

A Probabilistic Model of Learning in Games

by

Chris W. Sanchirico, Columbia University

November 1993 Revised August 1996

Discussion Paper Series No. 9596-22

dp 9596.02
pages 53

A PROBABILISTIC MODEL OF LEARNING IN GAMES

By Chris William Sanchirico¹

November 1993
(This Draft: August 1996)

NOTE: This draft includes proofs left out of the published version, “A Probabilistic Model of Learning in Games,” *forthcoming in Econometrica*.

This paper presents a new, probabilistic model of learning in games which investigates the often stated intuition that common knowledge of strategic intent may arise from repeated interaction. The model is set in the usual repeated game framework, but the two key assumptions are framed in terms of the likelihood of beliefs and actions conditional on the history of play. The first assumption formalizes the basic intuition of the learning approach; the second, the indeterminacy that inspired resort to learning models in the first place. Together the assumptions imply that, almost surely, play will remain almost always within one of the stage game’s “minimal inclusive sets.” In important classes of games, including those with strategic complementarities, potential functions, and bandwagon effects, all such sets are singleton Nash.

KEYWORDS: Learning in Games, Rationalizability

For almost half a century Nash equilibrium has been game theory's predominant solution concept. Yet in recent years foundational research on games has increasingly focused on the need to shore up the justification for equilibrium's fundamental assumption: in plain terms, that players correctly guess their opponents' strategies. Figuring large in this new literature is a resurgence in research on learning in games. The learning approach rests on a simple intuition: namely that players who play together repeatedly will eventually reach a common understanding of their strategic intentions. Translating this intuition into concrete convergence results, however, has proven to be no simple matter.

Consider, for example, "fictitious play." In this leading model of learning in games, agents' beliefs about their opponents' current actions are assumed to equal the empirical frequency of past opponent play. Agents, moreover, act "myopically" in choosing to play a best response to such beliefs without regard to the effect on their opponents' future beliefs. Shapley's (1964) well known example showed that the resulting empirical frequencies of play do not generally converge. More recent research, however, has focused on the model's proclivity to generate non-convergent sequences of actual play even when frequencies do converge. Such is the case in GAME 1.² A simple geometric argument confirms that for a range of initial conditions the generated sequence of actions fails to converge to the game's unique pure equilibrium, (Heads, Out), but instead behaves as if the game consisted solely of its Matching Pennies component (shaded). Beliefs cycle toward the Matching Pennies mixed equilibrium while actual play cycles among the four action profiles.

[INSERT Figure 1: Game 1]

Convergence of actual play, here and in general, requires the coordination of two complementary phenomena. First, (Heads, Out) must be in some sense “absorbing.” Second, there must be sufficient “entropy” in the system to insure that play reaches (Heads, Out) often enough for this absorption to occur. Under fictitious play, strict equilibria like (Heads, Out) are in fact immediately absorbing in the sense that if they are ever played, they are played forever after. Yet fictitious play fails to converge because the same rigid structure that produces absorption makes the process prone to eternally ignore its absorptive states.

The central problem of learning in games is to simultaneously generate both of these seemingly contradictory forces—absorption and entropy—in general games and from reasonable assumptions about how the history of play affects current beliefs. Such is the design here. The paper introduces a class of models defined by two mutually consistent assumptions, one each for absorption and entropy. Impotent on their own, the assumptions combine to imply convergence of actual play to one of the stage game’s “minimal inclusive sets” (Basu and Weibull (1991)). Roughly, a minimal inclusive set is one that includes all its own best responses, but no other sets with the same property. In GAME 1, for instance, {(Heads, Out)} is the only minimal inclusive set. More generally, such sets are generically singleton, Nash in games with strategic complementarities, potential functions, identical interests or bandwagon effects. At the other extreme, in Shapley’s example and matching pennies, the whole set of profiles is minimal inclusive; the paper leaves open the question of whether and how a common understanding of intent could develop in such “irreducible” games.

An example of a process satisfying the two assumptions for GAME 1 will help to introduce the general approach, describe the assumptions, and explain how they imply convergence. Given

$0 < \lambda \leq 1$, define for each player i and each subset of stage game action profiles,³ $E \subseteq A$ the set $\Delta_i^\lambda(E) = \{\psi_{-i} \in \Delta(A_{-i}) \mid \psi_{-i}(E_{-i}) \geq \lambda\}$, where $\Delta(A_{-i})$ represents the set of probability measures on the set of opponent actions, A_{-i} . The set $\Delta_i^\lambda(E)$ represents the event that i thinks her opponent likely to play within E_{-i} , with “likely” defined by λ . For each subset $B \subseteq \Delta(A_{-i})$ of stage game beliefs for i let $u_i(B)$ represent the measure placing unit weight uniformly on B and zero weight elsewhere. Combining notation, consider for a given history $\{a^1, \dots, a^{t-1}\}$, the measure⁴ $u_i(\Delta_i^\lambda(a^{t-r}, \dots, a^{t-1}))$. The measure puts probability one on the event that i believes her opponent likely to repeat an action he has taken within the last r periods.⁵ All beliefs consistent with this event are regarded as equally likely. The graph marked “A” in Figure 2, for instance, shows the frequency distribution of $u_{col}(\Delta_{col}^\lambda(\dots))$ when $r = 2$ and Row has just played Tails twice in a row.

In this example the measure describing i 's current beliefs after the partial history

$\{a^1, \dots, a^{t-1}\}$ is a convex combination of the measure $u_i(\Delta_i^\lambda(a^{t-r}, \dots, a^{t-1}))$ and a one period lag⁶:

$$(1) \quad p_i(a^1, \dots, a^{t-1}) = \alpha p_i(a^1, \dots, a^{t-2}) + (1 - \alpha) u_i(\Delta_i^\lambda(a^{t-r}, \dots, a^{t-1})).$$

Thus, the measure after the history $\{a^1, \dots, a^t, \dots, a^{t-1}\}$ is just a geometric average of past

$u_i(\Delta_i^\lambda(a^{t-r}, \dots, a^{t-1}))$'s. Figure 2, for example, shows the progression of $u_{col}(\Delta_{col}^\lambda(a^{t-r}, \dots, a^{t-1}))$

and $p_{col}(a^1, \dots, a^{t-1})$ for a given history of play for Row, starting from given initial conditions,

with $\alpha = \frac{1}{2}$ and $r = 2$. Finally, assume that players' current beliefs are drawn independently at

each partial history and that both players play a myopic best response to their current beliefs.

[INSERT Figure 2]

These specifications define a probability measure P on the set of all sequences $\{\psi', a'\}$ of beliefs and actions.⁷ The theorems of this paper imply that if we set $\lambda \geq \frac{3}{4}$ (which we do hereinafter), then P converges to (Heads, Out) in the sense that almost surely (Heads, Out) will be played almost always. In other words, P assigns probability one to the set of all sequences $\{\psi', a'\}$ in which a' remains at (Heads, Out) forever after some point.

Before explaining how this result obtains, it is worth emphasizing an important difference between this and more traditional learning models. Notice that the measure $p_i(a^1, \dots, a^{t-1})$ is not player i 's belief about her opponent's choice at t , but rather, a probability measure on such beliefs. While learning rules such as fictitious play stipulate for each history what players' beliefs certainly are, this model specifies a measure describing what such beliefs tend to be. This probabilistic approach rests on three general assertions:

1) Placing “extra-rational” restrictions on beliefs is unavoidable given the indeterminacy of rational strategic interaction. A corollary to the literature on extensive form rationalizability (Pearce (1984)) is that rationality alone will not yield convergence. In GAME 1, for instance, wherein all actions are stage game rationalizable, any sequence of action profiles $\{a^t\}$ is consistent with common knowledge of rationality in the extensive form of the repeated game.⁸

2) Such “extra-rational” restrictions should reflect the indeterminacy that necessitates their use. The usual way to impose extra-rational restrictions is to stipulate that players form their beliefs according to a particular learning rule, such as fictitious play. In contrast, the probabilistic approach introduced herein eschews rigid, formulaic algorithms in favor of probabilistic statements about the tendency of players to think the past repeats itself. The

necessary restrictions are thus cast in a manner that acknowledges and incorporates our agnosticism about how players form their beliefs. (C.f., Gul (1991) and Milgrom and Roberts (1991).)

3) Formalizing indeterminacy in a probabilistic framework is a natural way to generate entropy. If we believe that many things are possible at each time t , then we must also believe it likely that many things will happen over time. Thus, to the extent that convergence requires both absorption and entropy, the use of probability theory makes the model not only more palatable, but also more effective in generating convergence.

Returning to GAME 1, convergence here follows from two intermediate results, which follow in turn from two properties of the process⁹ (1). The first result is that (Heads, Out) is *eventually absorbing*: conditional on the event that (Heads, Out) is played infinitely often, it will be played almost always (i.e., always, after some point) with probability 1. Eventual absorption is in turn a consequence of the manner in which (Heads, Out) feeds back on itself under this process.

Feedback follows from jointly from (1) and the best response properties of (Heads, Out). The more (Heads, Out) is played, the more often and predominant it is in recent history, $(a^{t-r}, \dots, a^{t-1})$, and so—since each player i 's current beliefs $p_i(a^1, \dots, a^{t-1})$ are a geometric average of past $u_i(\Delta_i^\lambda(a^{t-r}, \dots, a^{t-1}))$ measures—the more probable that both players think their opponent likely to play (Heads, Out) again. This much is true of all action profiles. Since (Heads, Out) is an equilibrium (and λ is large enough), each player does in fact repeat her part of (Heads, Out) when she thinks her opponent likely to do the same. Hence, the feedback: the more (Heads, Out) is played, the more likely it is to be played yet again.

