increases volatility proportionately for any reasonable measure of volatility.

For the parameter $B$ we know that asymptotic stability holds for every $0 \leq B \leq 1$. Thus it seems reasonable to conjecture that the asymptotic volatility of the chain converges to 0 as $B \nearrow 0$. Conversely, we should
expect that volatility gets larger as $B$ becomes larger in absolute value (more negative), at least over some range. But we have no results of this form.

A change in $\lambda$ affects the speed at which the imitators adapt to the experimenter’s behavior. When $\lambda$ is very small (close to 0), the asymptotic volatility should be small as the chain will spend most of its time close to the origin. Intuitively, since the $y$ variable moves very slowly, the $x$ variable will be close to a best response $By$ most of the time; that is, the chain will be close to the best response line $\{(By, y)\}$. But then the tendency of the chain to move toward the 45 degree line must take it to the origin. A theorem to this effect, if such exists, would be hard to formulate and prove and, in any case, has eluded us so far.

5.3 Related Literature

Brock and Hommes (1997) study a situation which also gives rise to cycles of variable amplitude, although the mechanism that generates cycles in their model is quite different from ours. In their paper, agents can use different models to predict the future behavior of the economy. More precise models are more expensive to use and agents do a cost-benefit analysis to decide which model to use when calculating their optimal actions. When the economy is close to equilibrium, the value of information is low because the future will be similar to the present and agents are more likely to use the crude, inexpensive model. Far away from equilibrium, the changes are likely to be greater and agents are more likely to use the precise, expensive model. This switching from crude to precise models leads to an endogenous regime shift, analogous to the changes in the speeds of adjustment in our model, that produces cycles in the behavior of the economy. Even from this brief sketch, a number of important differences from our approach can be seen. The agents in the Brock-Hommes model are hypothesized to be rational, maximizing agents, even though they are boundedly rational insofar as they use imperfect models to predict the future. Also, there is no (explicitly modeled) uncertainty. Nonetheless, the two models have a similar flavor.

The comparison with the deterministic Brock-Hommes model raises the question of why we assumed random experimentation in the first place. There are good reasons why learning should be modeled as the result of random experimentation. In the first place, if agents knew exactly where to look for a better strategy, they would not be learning at all. Secondly, boundedly rational agents cannot master the computational complexity of Bayesian
learning; so we are left with the assumption that they search randomly for better strategies. Thirdly, as we have seen, our stochastic model produces effects that cannot arise from a deterministic model.

A related question is whether the cycles in the nonconvergent instances of our model should be regarded as “endogenous.” In the macroeconomic literature, models of cycles fall into two main categories. First, there are models of endogenous cycles, which use non-linearities in deterministic models to generate limit cycles and chaos (see Benhabib (1992) for a representative selection). Secondly, there is the real business cycle literature, which uses exogenous productivity shocks to generate cycles. We think of our model as belonging to the endogenous branch of the literature despite its stochastic elements. The random experimentation in our model ensures that the rate of successful experimentation varies with the distance from the equilibrium point; and the probability of finding a better strategy is sometimes high and sometimes low. As a result, experimenters are sometimes changing their strategies quickly and sometimes slowly. It is the variable rate of learning that is crucial for the dynamic properties of the model and distinguishes it from the rest of the endogenous-cycles literature. This endogenous variation in the rate of learning seems very intuitive to us, and we expect that it is likely to crop up in other contexts. It should also be noted that our stage game has a unique symmetric equilibrium (when $B \neq 1$) and that equilibrium is also the unique stationary point of our model of learning. In this sense, the stochastic process is capable of settling down; yet the system fluctuates indefinitely around the equilibrium point. For all these reasons we think it is fair to describe the fluctuations in our model as endogenous.

There is a large literature dealing with optimal experimentation (e.g., Aghion, Bolton and Harris (1991), Banks and Sundaram (1992), and Bolton and Harris (1995)). These are models of rational learning, rather than rule-of-thumb learning, and they assume that the underlying environment is stationary. As the agents’ beliefs converge to the true state of nature, they become insensitive to new information and experimentation dies out. Our model is motivated by an interest in non-stationary environments, so we assume that experimentation continues indefinitely.

Similarly, the literature on learning in games (e.g., Fudenberg and Kreps (1993), Jordan (1993), Kalai and Lehrer (1993a,b), Krishna and Sjostrom (1995), Marimon (1995), and Benaim and Hirsch (1996)) addresses the problem of how to play a best response to other players’ actions when other players are changing their actions. Since these authors assume that the un-
derlying game is fixed and are interested in convergence to equilibrium, their models also have the feature that learning eventually dies out. For example, in fictitious play or in models making use of stochastic approximation theory, the weight placed on current information declines asymptotically to zero as the length of the learning period goes to infinity. In our model, on the other hand, constant experimentation is both appropriate and crucial to the results.

In the closely related literature on evolutionary game theory, adaptation replaces learning. Evolution is analogous to group learning in which the group benefits from the experience of past generations. Similar issues arise. In particular, convergence to equilibrium implies that evolution eventually dies out.

Imitative behavior has been the subject of an extensive literature on social learning. Some of this deals with the free-rider problem that arises when there are informational externalities (e.g., Chamley and Gale (1994) and Caplin and Leahy (1994)). Although we have not modeled the free-riding decision explicitly (agents are exogenously assigned to be experimenters or imitators), the existence of free-riding imitators is crucial to the dynamic properties of the model.

The literature on social learning also deals with the problem of herd behavior, in which individuals suppress their own private information as they follow the herd (e.g., Banerjee (1992), Bikhchandani, Hirshleifer and Welch (1992), Chamley and Gale (1994), and Gul and Lundholm (1992)). This literature is firmly in the Bayesian tradition. Our bounded rationality approach makes it difficult to talk about information. However, there is an echo of herd behavior in the inertial effect of our imitators chasing the lone experimenter.

Social learning does not always lead to market failure. In some cases, it leads to optimal outcomes, at least in the limit as time goes to infinity (e.g., Ellison and Fudenberg (1993) and Banerjee and Fudenberg (1995)).

Appendix

Proof of Lemma 3
Set \( y_0 = y \), let \( y_i = y_{i-1} + \lambda (x - y_{i-1}) \) for \( i = 1, 2, \ldots \), and let \( p_i(x, y) \) denote the probability of a successful experiment on the \( i^{th} \) trial given that the previous \( i - 1 \) experiments have all been failures. Then, if \( \eta \) is sufficiently small,
\[ p_i(x, y) = \frac{\text{length of } B(x, y_{i-1}, 0)}{2}, \] and the probability that no successes have occurred in the first \( i - 1 \) experiments is
\[ q_i(x, y) = \prod_{j=1}^{i-1} (1 - p_j(x, y)). \]

Thus
\[ F(z; x, y) = \sum_{i=1}^{\infty} p_i(x, y)q_i(x, y)G(z; x, y). \]

From now on, we restrict attention to the case \( x \leq y \), the argument in the complementary case being symmetrical. There are then three regions to be considered (see Figure 3).

**Region A** \((0 < x \leq y)\)

In this region, it is immediate that \( G(z; x, y_i) \leq G(z; x, y_{i+1}) \) \( \forall z \in \mathbb{R} \). Since \( \lim_{i \to \infty} G(z; x, y_i) = G(z; x, x) \) \( \forall z \), it follows that \( F(z; x, y) \leq G(z; x, x). \)

**Region B** \((x < By, x \leq y)\)

First, recall the “two-step uniform” formula (1) for \( G(\cdot) = G(\cdot; x, x) \), with slope \((1 - B)^{-1}\) between 0 and 1. Since each \( G(\cdot; x, y) \) in this region is either a two- or one-step uniform with 1 being in the support of the first step, it follows that \( \exists M \) such that \( G(z; x, y_i) \leq (1 + M)G(z) \) \( \forall z \in \mathbb{R} \) and \( \forall i \). Next, since \( \lim_{i \to \infty} G(z; x, y_i) = G(z; x, x) \) \( \forall z \), for each \( z \) there is an integer \( k \) such that for \( i > k \), \( G(z; x, y_i) \leq (1 + (\eta/2))G(z) \). But, from a standard result in analysis, since cumulative distribution functions are monotone and since \( G \) is continuous, there is a \( k \) for which this is true uniformly in \( z \). Next, notice that by taking \( \eta \) sufficiently small, \( \sum_{i=1}^{k} p_i(x, y) \) can be made as small as we like. Combined with the \((1 + M)\) bound above, we can limit the size of the first \( k \) terms in (2) to \((\eta/2)G(z)\), and the result follows.