Feedback per se, however, is not enough to generate eventual absorption. For example, the probability that (Heads, Out) is repeated might increase to an asymptote of $\frac{1}{2}$, implying that almost surely (Heads, Out) is infinitely often not played. Indeed, the bare fact that the probability of (Heads, Out) approaches 1 is also not sufficient. THEOREM 1, however, establishes that eventual absorption does follow if the feedback is *uniformly summable*: i.e., the probability that players think (Heads, Out) likely to recur always increases (in the number of times in a row n that it has played) faster than $1 - x_n$ for some uniformly chosen (across t), nonnegative, summable sequence $\{x_n\}$. Such is the case here, where, as the reader can check, after $n + r - 1$ plays of (Heads, Out), the chance that the players jointly think it likely to recur is always at least
$$\left[(1 - \alpha)(1 + \alpha + \dots + \alpha^{n-1}) \right]^2 = [1 - \alpha^n]^2 = 1 - (2\alpha^n - \alpha^{2n}).$$

The proof that eventual absorption follows from uniformly summable feedback has two steps. By an argument related to the second Borel-Cantelli Lemma, the summability condition just discussed implies that every time we arrive at (Heads, Out) there is some chance we stay there forever. Since the summability is uniform, this chance is uniform as well. An argument from the first Borel-Cantelli Lemma then establishes that this small chance of absorption on each arrival translates into a long run certainty, so long as we arrive at (Heads, Out) sufficiently often.

Arriving at (Heads, Out) sufficiently often is the role of this process' second key property, *best response entropy*: namely, if the profile a is a best response (for both players) to \tilde{a} , and \tilde{a} is played at t , then the chance that a will be played at $t+1$ is never less than $\left(\frac{1}{16}(1 - \alpha)\right)^2$, i.e., the chance is uniformly (over t) bounded away from zero. The indeterminacy of rational strategic interaction might suggest that anything is possible at all times. Applied literally, this precludes any form of convergence. Best response entropy insists only that best responses to recent

history—actions that are in a sense still “in play”—be regarded as possible. The degree of entropy over current play is thus a function of whether recent history has many or few best responses, which depends in turn on whether recent history is itself diffuse or concentrated.

That best response entropy implies infinite plays of (Heads, Out) is the content of THEOREM 2. For intuition, note that if (Heads, Out) is not played infinitely many times, some other profile in this finite game, say (Tails, Tails), must be. But, because (Tails, Heads) is a best response to (Tails, Tails), each time the former is played, there is a chance the latter is played in the following period, implying that (Tails, Heads) is also played infinitely often. Continuing the argument to (Heads, Heads) and then (Heads, Out) we obtain a contradiction. There being such a “best response chain” from all profiles to (Heads, Out), the result follows.

Taken alone, uniformly summable feedback is consistent with the performance of fictitious play in GAME 1. Best response entropy on its own is consistent with drawing action profiles in i.i.d. fashion. Together, however, the properties imply almost sure, almost always convergence to (Heads, Out). Best response entropy gives $\Pr((\text{Heads, Out}) \text{ i.o.}) = 1$. Uniformly summable feedback yields $\Pr((\text{Heads, Out}) \text{ a.a.} | (\text{Heads, Out}) \text{ i.o.}) = 1$. The product of these is the convergence result.

It is worth noting that this convergence is not a special case of Kalai and Lehrer’s (1993) model of “rational learning.” Indeed, so long as $\lambda < 1$, one can show that players will almost surely not put positive weight on the true path of play--what is required by rational learning’s absolute continuity assumption in this context. Yet the main point of comparison with rational learning is perhaps more methodological than technical. Arguably, rational learning’s assumption that i puts positive weight on the true path of play¹⁰ is really just another way of

saying that as play unfolds, i becomes more and more certain and correct in her beliefs about the future course of the game¹¹—thus begging the question of why this might occur. In contrast, this paper is an explicit attempt to explain convergence. The result is not in any sense a mathematical restatement of the fact of convergence, but rather a mathematical formalization of an explanation for convergence that is fundamentally behavioral: namely, that the right combination of indeterminacy, on the one hand, and a self-intensifying tendency for players to think that history repeats itself, on the other, will lead players over time to a common understanding of their strategic intentions.

The theorems and lemmas of this paper generalize this introductory example along several dimensions. First, the model captures the two highlighted properties of this example in two assumptions on general measures over paths of play and beliefs. Any measure satisfying these assumptions is shown to converge. Second, as discussed, convergence is shown for general games, to one of the stage game's minimal inclusive sets. Third, in the general model it is the probability that a subset is “salient” that increases as the subset is played repeatedly, not necessarily the probability that players think it likely to recur. As explained within, salience generalizes the latter to full hierarchies of beliefs. A final generalization—not included here—is that players need not be myopic. As shown in the first Appendix, it is enough that discount factors are sufficiently small—i.e. that players are “sufficiently patient.” (Note that no matter how small a player's positive discount factor, her current beliefs may be so close to the break-even between two actions that future effects are decisive in her current choice.)

It is possible to construct many other examples satisfying the assumptions of the general model. To come full circle, we can use the framework to alter fictitious play so as to improve its performance in GAME 1. We make the weights geometric, rather than arithmetic, and

simultaneously add some constant positive probability that players play last period best response instead of what is dictated by the re-weighted fictitious play beliefs:

$$\mu'_{-i} = \begin{cases} \beta\mu'_{-i} + (1-\beta)a'^{-1}_{-i}, & \text{with probability } \alpha \\ a'^{-1}_{-i}, & \text{with probability } (1-\alpha) \end{cases}$$

where here “ a'^{-1}_{-i} ” represents the belief putting unit weight on the profile a'^{-1}_{-i} . This system of measures clearly satisfies best response entropy and, proceeding as if $\alpha = 1$, one can show that

$$x_n = \begin{cases} 0, & \text{if } n < \frac{\ln \frac{1}{\lambda}}{\ln \beta} \\ 1, & \text{otherwise} \end{cases}$$

is a uniform lower bound on the probability that the players think it likely that (Heads, Out) will be repeated again after it has been played n times in a row. ($n < \frac{\ln \frac{1}{\lambda}}{\ln \beta} \Leftrightarrow 1 - \beta^n \geq \frac{3}{4} = \lambda$.)

The bounding sequence here is almost always zero: if (Heads, Out) is ever played $\frac{\ln \frac{1}{\lambda}}{\ln \beta}$ times in a row, it is played forever after. Contrast this “lock-in” dynamic with the example above wherein there is always some chance of not playing (Heads, Out) no matter how many times it has been played. Hurkens’ (1994) model of learning by forgetful players also satisfies the assumptions of the general model with a bounding sequence that is almost always zero--always zero for all n larger than the bound on memory. Sonsino (1994) generates convergence to patterns of play (see the conclusion for more on patterns); convergence there also operates by a similar “lock-in” dynamic.

Section 2 of the paper sets out the general framework of the model. Section 3 proves the two intermediate results. The main convergence result is proven in Section 4, which also discusses the consistency of the assumptions. Section 5 concerns the size of minimal inclusive sets in special classes of games and Section 6 concludes the paper. A series of appendices

contain proofs that are referred to in the main body of the paper and have been left out of the published version of the paper.

1. GENERAL FRAMEWORK AND ASSUMPTIONS

Fix a stage game $G = (A_1, \dots, A_n; \pi_1, \dots, \pi_n)$, where A_i is player i 's finite set of *actions* and $\pi_i: A_1 \times \dots \times A_n \rightarrow \mathfrak{R}$ is i 's *payoff function*. For any subset $E_{-i} \subseteq A_{-i}$ of opponent action profiles,¹² let $\Delta(E_{-i})$ denote the set of all probability measures ψ_{-i} on A_{-i} with $\psi_{-i}(E_{-i}) = 1$. Extend π_i to an *expected payoff function* $u_i: A_i \times \Delta(A_{-i}) \rightarrow \mathfrak{R}$ in the usual manner. Denote the set of (*stage game*) *best responses for player i to the belief $\psi_{-i} \in \Delta(A_{-i})$* as $b_i(\psi_{-i})$. The set of (*stage game*) *best responses for player i to beliefs on any subset $E_{-i} \subseteq A_{-i}$* is

$b_i \circ \Delta(E_{-i}) \equiv \bigcup_{\psi_{-i} \in \Delta(E_{-i})} b_i(\psi_{-i})$. Finally, for any subset $E \subseteq A$ of action profiles whether or not

rectangular, define $b \circ \Delta(E) = (b_1 \circ \Delta(E_{-1}), \dots, b_n \circ \Delta(E_{-n}))$.