**Region C** \((0 > x \geq B y)\)

For small enough values of \( i \), \((x, y_i)\) remains in region C, a success increases \(|x|\) and, from the argument in region A, \( G(z; x, y_i) \leq G(z; -x, -x) = G(z) \).

For large \( i \), \((x, y_i)\) belongs to region B, so \( F(z; x, y_i) \leq (1 + \eta)G(z) \). Hence
\[ F(z; x, y) \leq \sum_{i=1}^{k} p_i(x, y)q_i(x, y)G_i(z; x, y) + q_{k+1}(x, y)F(z; x, y_{k+1}) \]
\[ \leq (1 + \eta)G(z). \]

31
Proof of Lemma 4
Substituting (1), for $B$ small and $\delta$ small,
\[
\int_0^\infty \ln zdF = \left[ \frac{1 + \delta}{2 - 2B} \right] \left[ \int_0^1 2 \ln zdz + \int_1^{1-2B-\delta} \ln zdz \right].
\]
Since $\int_0^1 \ln zdz > -\infty$, the second term dominates the first when $B$ is sufficiently small. ■

Proof of Lemma 5
The proof is by induction. For $n = 1$ there is nothing to prove. Suppose it is true for $n$. Then
\[
\Pr\left( \left| X_n \right| \leq x, \xi_{n+1} \leq r \right) = \int_{\left\{ \left| X_n \right| \leq x \right\}} \Pr(\xi_{n+1} \leq r \mid \left| X_n \right| = x) d\Pr(\left| X_n \right| \leq x)
\]
\[
\leq \int_{\left\{ \left| X_n \right| \leq x \right\}} \Pr(\xi_{n+1} \leq r \mid \left| X_n \right| = x) d\Pr(\left| X_n \right| \leq x) \text{ by hypothesis}
\]
\[
= \Pr(\xi_{n+1} \leq r) \Pr(\left| X_n \right| \leq x) \text{ by independence of the processes}
\]
\[
\leq \Pr(\xi_{n+1} \leq r) \Pr(Z_n \leq x) \text{ by the induction hypothesis}
\]
\[
= \Pr(Z_n \leq x, \xi_{n+1} \leq r) \text{ by the i.i.d. hypothesis.}
\]
Now let $r \to \infty$, and the result follows. ■

Proof of Proposition 7
If $\varphi$ is a measure on $(S, A)$, the chain $\{(X_t, Y_t)\}$ is $\varphi$-irreducible if $\forall s \in S, \varphi(A) > 0$ implies $\forall s \in S, P^n(s, A) > 0$ for some $n$, where $n$ can depend on both $s$ and $A$. The maximal such probability measure is called $\psi$, and the chain is $\psi$-irreducible (see [MT, Proposition 4.2.2]). The proof of the proposition is the next four lemmas.

Lemma 19 $\{(X_t, Y_t)\}$ is $\psi$-irreducible; and the support of $\psi$ has nonempty interior.

Proof: We shall first argue that if $(x, x)$ belongs to the set $N(\eta)$ and if $A$ is any open square centered at $(x, x)$ with sufficiently small diameter, then $\forall s \in S, P^n(s, A) > 0$ for some $n$. Starting from any $s \in A$, this is obvious.
Starting from any \( s \) in the open vertical strip \( A_1 \) (see Figure 4) generated by the square \( A \) this is true because a finite sequence of experimental failures moves any such \( s \) into \( A \), and the successive failures are independent and have probabilities bounded away from zero. Now from any \( s \in A_2 \)—the set of points that are to the right of \( A_1 \) and in the first quadrant—it is easy to see that the chain must enter the second quadrant in finite time with probability one; therefore, with positive probability, the chain enters \( A_1 \) first. From any \( s \) in \( A_3 \)—the part of the fourth quadrant to the right of \( A_1 \)—a sequence of successive experimental failures brings the chain into \( A \).

Now let \(-A\) be the mirror image of \( A \) in the third quadrant. If \( x \) is sufficiently small, there is a positive probability of moving from any \( s \in -A \) into \( A_1 \) in one step and hence (from the image of the argument in the previous paragraph) of moving from any \( s \) in or to the left of the vertical strip generated by \(-A\) into \( A \) in a finite number of steps.

Now let \( Q \subset S \) be the strip of states excluded from the constructions in the previous two paragraphs (i.e., first coordinate small in absolute value). From Theorem 2 the state leaves \( Q \) in finite time with probability one.

We have therefore showed that \( \forall s \in S, P^n(s, A) > 0 \) for some \( n \).

If \( d \) is the length of a side of \( A \), let \( z = x - (d/2) \), so that \((z, z)\) and \((z + d, z + d)\) are the two corners of \( A \) on the 45-degree line. Now let \( D \) be another square having side length \( d \), two of the corners of which are \((0, z-d)\) and \((d, z)\), and let \( \varphi \) be Lebesgue measure restricted to \( D \). (See Figure 5.) We shall show that if \( E \) is any set having positive \( \varphi \)-measure and \( a \in A \), then \( P^2(a, E) > 0 \). In light of the preceding paragraphs, this will establish the \( \varphi \)-irreducibility, and hence the \( \psi \)-irreducibility, of \( \Phi \). Since the support of \( \varphi \) has nonempty interior, so does the support of \( \psi \). Our argument requires

\[
0 < z < \frac{2}{2 + \lambda} \quad \text{and} \quad 0 < d < \frac{\lambda z}{2}.
\]

We need to find a positive lower bound for \( P^2(a, E) \). To do this we shall compute the probability of a successful experiment restricted to a particular interval in the first period, followed by the probability of a successful experiment in the second period restricted so as to carry the chain into \( E \). First notice that if the first experiment were unsuccessful, the action of the imitators would move any point \( a \) in \( A \) into a parallelogram subset of \( A \). The vertical coordinate of this point \( a' \) will define the success interval of interest; namely, the interval constructed so that the succeeding vertical move (absent a success) would then move the state onto a slanted line segment lying
to the right of $D$. For instance, if $a = (z, z)$, then $a' = a$; the first success interval is $[z - (d/\lambda), z]$ and the slanted line runs from $(z - (d/\lambda), z - d)$ to $(z, z)$. At the other extreme, if $a = (z + d, z + d)$, then $a' = a$ again; the first success interval is $[z + d - (2d/\lambda), z + d - (d/\lambda)]$, and the slanted line runs from $(z + d - (2d/\lambda), z - d)$ to $(z + d - (d/\lambda), z)$. Other cases, where $a' \neq a$, are dealt with analogously. Note that the uniform distribution on first-period experimental outcomes generates a uniform conditional density on each slanted line, and starting from such a conditional density on any of the slanted lines, the density generated by a second experimental success restricted to $D$ is again uniform, this time on $D$. So, since the first-period success intervals have lengths that are bounded away from zero, from any $a \in A$, $P^2(a, E)$ is bounded below by a positive constant multiplied by the uniform measure of $E$. Hence if $\varphi(E) > 0$, then $P^2(a, E) > 0 \forall a \in A$ and so $P^n(s, A) > 0 \forall s \in S$, for some $n.

A $\psi$-irreducible chain is a Feller chain if $\{s_n\} \to s$ implies the sequence of probability measures $\{P(s_n, \cdot)\}$ converges weakly to $P(s, \cdot)$.

**Lemma 20** $\{(X_t, Y_t)\}$ is a Feller chain.

**Proof:** It is enough to show that $\int h(s')P(s_n, ds') \to \int h(s')P(s, ds')$ for all bounded continuous functions $h : \mathbb{R}^2 \to \mathbb{R}$. Recall that the probability measure $P(s, \cdot)$ is composed of two parts—an atom at the point on the vertical line through $s$ a fraction $\lambda$ of the distance to the 45-degree line, and a uniform measure on a horizontal line segment. Since $\{s_n\} \to s$, the locations of the atomic parts converge to the atom of the limit distribution; furthermore, the probabilities assigned to the atoms in the sequence converge to that of the limiting atom. In addition, the end points of the uniform supports converge to those of the limiting measure. For bounded continuous $h$, therefore, the sequence of integrals converges to the limit integral by the dominated convergence theorem.

A set of states $A$ is Harris recurrent if, starting from any initial condition in $A$, the number of visits to $A$ is infinite with probability one. A $\psi$-irreducible chain is Harris recurrent if every set having positive $\psi$-measure is Harris recurrent.