The model's assumptions concern the manner in which the history of play affects the likelihood of players' belief hierarchies regarding opponents' current actions. A formal statement of these assumptions, then, requires both a definition of such belief hierarchies and a probability space in which to cast statement about likelihoods. For the first task we borrow from Tan and Werlang's (1988) adaptation of "types" to uncertainty regarding strategic intent, let Θ_i denote the topological space of *stage game types* for player i with respect to the set of opponent action profiles A_{-i} , as in their DEFINITION 3.9. For the second task we provide

DEFINITION 1: Define the *probability space of action/ belief paths* for the game G to be the tuple $([\Theta \times A]^\infty, \mathfrak{F}, P)$, where: i) $[\Theta \times A]^\infty$ denotes the set of all sequences $\{\theta^t, a^t\}$ of profiles of stage game types and actions, ii) \mathfrak{F} is the product σ -algebra on $[\Theta \times A]^\infty$ constructed from the Borel sets on each copy of Θ and the power set on each copy of A , and iii) $P: \mathfrak{F} \rightarrow \mathfrak{R}$ is a probability measure.¹³

The object here is to generate common knowledge of strategic intent from repeated play; common knowledge of rationality is assumed from the onset. ASSUMPTION 0, which translates the assumption of common knowledge of rationality into our probability space, borrows more from Tan and Werlang (1988). First, the subset $K_i \subseteq \Theta_i$ (from their DEFINITION 5.2) represents the set of all types for player i consistent with common knowledge of rationality. Second, since by their THEOREM 3.1 each θ_i may be regarded as a probability measure on $A_{-i} \times \Theta_{-i}$, let $\theta_i(A_{-i})$ and $\theta_i(\Theta_{-i})$ denote the marginal of θ_i on A_{-i} and Θ_{-i} , respectively.

$$\text{ASSUMPTION}^{14} 0: \forall t \geq 1, \forall i = 1, \dots, n, P\left(\left[a^t \in b_i(\theta_i(A_{-i}))\right] \cap \left[\theta_i \in K_i\right]\right) = 1.$$

In the introduction's example the history of play affected the probability that players thought a given subset E_{-i} of profiles (there, a singleton) likely to recur. In general, it need only affect the probability that the set is "salient:" every player i either believes it likely that his opponents will play in E_{-i} in the incipient stage game, or believes it likely that his opponents hold such beliefs, or, believes it likely that his opponents believe *their* opponents hold such beliefs, or,

etc... up to any order. (For simplicity the parameter λ and hence the qualifier “likely” is left out of the formal definition of salience. The generalization is easily conceived, yet tedious to denote.)

For any product $X_1 \times \dots \times X_m$, and any subset S_k of any factor X_k , let $\langle S_k \rangle$ denote the “slab” of S_k , that is, the subset $\{x \in X_1 \times \dots \times X_m \mid x_k \in S_k\}$.

DEFINITION 2: Fix a rectangular subset of action profiles $E = E_1 \times \dots \times E_n \subseteq A$. Define, for all i , the set $S_i(1)(E) = \{\theta_i \in \Theta_i \mid \theta_i(A_{-i}) \in \Delta(E_{-i})\}$. Continuing inductively, given $S_j(n-1)(E)$ for each player j , define

$$S_i(n)(E) = \left\{ \theta_i \in \Theta_i \mid \forall j \neq i, \forall a_j \in \text{supp} \theta_i(A_j), \text{either } a_j \in E_j \text{ or } \theta_i(\langle \{a_j\} \rangle \cap \langle S_j(n-1)(E) \rangle) > 0 \right\}.$$

Then, define for all i , the set $S_i(E) = \bigcup_{n=1}^{\infty} S_i(n)(E)$. Lastly, define $S(E) \equiv S_1(E) \times \dots \times S_n(E)$. The subset E is said to be *salient* at time t , if and only if $\theta^t \in S(E)$.

ASSUMPTION 1 is the source of the uniformly summable feedback discussed in the introduction. ASSUMPTION 2 is the source of best response entropy. Both assumptions are parameterized, the former by the class of subsets to which it applies, the latter by the length of recent history, ρ . This parameterization allows for two modes of convergence in the main theorem. For each $\Gamma \subseteq 2^A$ define:

ASSUMPTION 1(Γ): For all subsets of action profiles $E \in \Gamma$, there exists a summable sequence $\{x_n\}$ such that:

$$\forall t \geq 1, \forall 1 \leq n \leq t-2, \quad P(\theta^t \in S(E) | a^{t-1}, a^{t-2}, \dots, a^{t-n} \in E; a^{t-n-1} \notin E) \geq 1 - x_n, \text{ if defined (2)}$$

For each $\rho \geq 1$ define:

ASSUMPTION 2(ρ): There exists $\varepsilon > 0$ such that for all $t \geq 1$ and all $\{a^1, \dots, a^{t-1}\}$, if $a \in b \circ (\{a^{t-\rho}, \dots, a^{t-1}\})$, then $P(a^t = a | \{a^1, \dots, a^{t-1}\}) \geq \varepsilon$, if defined.

2. INTERMEDIATE RESULTS

2.1 Inclusive Sets, Feedback and Eventual Absorption

Following Basu and Weibull (1991) let us say that a nonempty subset of action profiles $E = E_1 \times \dots \times E_n \subseteq A$ is (best response) inclusive, if $b_i \circ \Delta(E_{-i}) \subseteq E_i, \forall i$. In GAME 1 the entire set of profiles and (Heads, Out) as a singleton are the only inclusive sets. The importance of inclusive sets in this model of learning lies in the following lemma, which says that if rationality is common knowledge, then whenever an inclusive set is salient it will in fact be played in. The lemma thus establishes that inclusive sets feed back on themselves under ASSUMPTION 1. That this implies eventual absorption is the content of THEOREM 1.

LEMMA 1: If P satisfies ASSUMPTION 0 and I is inclusive, then for all $t \geq 1$,

$$P(a^t \in I | \theta^t \in S(I)) = 1, \text{ if defined.}$$

The proof, which appears in an appendix, is inductive on orders of belief. For intuition, note that if i believes her opponents will play in I , then she, being rational, will herself play in I , since I contains all best responses to itself. Similarly, if i believes both that her opponents are rational and that they think their opponents will play in I , then she must think that her opponents will themselves play in I . Then, again, she will play in I .

Let $[a' \in I \text{ i.o.}]$ denote the event that play is in I “infinitely often”--in set notation

$$\bigcap_{t=1}^{\infty} \bigcup_{s=t}^{\infty} [a^s \in I]. \text{ Let } [a' \in I \text{ a.a.}] \text{ be the event that play is in } I \text{ “almost always,” } \bigcup_{t=1}^{\infty} \bigcap_{s=t}^{\infty} [a^s \in I].$$

THEOREM 1 [Eventual Absorption]: *For all $\rho \geq 1$ and all $\Gamma \subseteq 2^A$, if ASSUMPTIONS 0, 1(I) and 2(ρ) hold and I is an inclusive set in Γ , then $P(a' \in I \text{ a.a.} | a' \in I \text{ i.o.}) = 1$, if defined.*

A sketch of the following proof appears in the introduction.

Proof: Let I satisfy ASSUMPTION 1 with sequence $\{x_n\}$. From LEMMA 1 it follows that $\forall t \geq 1$,

$$\forall 1 \leq n \leq t - 2$$

$$(3) \quad P(a', \dots, a^{t-n} \in I; a^{t-n-1} \notin I) \geq P(\theta' \in S(I); a^{t-1}, \dots, a^{t-n} \in I; a^{t-n-1} \notin I).$$

Combining (3) with ASSUMPTION 1 yields: $\forall t \geq 1, \forall 1 \leq n \leq t - 2$

$$(4) \quad P(a', \dots, a^{t-n} \in I; a^{t-n-1} \notin I) \geq (1 - x_n) P(a^{t-1}, \dots, a^{t-n} \in I; a^{t-n-1} \notin I).$$

Re-indexing, write (4) as: $\forall t \geq 1, \forall n \geq 1$

$$(5) \quad P(a^{t+n+1}, \dots, a^{t+1} \in I; a^t \notin I) \geq (1 - x_n) P(a^{t+n+1}, \dots, a^{t+1} \in I; a^t \notin I).$$

By ASSUMPTION 2 and the fact that I is inclusive, we may take $x_n < 1$, all n . Now for any t , the family of inequalities in (5) indexed by n yields, by iterative substitution:

$$(6) \quad P\left(\bigcap_{n=1}^m I^{t+n+1} \cap I^{t+1} - I^t\right) \geq \prod_{n=1}^m (1 - x_n) P(I^{t+1} - I^t)$$

where I have written, and will henceforth write I^t for the event $[a^t \in I]$. The fact that

$\{x_n\} \subseteq [0,1)$ and $\sum x_n < \infty$ implies that $\lim_{m \rightarrow \infty} \prod_{n=1}^m (1 - x_n)$ exists and is a strictly positive number,

call it ξ . (See, e.g., Knopp (1971)) Taking the limit of both sides in (6) yields: $\forall t \geq 1$,

$$(7) \quad P\left(\bigcap_{n=1}^{\infty} I^{t+n+1} \cap I^{t+1} - I^t\right) \geq \xi \cdot P(I^{t+1} - I^t).$$

Now the sequence of sets $\left\{\bigcap_{n=1}^{\infty} I^{t+n+1} \cap I^{t+1} - I^t\right\}_{t=1}^{\infty}$ is disjoint. Hence, summing the left side of (7)

over all $t \geq 1$ yields a number less than 1. Then since $\xi > 0$, (7) implies $\sum_{t=1}^{\infty} P(I^{t+1} - I^t) < \infty$.

Therefore, by the first Borel-Cantelli lemma $P\left(\left(I^{t+1} - I^t\right) \text{ i.o.}\right) = 0$. A standard argument shows

$[I^t \text{ i.o.}] - [I^t \text{ a.a.}] \subseteq \left[\left(I^{t+1} - I^t\right) \text{ i.o.}\right]$ and the result follows. ■

2.2 Plateaus, Entropy and Eventual Repulsion

An inclusive set is said to be *minimal* if it does not strictly contain another inclusive set.

The *plateau* of an inclusive set is constructed by removing from the inclusive set all smaller inclusive sets nested therein. (By convention the plateau of a minimal inclusive set is the empty set.) Thus, the plateau of the entire set of profiles in GAME 1 is the Matching Pennies component plus (Tails, Out). In GAME 1, but not in general, the plateau of the entire set of profiles

corresponds to the *grand plateau*: what remains of the entire set of profiles when we remove all minimal inclusive sets.