**Lemma 21** $\{(X_t, Y_t)\}$ is Harris recurrent.
Proof: Proposition 9.1.7 and Proposition 6.2.8 of [MT] imply that \{({X}_t, {Y}_t)\} is Harris recurrent if there exists a compact set that is reached with probability one from any initial condition. We can take as our compact set \(K \setminus N'(\eta)\), where
\[
N'(\eta) = \{(x, y) | |x| \leq \eta\}.
\]

We know that \(K\) is reached with probability one, so all we need to show is that \(N'(\eta)\) is left with probability one. First note that if at the initial state \(|y| \geq 2\eta\), then after a finite number of steps if \(\Phi\) has not left \(N'(\eta)\), it must have entered \(N(\eta)\). If it subsequently leaves \(N(\eta)\), it can only be because a successful experiment takes it outside of \(N'(\eta)\). Since Theorem 2 tells us that \(\Phi\) leaves \(N(\eta)\) with probability one, we are done.

A \(\psi\)-irreducible chain has a cycle of length \(d\) if for a set \(A\) having positive \(\psi\)-measure, the greatest common divisor of all the values of \(n\) for which \(P^n(s, A) > 0, \forall s \in A\), is \(d\). To see that our chain is aperiodic—i.e., no cycles of length greater than one—it is enough to notice that once the chain enters a square centered on the 45-degree line, it can stay there for any finite number of periods with positive probability. Finally we have

Lemma 22 \{({X}_t, {Y}_t)\} is a null chain.

Proof: Recurrent chains have invariant measures that are unique up to constant multiples ([MT],Theorem 10.4.4). If the measure can be normalized to be a probability measure, the chain is positive; otherwise it is null. By Proposition 10.4.9 of [MT] any invariant measure has the same null sets as \(\psi\), and by Proposition 10.4.10 of [MT], if \(\mu\) is any invariant measure and if \(\mu(A') > 0\), then
\[
\int_{A'} \mu(ds)E_s[\tau_{A'}] < \infty
\]
if the chain is positive, where \(E_s[\tau_{A'}]\) is the expected time it takes the chain to hit \(A'\) for the first time given that it started at \(s\) and at least one unit of time has passed. We shall show that whenever \(x\) is sufficiently small, there are open subsets \(A'\) of the set \(D\) in Figure 5, which is contained in the support of \(\psi\) by the proof of Lemma 19, for which \(E_s[\tau_{A'}] = \infty \forall s \in A'\). Therefore, the above integral with \(\mu = \psi\) is infinite, thus proving that the chain is null.

Start at \((\pi, \gamma)\) \(\in D\). If \(A'\) is a sufficiently small neighborhood of \((\pi, \gamma)\), then the chain leaves \(A'\) at time \(1\), and the sign of \(X_t\) must be negative at some time \(t\) before the chain can possibly return to \(A'\). We shall show that
the expected first time for \( X_t \) to be negative is infinite, thereby establishing that \( E(\tau | \tau < \infty) \) = \( \infty \) and proving the lemma.

We begin by calculating a lower bound for \( E(\tau | (x_0, y_0)) \), the expected time till the first successful experiment starting from any state \( (x_0, y_0) \in D \). Let \( \gamma = (1 - \lambda) \); and let \( y_j = \gamma^j y_0 + (1 - \gamma^j)x_0 \), so \( (x_0, y_j) \) is the position of the chain at time \( j \) if there have been no successful experiments. Let \( p_j = By_j + x_0 \)—the probability of a success at \( (x_0, y_j) \). Then
\[
p_j = B\gamma^j(y_0 - x_0) + (B + 1)x_0.
\]

Let \( N \) be a large integer. If
\[
0 < x_0 \leq \left( \frac{1 - By_0}{B + 1} \right) \gamma^N
\] (3)
and if \( j \geq N \), then
\[
p_j \leq p_N = B\gamma^N(y_0 - x_0) + (B + 1)x_0 \leq By_0\gamma^N + (1 - By_0)\gamma^N = \gamma^N.
\]

Now replace \( p_j \) with
\[
p'_j = \begin{cases} p_j & \text{if } j \leq N \\ \gamma^N & \text{if } j > N \end{cases}
\]

Since the distribution of the first success time using \( \{p_j\} \) stochastically dominates that same distribution using \( \{p'_j\} \), if \( \tau \) is the time of the first success, then
\[
E(\tau | (x_0, y_0)) = \sum_{j=1}^{\infty} j \prod_{k=1}^{j-1} (1 - p_k)p_j \geq \sum_{j=1}^{\infty} j \prod_{k=1}^{j-1} (1 - p'_k)p'_j
\]
\[
= \sum_{j=1}^{N} j \prod_{k=1}^{j-1} (1 - p_k)p_j + \prod_{k=1}^{N} (1 - p_k) \sum_{j=1}^{\infty} (j + N)(1 - \gamma^N)^j \gamma^N.
\]

By direct calculation,
\[
N\gamma^N \sum_{j=1}^{\infty} (1 - \gamma^N)^j = N \quad \text{and} \quad \gamma^N \sum_{j=1}^{\infty} j(1 - \gamma^N)^j = \frac{1}{\gamma^N}.
\]

Thus
\[
E(\tau | (x_0, y_0)) \geq \sum_{j=1}^{N} j \prod_{k=1}^{j-1} (1 - p_k)p_j + \prod_{k=1}^{N} (1 - p_k)(N + \frac{1}{\gamma^N}). \quad (4)
\]
Using (3) again,

\[
\prod_{k=1}^{N} (1 - p_k) = \prod_{k=1}^{N} (1 - (B + 1)x_0 - \gamma^k B(y_0 - x_0)) \geq \prod_{k=1}^{N} (1 - (B + 1)x_0 - \gamma^k B y_0)
\]

\[
\geq \prod_{k=1}^{N} (1 - (1 - B y_0)\gamma^N - \gamma^k B y_0) \geq \prod_{k=1}^{N} (1 - (1 - \gamma^k B y_0)\gamma^N - \gamma^k B y_0)
\]

\[
= \prod_{k=1}^{N} (1 - \gamma^N)(1 - \gamma^k B y_0) = (1 - \gamma^N)^N \prod_{k=1}^{N} (1 - \gamma^k B y_0).
\]

We establish next that

\[
\lim_{N \to \infty} \prod_{k=1}^{N} (1 - p_k) > 0
\]

by showing that

\[
\lim_{N \to \infty} \prod_{k=1}^{N} (1 - \gamma^k B y_0) > 0 \quad \text{and} \quad \lim_{N \to \infty} (1 - \gamma^N)^N > 0.
\]

The first of these follows from standard results on the convergence of infinite products. For the second,

\[
\log(1 - \gamma^N)^N = N \sum_{k=1}^{\infty} \frac{-\gamma^{Nk}}{k} \geq -N \frac{\gamma^N}{1 - \gamma^N} \to 0.
\]

So, \( \lim_{N \to \infty} (1 - \gamma^N)^N \geq 1 \), and (5) holds. Now, from (4),

\[
E(\tau|(x_0, y_0)) \geq C(N + \frac{1}{\gamma^N}),
\]

whenever \( 0 < x_0 \leq \left(\frac{1 - B y_0}{B + 1}\right) \gamma^N \), where \( C > 0 \) is the limit in (5).

Now, beginning from \((\pi, \gamma) \in A'\) with positive probability, say \( C'\), the first successful experiment will move the \( x \)-coordinate of the chain only into the interval \([0, \pi]\), and the conditional density of such \( x \)-coordinates is uniform. Therefore, the expected time for the first sign change of the \( x \)-coordinate is at least

\[
C' \int_{0}^{\pi} \inf_{y_0 \in [\pi, \gamma]} E(\tau|(x_0, y_0)) dx_0 = C' \int_{0}^{\pi} E(\tau|(x_0, \gamma)) dx_0.
\]
But, using (6), we can proceed by constructing a step-function lower bound for the integrand; namely, if \( N \) is the largest \( N \) for which
\[
\bar{x} \leq \left( \frac{1 - B\eta}{B + 1} \right) \gamma^N,
\]
we have
\[
C' \int_0^{\bar{x}} E(\tau| (x_0, \eta)) dx_0 \geq CC' \sum_{N=N}^{\infty} (\gamma^N - \gamma^{N+1}) \left( \frac{1 - B\eta}{B + 1} \right) (N + \frac{1}{\gamma^N})
\]
\[
= CC' \sum_{N=N}^{\infty} (1 - \gamma) \left( \frac{1 - B\eta}{B + 1} \right) (N\gamma^N + 1) = \infty,
\]
and we are finished. \( \blacksquare \)