Plateaus, and perhaps also the grand plateau, will be eventually repelling if recent history is sufficiently long. What length suffices depends on certain properties of the game's best response correspondence, which are summarized the following notion of "size." (All results hold if we take the size of the game to be the number of profiles.)

DEFINITION 3: For all subsets $E \subseteq A$ and all action profiles $a \in E$, a *best response chain* from a to E is a finite sequence of action profiles $\{a(1), \dots, a(n)\}$ satisfying: (1) $a(1) = a$, (2) $\forall 1 \leq k \leq n$, $a(k) \in b \circ \Delta(\{a(1), \dots, a(k-1)\})$, and (3) $a(n) \in E$. The *span* of the chain $\{a(1), \dots, a(n)\}$ is defined as $\max_{1 \leq k \leq n} \{ \min j : a(k) \in b \circ \Delta(\{a(k-j), \dots, a(k-1)\}) \}$. Let a be an element of a nonempty plateau E . The *size of the profile*, a , is the smallest span across all chains from a to $\neg E$. The *size of the game* $s(G)$ is the largest size across all profiles.

Recall that in discussing the application of THEOREM 2 to GAME 1, we noted the existence of a best response chain (of span 1) from every profile to (Heads, Out). More generally, one can show for all finite games the existence of such a chain (not necessarily of span 1) from every point in every plateau to that plateau's complement. This insures that "size" is well defined and, though implicit in the proof, is central to eventual repulsion.

THEOREM 2 [Eventual Repulsion from Plateaus]: *Let P satisfy ASSUMPTION 2(ρ). (1) If $\rho \geq s(G)$, then for all plateaus $E \subseteq A$, $P(a' \in E \text{ a.a.}) = 0$. (2) If $\rho \geq |A|$, then $P(a' \in \Pi \text{ a.a.}) = 0$, where Π is the grand plateau.*

A sketch of the following proof appears in the introduction.

Proof: I prove only part (1) of the theorem. Part (2) follows in the same manner. Since the plateau $E \subseteq A$ is finite, at least one of its members is played infinitely often, so that

$[E' a.a.] = [E' a.a.] \cap \bigcup_{a \in E} [a' = a \text{ i.o.}]$. Then by the subadditivity of P

$$(8) \quad P(E' a.a.) \leq \sum_{a \in E} P([a' = a \text{ i.o.}] \cap [E' a.a.]).$$

Now take any $a \in E$. Let $\{a(1), \dots, a(n)\}$ be a chain from $a(1) = a$ to $\neg E$ that has the smallest span of all chains from a to $\neg E$ (i.e. a chain whose span is the size of a). We show

$$(9) \quad P(\{a^{t-n}, \dots, a^{t-1}\} = \{a(1), \dots, a(n)\} \text{ i.o.}) = P(a' = a \text{ i.o.}).$$

The “ \leq ” direction is obvious since $a(1) = a$. To show “ \geq ” suppose, contra, that $r < n$ is the largest index for which (9) does hold with “ \geq ”. Now $a(r+1) \in b \circ \Delta(\{a(r-\rho+1), \dots, a(r)\})$, since $\{a(1), \dots, a(n)\}$ is a chain with smallest span among those from a to $\neg E$ and $s(G) \leq \rho$.

Hence, ASSUMPTION 2(ρ) insures the existence of $\varepsilon > 0$ such that $\forall t \geq 1$ and each individual history $\{a^1, \dots, a^{t-1}\}$ that has $\{a^{t-r}, \dots, a^{t-1}\} = \{a(1), \dots, a(r)\}$, $P(a' = a(r+1) | \{a^1, \dots, a^{t-1}\}) \geq \varepsilon$.

Therefore, by a standard result (See the Appendix),

$$P(\{a^{t-r}, \dots, a^{t-1}\} = \{a(1), \dots, a(r)\} \text{ i.o.})$$

$$\begin{aligned}
&= P\left(\left([a' = a(r+1)] \cap [\{a^{t-r}, \dots, a^{t-1}\} = \{a(1), \dots, a(r)\}]\right) \text{ i.o.}\right) \\
&= P\left(\{a^{t-r}, \dots, a^t\} = \{a(1), \dots, a(r+1)\} \text{ i.o.}\right),
\end{aligned}$$

contradicting our supposition. This proves equation (9), which in turn implies

$$P\left([a' = a \text{ i.o.}] \cap [E' \text{ a.a.}]\right) = P\left([\{a^{t-n}, \dots, a^{t-1}\} = \{a(1), \dots, a(n)\}] \text{ i.o.} \cap [E' \text{ a.a.}]\right)$$

But since $a(n) \notin E$, $P\left([\{a^{t-n}, \dots, a^{t-1}\} = \{a(1), \dots, a(n)\}] \text{ i.o.} \cap [E' \text{ a.a.}]\right) = 0$, and so

$P\left([a' = a \text{ i.o.}] \cap [E' \text{ a.a.}]\right) = 0$ also. This holding for all $a \in E$, the result follows from (8). ■

3. CONVERGENCE AND CONSISTENCY

Neither ASSUMPTION 1 nor 2 guarantees convergence on its own. Fictitious play's performance in GAME 1--as observed in the opening paragraphs of this paper--is fully consistent with ASSUMPTIONS 0 and 1(Γ), even with Γ taken to be the power set of A . Moreover, any probability measure that draws actions from GAME 1 in i.i.d. uniform fashion is consistent with ASSUMPTIONS 0 and 2(ρ), for any ρ .

The interaction of these assumptions, however, produces a strong form of convergence to minimal inclusive sets, as proven below in THEOREM 3. The theorem is perhaps best understood visually. Figure 3 depicts a stage game ((0,0) payoffs are not shown) as a contour map with smaller inclusive sets marked with darker shading. Figure 4 translates this contour map into three dimensions (making clear the choice of the term "plateau.") The second intermediate result, eventual repulsion guarantees that we do not remain on the highest plateau almost always, implying that we are infinitely often in one of its "holes." In particular, calling the smaller hole S

and the larger L , eventual repulsion yields $P([a' \in S i.o.] \cup [a' \in L i.o.]) = 1$. (Note that this is not the same as “ $P([a' \in S i.o.]) = 1$ or $P([a' \in L i.o.]) = 1$.”). Now split the event $[a' \in S i.o.] \cup [a' \in L i.o.]$ into two (intersecting) sections, $[a' \in S i.o.]$ and $[a' \in L i.o.]$, and consider first $[a' \in S i.o.]$. The first intermediate result, eventual absorption says that *conditional* on $[a' \in S i.o.]$, the event $[a' \in S a.a.]$ has probability one. Similarly, conditional on $[a' \in L i.o.]$, $[a' \in L a.a.]$ receives probability one. Together with $P([a' \in S i.o.] \cup [a' \in L i.o.]) = 1$, these conditions imply (with some Boolean manipulation) that $P([a' \in S a.a.] \cup [a' \in L a.a.]) = 1$. In words, probability is divided between those sequences of play that stay in the larger hole always after some point and those sequences that stay in the smaller hole always after some point. Now, conditional on either of these (disjoint) “almost always” events we reapply eventual repulsion and absorption in parallel fashion to whatever holes (even smaller inclusive sets) there may be in either S and L , respectively. (Since there are no holes in S , we work again with S itself.) When we eventually reach a level at which none of the inclusive sets has a hole, we have established that there is probability one on those sequences of play that after some point remain in one of the stage game’s minimal inclusive sets.

[Insert Figure 3 and Figure 4, side by side]

This describes the process of convergence when ASSUMPTION 2 holds with recent history of relatively short duration and ASSUMPTION 1 applies to all inclusive sets. Only one iteration is necessary if recent history is long enough to guarantee that the grand plateau does not absorb, in which case ASSUMPTION 1 need only be applied to minimal inclusive sets.

THEOREM 3 [Convergence]: *Let P satisfy ASSUMPTIONS 0, 1(Γ), and 2(ρ). If either (1) Γ contains all G 's inclusive sets and $\rho \geq s(G)$, or (2) Γ contains all G 's minimal inclusive sets and $\rho \geq |A|$, then $P\left(\bigcup\left\{\left[a' \in \underline{I} \text{ a.a.}\right] \mid \underline{I} \text{ is minimal inclusive}\right\}\right) = 1$.*

Proof: Case (2) follows directly from THEOREMS 1 and 2. For case (1) see the Appendix.

If ρ is smaller than the size of the game, then ASSUMPTIONS 0, 1(Γ) and 2(ρ) do not imply convergence. GAME 2, for instance, is of size 2. One can construct a measure P satisfying ASSUMPTION 0, 1($\Gamma = A$) and 2($\rho = 1$), that puts probability 1 on the sequence of actions generated by “last-period-best-response” starting from (Heads, Heads): namely, (Heads, Heads), (Heads, Tails), (Tails, Tails), (Tails, Heads), (Heads, Heads), etc....

[INSERT Figure 5: Game 2]

Convergence is of little interest if no probability measure could possibly satisfy the assumptions from which it is shown to follow. Mere existence, however, is trivial. For any game G , let ψ_1, \dots, ψ_n be a mixed equilibrium in one of G 's minimal inclusive sets. The measure P that, at each t , (1) puts unit weight on the type profile θ such that (a) each player's belief about incipient opponent play is $\psi_1 \times \dots \times \psi_{i-1} \times \psi_{i+1} \times \dots \times \psi_n$ and (b) “(a)” is common knowledge, and (2) draws play according to $\psi_1 \times \dots \times \psi_n$, is consistent with ASSUMPTION 0, ASSUMPTION 1 applied to all subsets, and ASSUMPTION 2 for any length of recent history. The Appendix shows that non trivial measures exist for all games.

4. MINIMAL INCLUSIVE SETS IN SPECIAL CLASSES OF GAMES

This section concerns the size of minimal inclusive sets in classes of games for which other learning processes have been shown to converge (in a manner weaker than proven here). The main result relies on a general theorem whose proof is straightforward and so omitted. (See Sanchirico (1996b) for details.)

Let E be a rectangular subset of action space A . The *restriction* of stage game G to E , denoted G_E is the finite game with strategy sets E_i and payoffs $\pi_i|E$. Let R be a property defined on the set of all finite games. We say R is *restrictable to inclusive sets*, if for all games G with property R and all inclusive sets I in G , the restriction G_I also has the property. The property: “has no more than one (pure strategy) equilibrium,” for instance, is restrictable to inclusive sets, since all equilibria in the restriction of G to an inclusive set are equilibria in G as well. The property is not, however, restrictable to general subsets of profiles. The property “has no less than one (pure) equilibrium” is not restrictable even to inclusive sets.

THEOREM 4: *If property R implies the existence of a (pure strategy) Nash equilibrium and is restrictable to inclusive sets, then in all games with property R for which all pure equilibria are strict,¹⁵ all minimal inclusive sets are singletons consisting of strict equilibria.*

For lack of space the reader is referred to the papers cited for formal definitions of the classes in the following corollary. However, one new concept is required. Because the subset of a complete lattice may not be a complete lattice in its own right, general supermodular games are not restrictable to inclusive sets and so not subject to THEOREM 4. Let us say, then, that an

ordinal supermodular game (Milgrom and Shannon (1994)) is *restrictable*, if for all inclusive sets I , each I_i is a complete lattice in its own right. Contained in this subclass are all supermodular games whose strategy sets may be completely ordered, including all those with diminishing increasing differences, as defined and analyzed by Krishna (1991).

COROLLARY 1 [Singleton Minimal Inclusive Sets in Special Classes of Games]: *Let G be a finite game. All of G 's minimal inclusive sets are singletons consisting of strict equilibria, if any of the following hold:*

(1) *G is a restrictable ordinal supermodular game all of whose pure equilibria are strict,*

(2) *G is an ordinal potential game (Monderer and Shapley (1993b)) all of whose pure equilibria are strict,*

(3) *G is a game with identical interests (Monderer and Shapley (1993a)) all of whose pure equilibria are strict,*

(4) *G has the marginal bandwagon property (Kandori and Rob (1992)).*

Proof: By THEOREM 4, we need only show that each property (1)-(4) is restrictable to inclusive sets and implies the existence of a pure equilibrium (with a slight variation in case (4)). In all cases, restrictability follows directly from the definitions. For existence: (1) Existence of a pure equilibrium is given by Milgrom and Shannon (1991) Theorem 15. (2) Existence is given by Monderer and Shapley (1993b), COROLLARY 2.2. (3) Existence is noted by Monderer and Shapley (1993b), p. 9. (4) One can show that the definition of the marginal bandwagon property implies the existence of a *strict* equilibrium.

5. CONCLUSION

One direction for future research would be to extend the model to non-simultaneous stage games, including several plays of a given normal form. This would in turn allow convergence to patterns of equilibria in the simultaneous game, as in Sonsino (1994), but without the lock-in dynamic that operates there. It would also be broad enough to encompass convergence to equilibria that are not simple sequences of equilibria in the simultaneous game. Extending the model in this direction would also force confrontation with the observability problems identified in Fudenberg and Kreps (1988).

Other directions for future research include discovering the precise relationship between the “parameters” in the two assumptions (Γ , $\{x_n\}$, ρ and ε) and the speed of convergence, and analyzing the issue of which minimal inclusive sets are likely to be selected.

Columbia University

6. APPENDIX: RELAXING THE ASSUMPTION OF MYOPIA

This appendix establishes that player’s need not be completely myopic in this model of probabilistic learning. The main lemma states that players will still play within an inclusive set (determined with respect to myopic best response) when that set is salient, so long as their discount rate is sufficiently small (i.e., the players are “sufficiently impatient”). The reason that this is not immediately obvious is that, however small their positive discount factor, the future may be decisive in players’ current choice. And, even when discount factors are sufficiently small, future expectations may determine a player’s choice of action within the given inclusive

set. The point is that, with sufficiently impatience, future expectations will never “push” a player out of an inclusive set.

Let μ_{-i} be agent i 's prior on opponent strategies. Recall that μ_{-i} is a measure on

$$\begin{aligned} S_{-i} &= \left(\prod_{h \in H} A_1 \right) \times \dots \times \left(\prod_{h \in H} A_{i-1} \right) \times \left(\prod_{h \in H} A_{i+1} \right) \times \dots \times \left(\prod_{h \in H} A_n \right) \\ &= (A_1 \times \dots \times A_{i-1} \times A_{i+1} \times \dots \times A_n) \times \left(\prod_{h \in H \setminus \{\emptyset\}} A_1 \right) \times \dots \times \left(\prod_{h \in H \setminus \{\emptyset\}} A_{i-1} \right) \times \left(\prod_{h \in H \setminus \{\emptyset\}} A_{i+1} \right) \times \dots \times \left(\prod_{h \in H \setminus \{\emptyset\}} A_n \right). \end{aligned}$$

Let μ_{-i}^1 denote the marginal of μ_{-i} on first period opponent action profiles, i.e.,

$(A_1 \times \dots \times A_{i-1} \times A_{i+1} \times \dots \times A_n)$ in the large product above. For any strategy s_i for player i , let $s_i(\emptyset)$

denote the action prescribed for period 1.

LEMMA [For Definition of Sufficient Impatience]: *For all (finite) stage games G (whether or not generic), there exists some discount factor $\bar{\delta} > 0$ s.t. if $\delta \leq \bar{\delta}$, then for all players i and all subsets $E_{-i} \subseteq A_{-i}$ of stage game opponent action profiles, if $s_i \in B_i^\delta(\mu_{-i})$ and $\mu_{-i}^1 \in \Delta(E_{-i})$, then $s_i(\emptyset) \in b_i \circ \Delta(E_{-i})$.*

REMARK [There is No Positive Discount Rate Small Enough to Guarantee Myopic Best Response]: The lemma says that if a player is sufficiently impatient and she believes that her opponents will play within $E_{-i} \subseteq A_{-i}$ in period 1, then she will play a best response to some beliefs on E_{-i} in the first period. The lemma does *not* state that this player's first period action will be a stage game best response to the marginal μ_{-i}^1 of repeated game beliefs on his opponents' first period actions. Indeed, it is not generally true, even *generically*, that we can find

$\bar{\delta}$ small enough (and still positive) so that, using the notation of the lemma, $s_i(\emptyset) \in b_i(\mu_{-i}^1)$. As shown in the following example, however small is δ , we can usually find a repeated game belief μ_{-i} such that the best response repeated game strategy does not prescribe an action in period 1 which is best response to the marginal μ_{-i}^1 of μ_{-i} on first period play. However, since the results in this paper are based on the best response properties of *subsets*, the weaker result in the lemma is all that we shall need.

EXAMPLE [Non-optimal First Period Play with Arbitrarily Small Discount Factors]: Consider, for example, MATCHING PENNIES and any discount factor δ , however small. In the stage game,

	H	T
H	1, -1	-1, 1
T	-1, 1	1, -1

MATCHING PENNIES

the break even belief for both players is $(\frac{1}{2}H, \frac{1}{2}T)$. Let V_1 and V_2 , with $V_1 > V_2$, be two distinct payoffs in the *repeated* game $G^\infty(\delta)$, each of which are produced by the extensive form

rationalizable repeated game beliefs μ_{-i}^1 and μ_{-i}^2 . (Assume for the purpose of the example that two distinct payoffs of this type exist. Indeed, this will be the case.) We construct an additional extensive form rationalizable belief μ_{-i} for Row as follows. (Again postpone the question of whether the constructed belief is consistent with extensive form rationalizability.) First, we construct beliefs over the time 1 stage game. We specify that Row believes that Column plays H with probability $\frac{1}{2} + \varepsilon$, where $0 < 4\varepsilon < \delta(V_1 - V_2)$. Notice that the best response for Row *in the time 1 stage game* is strictly H , since, no matter how small δ , epsilon ε may be set to a positive number. Next we construct continuation beliefs for each of the four subgames following time 1 play. In the two subgames where Row has played H , assign Row the belief μ_{-i}^2 , as defined above. In the two subgames following Row's play of T , assign Row the belief μ_{-i}^1 . By the Caratheodory extension theorem of measure theory, these five measures, 1 for time 1 play, and four for each of the successive subgames, induce a unique belief (measure) over Column's repeated game strategies. What then is a best response *extensive form strategy* for this constructed belief? What, in particular, is first period play in any of these best responses?. If Row chooses H in the first period current payoffs are $(\frac{1}{2} + \varepsilon) \cdot 1 + (\frac{1}{2} - \varepsilon) \cdot (-1) = 2\varepsilon$ and discounted future payoffs are δV_2 . If Row chooses T in the first period, current payoffs are -2ε , while discounted future payoffs are δV_1 . By choice of ε , the 4ε gain in the first period from playing H instead of T is outweighed by the $\delta(V_1 - V_2)$ discounted loss in future payoffs.