**Proof of Corollary A**

From the proof of Lemma 22, from any initial state \((\pi, \eta) \in D\) of Figure 5, the expected time till the first sign change in the first coordinate of the chain is infinite. The same is true, by the same argument, for any initial state in the open triangle \( \Delta(\eta) \) formed by \((0, 0), (\eta, \eta), \) and \((0, \eta), \) if \( \eta \) is sufficiently small. So, from such starting states, the expected time till the first departure from \( N(\eta) \) is infinite. From an initial state in \( N(\eta) \) above \( \Delta(\eta), \) consider the states to be elements of \( \Delta(2\eta) \) and repeat the argument. From an initial state in the open triangle \( \Delta'(\eta) \) formed by \((0, 0), (\eta, \eta), \) and \((\eta, 0), \) the chain remains in \( \Delta'(\eta) \) until the first success, which moves it into \( \Delta(\eta) \) with probability bounded away from zero, and hence the same conclusion holds for the expected first departure time from \( N(\eta) \). The same is obviously true of the open line segment separating the triangles. From initial states in \( N(\eta) \) and the lower right quadrant, the probability that the chain reaches \( \Delta'(\eta) \) before the first success is bounded away from zero, and the conclusion is the same. For initial states having negative first coordinates, the symmetric argument works; and for any initial state on the vertical axis, there is a positive probability of no successes ever and hence the desired conclusion holds. \( \blacksquare \)

**Proof of Lemma 13**

Suppose that \( \Phi_i \) lies above the line \( y = x + d/\lambda. \) Then the \( y \)-coordinate must fall by more than \( d, \) but the \( x \)-coordinate can fall by at most \( d. \) But the \( x \)-coordinate cannot fall unboundedly over time, because eventually it would
lie to the left of the best response line $x = By - \varepsilon$, at which point it would begin rising again. Hence $\Phi_t$ must fall below the line $y = x + d/\lambda$ in finite time. The same argument shows that if $\Phi_t$ lies below the line $y = x - d/\lambda$, then within a finite number of periods it must lie above it.

Suppose now that $\Phi_t = (x, y) \in C_d$. If $y' = \lambda x + (1 - \lambda)y$, then $(x, y') \in C_d$. If $x' \in B(x, y', \varepsilon_t)$ then $x' \in [x - d, x + d]$ and a direct calculation shows that

$$|x' - y'| \leq |x' - \lambda x - (1 - \lambda)y| \leq |x \pm d - \lambda x - (1 - \lambda)y| = |(1 - \lambda)(x - y) \pm d| \leq (1 - \lambda)d/\lambda + d = d/\lambda.$$

**Proof of Lemma 14**

We have already shown that for any sample path and any $t > T$, if $(X_t, Y_t) \in C_d$ then $(X_s, Y_s) \in C_d$ for all $s > t$. So it is only necessary to show that the $x-$coordinate of the chain enters and remains in the interval $[-x_{sd} - d, x_{sd} + d]$. Suppose that $(x, y) \in K_{sd}$ and $x \leq By + \varepsilon$. Then $x \leq x_{sd}$ and the better responses lie to the right of $x$. Since the maximum movement to the right is $d$, if the next value of the chain is $(x', y')$, say, then $x' \leq x_{sd} + d$, that is, $(x', y') \in K_{sd}$. The argument for the case $x \geq By - \varepsilon$ is symmetrical.

Because of Lemma 13 we can assume without loss of generality that the initial condition $(x_0, y_0) \in C_d$ and that $d_t \leq d$ for all $t$. Suppose that $x \leq By - \varepsilon$. (The other case is exactly symmetrical.) As long as the chain lies to the left of $K_{sd}$, the experimenter’s position can only move to the right. Since the width of $K_{sd}$ is at least $2d$, it is impossible for the chain to jump over it. If $X_{t-1} \leq -x_{sd} - d$, then

$$X_t = X_{t-1} + \max\{0, \omega_t\} = x_0 + \sum_{s=1}^{t} \max\{0, \omega_t\},$$

so the chain can fail to reach $K_{sd}$ only if $\sum_{s=1}^{\infty} \max\{0, \omega_t\} < \infty$. Thus the lemma is proved if we can show that

$$\sum_{t=1}^{\infty} \max\{0, \omega_t\} = \infty$$
with probability one. However, this follows from Kolmogorov’s “three-series theorem” [Feller (1966), p.317, Theorem 3] which tells us that the series \( \sum_{t=1}^{\infty} \max\{0, \omega_t\} \) diverges with probability one if the sum of the expectations \( \sum_{t=1}^{\infty} E[\max\{0, \omega_t\}] \) is infinite. Since \( E[\max\{0, \omega_t\}] = d_t/4 \) and we have assumed that \( \sum_{t=1}^{\infty} d_t = \infty \), the desired result is immediate.

References