In sum, however small the importance of the future, we can find current beliefs close enough to the break even point between two actions in the stage game, which make the current choice between these actions even less important.

Proof of Lemma:

Step 1:

Suppose that $a_i \notin b_i \circ \Delta(E_{-i})$. I claim that we can find $\varepsilon > 0$ such that $\forall \psi_{-i} \in \Delta(E_{-i})$, $v_i(\psi_{-i}) - E_{\psi_{-i}} \pi_i(a_i, \cdot) > \varepsilon$. Notice that $v_i(\psi_{-i}) - E_{\psi_{-i}} \pi_i(a_i, \cdot)$ is continuous in ψ_{-i} . Also, since $a_i \notin b_i \circ \Delta(E_{-i})$, we have that $v_i(\psi_{-i}) - E_{\psi_{-i}} \pi_i(a_i, \cdot) > 0$, $\forall \psi_{-i} \in \Delta(E_{-i})$. The claim then follows from the well-known result that a continuous and strictly positive function on a compact set is uniformly bounded away from zero.

Since G is finite in players and actions, we can take $\varepsilon > 0$ to be uniform across players i , subsets $E_{-i} \subseteq A_{-i}$ and $a_i \notin b_i \circ \Delta(E_{-i})$.

Step 2

Given δ , let \bar{V} and \underline{V} be the largest and smallest repeated game payoffs, respectively, chosen across all players. These exist and are precisely $\frac{\bar{v}}{1-\delta}$ and $\frac{\underline{v}}{1-\delta}$ respectively, where \bar{v} and \underline{v} are the largest and smallest payoffs across all players in the (finite) *stage game*, respectively. Since $\frac{\delta}{1-\delta}$ approaches 0 as δ approaches zero, we may then set $\delta > 0$ small enough so that

$$\delta(\bar{V} - \underline{V}) = \frac{\delta}{1-\delta}(\bar{v} - \underline{v}) < \varepsilon.$$

Notice that nothing in the above argument depended on \bar{V} and \underline{V} being distinct: in particular, the argument here does not rely on genericity. Moreover, notice that we need not

choose \bar{V} and \underline{V} from among the extensive form rationalizable payoffs; the spread between the largest and smallest values in this set will only be smaller.

Now suppose that for some belief μ_{-i} , we have $\mu_{-i}^1 \in \Delta(E_{-i})$. Take any $a_i \notin b_i \circ \Delta(E_{-i})$ and any $s_i \in S_i$ s.t. $s_i(\emptyset) = a_i$. First,

$$U(s_i, \mu_{-i}) \leq E_{\mu_{-i}^1} \pi(a_i, \cdot) + \delta \bar{V}$$

Second, for any $a_i' \in b_i \circ \Delta(E_{-i})$ and any $s_i' \in S_i$ s.t. $s_i'(\emptyset) = a_i'$,

$$U(s_i', \mu_{-i}) \geq v_i(\mu_{-i}^1) + \delta \underline{V}$$

Therefore, by choice of δ ,

$$\begin{aligned} U_i(s_i', \mu_{-i}) - U_i(s_i, \mu_{-i}) &\geq (v_i(\mu_{-i}^1) - E_{\mu_{-i}^1} \pi(a_i, \cdot)) - \delta(\bar{V} - \underline{V}) \\ &\geq \varepsilon - \delta(\bar{V} - \underline{V}) \\ &> 0. \end{aligned}$$

(“ ε “ was defined in step 1). Therefore, s_i is not best response to μ_{-i} . ■

DEFINITION [Sufficient Impatience]: Given repeated game $G^\infty(\delta)$, we say that players are *sufficiently impatient*, if δ is no greater than $\bar{\delta}$ as defined in the preceding LEMMA.

EXAMPLE [How Much Impatience Will Suffice?]: How small does $\bar{\delta}$ have to be? Naturally, the answer depends on payoffs. In particular, it depends on a) the spread between the largest and smallest extensive form rationalizable repeated game payoffs, (which will simultaneously depend on δ) and b) the size of the smallest “ ε -room” around a subset in the stage game (See, the proof of the lemma for an intuitive meaning of “ ε -room”).

	<i>P</i>	<i>G</i>
<i>P</i>	<i>1,1</i>	<i>0,0</i>
<i>G</i>	<i>0,0</i>	<i>1,1</i>

Game of Pure Coordination

Consider, for example, the following game of pure coordination. The largest repeated extensive form rationalizable payoff is $\sum_{t=1}^{\infty} \delta^{t-1} \cdot 1 = \frac{1}{1-\delta}$, corresponding to the extensive form rationalizable strategy of always showing up at Penn and believing one’s opponent will do the same. The lowest extensive form rationalizable payoff is no less than (indeed, is equal to)

$\sum_{t=1}^{\infty} \delta^{t-1} \cdot \frac{1}{2} = \frac{\frac{1}{2}}{1-\delta}$. Since a player can always play myopic best response to any belief, and the

minimum value of $v(\psi_{-i})$ is $\frac{1}{2}$, no extensive form rationalizable strategy can pay less than this amount. (One can show that this is in fact the lowest value among extensive form rationalizable strategies and beliefs.) The discounted difference is thus, for any given δ , no more than

$$\left(\frac{\delta}{1-\delta} \right)^{\frac{1}{2}}.$$

What about the “ ε -room” around subsets in the stage game? There are only three opponent subsets for each player, $\{P\}$, $\{G\}$ and $\{P, G\}$. The last need not be checked. The first two are the same. In particular, if Row believes that Column is playing P, and she herself plays G instead of P, the loss is 1.

The discount factor δ must be set so that differences in future payoffs never outweigh this potential current loss, in other words, so that $\left(\frac{\delta}{1-\delta} \right)^{\frac{1}{2}} < 1$. Hence, any δ less than $\frac{2}{3}$ is sufficient impatience in this game. For the purpose of example, this corresponds to an interest rate of 10% compounded over 4-5 periods.

7. APPENDIX: PROOF OF LEMMA 1

Take any inclusive set I . I give the proof in two steps.

Step 1: First I show that $\theta \in K \cap S(I)$ implies $\theta_i(A_{-i}) \in \Delta(I_{-i})$ for all i . For this it suffices to show that for all n and all i , $\theta_i \in K_i \cap S_i(n)(I)$ implies $\theta_i(A_{-i}) \in \Delta(I_{-i})$. The proposition is true for all i and $n = 1$ by definition of $S_i(1)(I)$. Continuing inductively, suppose that, for all j ,

$\theta_j \in K_j \cap S_j(n-1)(I)$ implies $\theta_j(A_{-j}) \in \Delta(I_{-j})$. Take any i and any $\theta_i \in K_i \cap S_i(n)(I)$. I must show $\theta_i(A_{-i}) \in \Delta(I_{-i})$. To this end take any $j \neq i$ and any $a_j \in \text{supp } \theta_i(A_j)$. I claim $a_j \in I_j$.

By definition of $S_i(n)(I)$ either *i*) $a_j \in I_j$, in which case we are done, or *ii*)

$\theta_i(\langle \{a_j\} \rangle \cap \langle S_j(n-1)(I) \rangle) > 0$. (Recall that θ_i may be regarded as a measure on $\Theta_{-i} \times A_{-i}$ and that the notation " $\langle \rangle$ " denotes the inverse image of the projection mapping.) Consider case *ii*).

First, since $\theta_i \in K_i$, we know that $\theta_i(\langle K_j(m-1) \rangle) = 1$, for all m . Therefore, $\theta_i(\langle K_j \rangle) = 1$ and, in

turn, $\theta_i(\langle \{a_j\} \rangle \cap \langle K_j \cap S_j(n-1)(I) \rangle) > 0$. Then by the inductive hypothesis,

$\theta_i(\langle \{a_j\} \rangle \cap \langle \{\theta_j \in \Theta_j \mid \theta_j(A_{-j}) \in \Delta(I_{-j})\} \rangle) > 0$. Second, $\theta_i \in K_i \subseteq K_i(1)$ also means that

$\theta_i(\langle \langle \{(\theta_j, a_j) \in \Theta_j \times A_j \mid a_j \in b_j(\theta_j(A_{-j}))\} \rangle \rangle) = 1$, where the inverse projection here is with respect

to the factor Θ_j in $\Theta_{-i} \times A_{-i}$. Combining these two implications yields $a_j \in b_j \circ \Delta(I_{-j})$. Then

since I is inclusive, $a_j \in I_j$.

Step 2: For all $t \geq 1$,

$$\begin{aligned}
& P([\theta' \in S(I)]) \\
&= P([\theta' \in K \cap S(I)]) && \text{[ASSUMPTION 0]} \\
&\leq P\left(\bigcap_{i=1}^n [\theta'_i(A_{-i}) \in \Delta(I_{-i})] \cap [\theta' \in S(I)]\right) && \text{[Step 1]} \\
&= P\left(\bigcap_{i=1}^n [a'_i \in b_i(\theta'_i(A_{-i}))] \cap \bigcap_{i=1}^n [\theta'_i(A_{-i}) \in \Delta(I_{-i})] \cap [\theta' \in S(I)]\right) && \text{[ASSUMPTION 0]}
\end{aligned}$$

$$\leq P\left([a' \in b \circ \Delta(I)] \cap [\theta' \in S(I)]\right) \quad [\text{Definition of } \Delta(E_{-i})]$$

$$\leq P\left([a' \in I_i] \cap [\theta' \in S(I)]\right). \quad [I \text{ is inclusive}]$$

Hence, $P\left([a' \in I_i] \cap [\theta' \in S(I)]\right) = P([\theta' \in S(I)])$ or $P\left([a' \in I_i] \mid [\theta' \in S(I)]\right) = 1$, if defined. ■

8. APPENDIX: LEMMA FOR THEOREM 2 (EVENTUAL REPULSION)

REMARK 1: The following lemma is stronger than what is needed for the proof of THEOREM 2.

ASSUMPTION 2 could be weakened accordingly.

LEMMA 2: *If $\{A_n\}$ and $\{B_n\}$ are sequences of events in some probability space $(\Omega, \mathfrak{F}, P)$*

and for infinitely many $m \geq 1$, $\exists \varepsilon_m > 0$ such that

$$\forall n \geq m \quad P\left(B_n \mid A_n - \bigcup_{k=m}^{n-1} [A_k \cap B_k]\right) \geq \varepsilon_m, \text{ if defined,} \quad (10)$$

then

$$P\left([A_n \cap B_n \text{ i.o.}] \mid [A_n \text{ i.o.}]\right) = 1, \text{ if defined.}$$

Proof: The result is trivial if $P([A_n \text{ i.o.}]) = 0$. Assume hereafter that $P([A_n \text{ i.o.}]) > 0$ and

redefine P to be its conditional on $[A_n \text{ i.o.}]$. Then from this redefinition of P and the definition

of the event $[A_n \cap B_n \text{ i.o.}]$ it suffices to show that for infinitely many $m \geq 1$, $P\left(\bigcup_{n=m}^{\infty} A_n \cap B_n\right) = 1$.

I show this for each of the infinitely many m which satisfy (10) in the statement of the lemma.

Take any $m \geq 1$ satisfying (10). Since $P([A_n \text{ i.o.}]) = 1$, we have: $\forall r \geq 1$,

$\lim_{M \rightarrow \infty} P\left(\bigcup_{n=r}^M A_n\right) = P\left(\bigcup_{n=r}^{\infty} A_n\right) = 1$. We can therefore find a subsequence $\{m_k\}$ of indices such that

$$m_1 = 1 \text{ and } \forall k \geq 1, P\left(\bigcup_{n=m_k}^{m_{k+1}-1} A_n\right) \geq 1 - \frac{1}{k}.$$

Now, for notational convenience, we may write, $C_0 = \emptyset$ and $\forall k \geq 1, C_k = \bigcup_{n=m}^{m_{k+1}-1} [A_n \cap B_n]$. In

this notation, our object is to show $\lim_{k \rightarrow \infty} P(C_k) = 1$. Now $\forall k \geq 1$, the union C_k may be written as

a disjoint union in the usual manner: $C_{k-1} \cup \left(\bigcup_{n=m_k}^{m_{k+1}-1} [A_n \cap B_n - C_{k-1}]\right)$. Then

$$P(C_k) - P(C_{k-1}) = P\left(\bigcup_{n=m_k}^{m_{k+1}-1} [A_n \cap B_n - C_{k-1}]\right).$$

Similarly “disjointizing” the union on the right hand side yields

$$P\left(\bigcup_{n=m_k}^{m_{k+1}-1} \left[A_n \cap B_n - \left[\bigcup_{j=m_k}^{n-1} A_j \cap B_j\right] - C_{k-1}\right]\right),$$

which by the additivity of P is the same as

$$\sum_{n=m_k}^{m_{k+1}-1} P\left(A_n \cap B_n - \left[\bigcup_{j=m_k}^{n-1} A_j \cap B_j\right] - C_{k-1}\right) = \sum_{n=m_k}^{m_{k+1}-1} P\left(A_n \cap B_n - \left[\bigcup_{j=m}^{n-1} A_j \cap B_j\right]\right).$$

By (10) in the statement of the lemma this is at least

$$\varepsilon_m \sum_{n=m_k}^{m_{k+1}-1} P\left(A_n - \left[\bigcup_{j=m}^{n-1} A_j \cap B_j\right]\right) = \varepsilon_m \sum_{n=m_k}^{m_{k+1}-1} P\left(A_n - \left[\bigcup_{j=m_k}^{n-1} A_j \cap B_j\right] - C_{k-1}\right).$$

(Hereafter, I drop the subscript m on ε_m). Now $A_n - \left[\bigcup_{j=m_k}^{n-1} A_j \right] - C_{k-1} \subseteq A_n - \left[\bigcup_{j=m_k}^{n-1} A_j \cap B_j \right] - C_{k-1}$

and so by the monotonicity of P , the right hand side of the foregoing equation is at least

$$\begin{aligned} & \varepsilon \sum_{n=m_k}^{m_{k+1}-1} P \left(A_n - \left[\bigcup_{j=m_k}^{n-1} A_j \right] - C_{k-1} \right) \\ & \geq \varepsilon \cdot P \left(\bigcup_{n=m_k}^{m_{k+1}-1} \left[A_n - \left[\bigcup_{j=m_k}^{n-1} A_j \right] - C_{k-1} \right] \right) \quad [\text{Subadditivity of } P] \\ & = \varepsilon \cdot P \left(\left[\bigcup_{n=m_k}^{m_{k+1}-1} A_n \right] - C_{k-1} \right). \end{aligned}$$

Some manipulation shows that this is at least

$$\varepsilon \left[P(-C_{k-1}) - \left(1 - P \left(\bigcup_{n=m_k}^{m_{k+1}-1} A_n \right) \right) \right].$$

And by choice of the sequence $\{m_k\}$, this is no less than

$$\varepsilon \left[P(-C_{k-1}) - \left(1 - \left(1 - \frac{1}{k} \right) \right) \right] = \varepsilon \left(P(-C_{k-1}) - \frac{1}{k} \right) = -\varepsilon \cdot P(C_{k-1}) + \varepsilon \left(1 - \frac{1}{k} \right).$$

Hence, the upshot of the foregoing chain of equalities and inequalities is the “difference inequation”

$$P(C_k) \geq (1 - \varepsilon) P(C_{k-1}) + \varepsilon \left(1 - \frac{1}{k} \right),$$

which by iterated substitution yields

$$P(C_k) \geq (1 - \varepsilon)^k P(C_1) + \varepsilon \left(1 - \frac{1}{k} \right) \sum_{n=0}^{k-1} (1 - \varepsilon)^n.$$

Taking the limit of both sides yields

$$\lim_{k \rightarrow \infty} P(C_k) = 1,$$

as was to be shown.

9. Appendix: Proof of Theorem 3, Convergence, CASE (I)

We will need a way to count through the nesting structure of inclusive sets. To this end first define the following relation on the family of inclusive sets for G : write $I < I'$, if and only if either I is strictly contained in I' or, $I = I'$ and both are minimal. Second, let $C^0 = \{A\}$. Third, for all $n \geq 1$, let C^n be the family of all inclusive sets I for which there exists a chain, with $<$, of length n from I to A . More specifically, an inclusive set belongs to C^n , if and only if there exist $n - 1$ inclusive sets, I^2, \dots, I^n such that $I < I^2 < I^3 < \dots < I^n < A$.

There are several things to note about this construction. First, if I is minimal, then by the definition of $<$, I may equal I^2 or both I^2 and I^3 , etc., in such a chain. Second, the definition of C^n also implies that if $I \in C^n$, then $I \in C^k$, for all $k \leq n$. Third, these two remarks imply that minimal inclusive sets are in every C^n . Fourth, since every game G has at least one minimal inclusive set, the C^n are never empty. And last, since there are finitely many inclusive sets, for large enough n , all $I \in C^n$ are minimal.

Therefore, it suffices to prove that $\forall n \geq 1$, $P\left(\bigcup_{I \in C^n} [a' \in I \text{ a.a.}]\right) = 1$. I do so by induction.

The case of $n = 0$ is trivial since, of course, $P(a' \in A \text{ a.a.}) = 1$.

Now suppose that $P\left(\bigcup_{I \in C^n} [a' \in I \text{ a.a.}]\right) = 1$. I will show $P\left(\bigcup_{I \in C^{n+1}} [a' \in I \text{ a.a.}]\right) = 1$ in two steps.

The first step uses THEOREM 2 (Eventual Repulsion) and the inductive hypothesis to prove

$P\left(\bigcup_{I \in C^{n+1}} [a' \in I \text{ i.o.}]\right) = 1$. The second step combines this result with THEOREM 1 (Eventual

Absorption) to show $P\left(\bigcup_{I \in C^{n+1}} [a' \in I \text{ a.a.}]\right) = 1$, as desired.

I first simplify the notation. Given any set of action profiles E write $[E \text{ a.a.}]$ and $[E \text{ i.o.}]$ for $[a' \in E \text{ a.a.}]$ and $[a' \in E \text{ i.o.}]$, respectively. Moreover, for each inclusive set I , write $\mathfrak{I}(I)$ for the family of inclusive sets strictly contained in I , or if there are none, then the family consisting of I alone. Lastly, write $\bigcup \mathfrak{I}(I)$ for the union of this family $\mathfrak{I}(I)$.

Step 1: From THEOREM 2 (Eventual Repulsion) we know that

$$P([I \text{ a.a.}] - [\bigcup \mathfrak{I}(I) \text{ i.o.}]) = P([(I - \bigcup \mathfrak{I}(I)) \text{ a.a.}]) = 0,$$

for all inclusive sets I . This allows the first crucial link (11) in the following chain of equalities.

The other crucial link is the last and is given by the inductive hypothesis:

$$\begin{aligned} & P\left(\bigcup_{I \in C^n} \bigcup_{J \in \mathfrak{I}(I)} [J \text{ i.o.}]\right) \\ &= P\left(\bigcup_{I \in C^n} \left[\left(\bigcup_{J \in \mathfrak{I}(I)} J\right) \text{ i.o.}\right]\right) \quad [\text{G is finite}] \\ &= P\left(\bigcup_{I \in C^n} [\bigcup \mathfrak{I}(I) \text{ i.o.}]\right) \end{aligned}$$

$$\begin{aligned}
&\geq P\left(\bigcup_{I \in C^n} [[I \text{ a.a.}] \cap [\bigcup \mathfrak{Z}(I) \text{ i.o.}]]\right) && \text{[Monotonicity of } P\text{]} \\
&= P\left(\bigcup_{I \in C^n} [[I \text{ a.a.}] \cap [\bigcup \mathfrak{Z}(I) \text{ i.o.}]]\right) + P\left(\left(\bigcup_{I \in C^n} [[I \text{ a.a.}] - [\bigcup \mathfrak{Z}(I) \text{ i.o.}]]\right) - \left(\bigcup_{I \in C^n} [[I \text{ a.a.}] \cap [\bigcup \mathfrak{Z}(I) \text{ i.o.}]]\right)\right) \\
& && (11) \\
&= P\left(\left(\bigcup_{I \in C^n} [[I \text{ a.a.}] \cap [\bigcup \mathfrak{Z}(I) \text{ i.o.}]]\right) \cup \left(\bigcup_{I \in C^n} [[I \text{ a.a.}] - [\bigcup \mathfrak{Z}(I) \text{ i.o.}]]\right)\right) && [P(A) + P(B - A) = P(A \cup B)] \\
&= P\left(\bigcup_{I \in C^n} \left[\left[[I \text{ a.a.}] - [\bigcup \mathfrak{Z}(I) \text{ i.o.}] \right] \cup \left[[I \text{ a.a.}] \cap [\bigcup \mathfrak{Z}(I) \text{ i.o.}] \right] \right] \right) && \text{[Associativity of Union]} \\
&= P\left(\bigcup_{I \in C^n} [I \text{ a.a.}]\right) \\
&= 1. && \text{[Inductive Supposition]}
\end{aligned}$$

Now if I' is in $\mathfrak{Z}(I)$ for some I in C^n , then I' must be in C^{n+1} , for if there is a chain of length n from I to A and I' is strictly contained in I (or is identical to I if I is minimal), then there is a chain of length $n+1$ from I' to A . Combining this fact with the chain of equalities above gives:

$$\begin{aligned}
&P\left(\bigcup_{I \in C^{n+1}} [I \text{ i.o.}]\right) \\
&\geq P\left(\bigcup_{I \in C^n} \bigcup_{J \in \mathfrak{Z}(I)} [J \text{ i.o.}]\right) \\
&= 1. && (12)
\end{aligned}$$

This completes step 1.

Step 2: The second step is similar to step 1, with (12) playing the role of the inductive hypothesis and THEOREM 1 (Eventual Absorption) playing the role of THEOREM 2 (Eventual Repulsion).

$$\begin{aligned}
& P\left(\bigcup_{I \in C^{n+1}} [I \text{ a. a.}]\right) \\
& \geq P\left(\bigcup_{I \in C^{n+1}} [[I \text{ i. o.}] \cap [I \text{ a. a.}]]\right) && \text{[Monotonicity of } P\text{]} \\
& = P\left(\bigcup_{I \in C^{n+1}} [[I \text{ i. o.}] \cap [I \text{ a. a.}]]\right) + P\left(\left(\bigcup_{I \in C^{n+1}} [[I \text{ i. o.}] - [I \text{ a. a.}]]\right) - \left(\bigcup_{I \in C^{n+1}} [[I \text{ i. o.}] \cap [I \text{ a. a.}]]\right)\right) && \text{[THEOREM 1]} \\
& = P\left(\left(\bigcup_{I \in C^{n+1}} [[I \text{ i. o.}] - [I \text{ a. a.}]]\right) \cup \left(\bigcup_{I \in C^{n+1}} [[I \text{ i. o.}] \cap [I \text{ a. a.}]]\right)\right) && \text{[} P(A) + P(B - A) = P(A \cup B)\text{]} \\
& = P\left(\bigcup_{I \in C^{n+1}} \left[\left([I \text{ i. o.}] - [I \text{ a. a.}]\right) \cup [[I \text{ i. o.}] \cap [I \text{ a. a.}]]\right]\right) && \text{[Associativity of Union]} \\
& = P\left(\bigcup_{I \in C^{n+1}} [I \text{ i. o.}]\right) \\
& = 1. && \text{[Step 1]}
\end{aligned}$$

10. APPENDIX: EXISTENCE OF NON-TRIVIAL MEASURES

I construct a “non-trivial” measure in 6 steps. (Compare this construction to the example in the introduction.) First, I construct a probability measure over incipient beliefs for each partial history of beliefs and play whose last element is an action profile. Next, I construct a probability measure over incipient play for each partial history of beliefs and play ending in beliefs. Third, I apply the Caratheodory extension theorem to extend this collection of measures to a single measure on the set of full sequences of play and beliefs $[\Theta \times A]^\infty$. Fourth, fifth, and sixth, I show that the resulting measure satisfies ASSUMPTIONS 0, 1 (with the power set of A), and 2 (with any length of recent history), respectively.

Step 1: Take any $t \geq 2$ and any partial history of play and beliefs, $\{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}\}$ which ends in an action profile and whose action profiles are in the (nonempty) set of rationalizable profiles for the stage game. Taken as subsets of A , the corresponding sequence of histories of play back from t , $\{(a^s, \dots, a^{t-1})\}_{s=1}^{t-1}$ is clearly decreasing in index¹⁶ s . Then, since S is monotonic, the sequence $\{S(a^s, \dots, a^{t-1})\}_{s=1}^{t-1}$ of (measurable) subsets of the set of stage game type profiles, Θ , is also decreasing in s . For each $s = 2, \dots, t-1$ assign

$\rho(S(a^s, \dots, a^{t-1}) \cap H) = 1 - (\frac{1}{2})^{t-r}$, where $1 \leq r \leq s$ is the smallest number (thus indexing the longest history) with¹⁷ $(a^r, \dots, a^{t-1}) \subseteq (a^s, \dots, a^{t-1})$. There is possibly some duplication in each sequence $\{a^s, \dots, a^{t-1}\}$ and use of r insures that each subset is assigned a single number. Since all a^q , $q = 1, \dots, t-1$ are stage game rationalizable, $S(a^s, \dots, a^{t-1}) \cap H$ is always nonempty. Set-or reset, if $(a^1, \dots, a^{t-1}) = (a^2, \dots, a^{t-1})$ -- $\rho(S(a^1, \dots, a^{t-1}) \cap H) = 1$. The upshot of this assignment is:

$$\forall s = 1, \dots, t-1 \quad \rho(S(a^s, \dots, a^{t-1}) \cap H) \geq 1 - (\frac{1}{2})^{t-s} \quad (13)$$

Extend ρ to the partition of Θ whose s^{th} cell is

$$C^s = S(a^{s-1}, \dots, a^{t-1}) \cap H - [S(a^s, \dots, a^{t-1}) \cap H], \text{ by setting}$$

$$\rho(C^s) = \rho(S(a^{s-1}, \dots, a^{t-1}) \cap H) - \rho(S(a^s, \dots, a^{t-1}) \cap H). \text{ Note that this assignment is consistent}$$

with the additivity property of a probability measure. Continuing, for each cell C^s consider the

partition $\{b^{-1}(a) \cap C^s \mid a \in A\}$ where I have written $b^{-1}(a)$ for $\{\theta \in \Theta \mid \forall i, a_i \in b_i(\theta_i(A_{-i}))\}$. If

there are m distinct nonempty cells in this partition, assign each such cell the number

$\rho(b^{-1}(a) \cap C^s) = \frac{1}{m} \rho(C^s)$. In words, “distribute” the probability of C^s equally among its nonempty cells.

I have to this point specified a map taking each (measurable) cell in the partition

$\{b^{-1}(a) \cap C^s \mid s = 2, \dots, t-1, a \in A\}$ of Θ onto a real number. Moreover, this set valued map sums to one over its domain. I may then extend the map to a probability measure $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}$ on all measurable subsets of Θ and this measure will also have the property (13).

If I then choose a measure on Θ arbitrarily for $t = 1$, I have a measure P_h over Θ for the null history as well as each partial history $h = \{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}\}$ which ends in stage game rationalizable play.

Step 2: Now for all histories of play and beliefs $(\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}, \theta^t)$ which end in beliefs, define a probability measure $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}, \theta^t}$ on A (rather than Θ) by assigning equal probability to each a with $a_i \in b_i \circ \Delta(\theta^t(A_{-i}))$, all i . Note that each a receiving positive probability will be stage game rationalizable.

Step 3: To this point I have defined, for each partial history of (stage game rationalizable) play and beliefs, whether it end in play or beliefs, a probability measure over the incipient factor, Θ or $R(A)$, in the product $[\Theta \times R(A)]^\infty$, where $R(A) \subseteq A$ stands for the set of stage game rationalizable strategies. By an application of the Caratheodory extension theorem (see Ash p. 109), I may (uniquely) extend this family of measures to a measure P on $[\Theta \times A]^\infty$ which has the

property, *inter alia*, that for all finite dimensional measurable rectangles,

$[T_1 \times S_1 \times \dots \times T_n \times S_n] \times T_{n+1} \times [A_{n+1} \times \Theta_{n+2} \times A_{n+2} \dots]$ with each $T_k \subseteq \Theta$ and each $S_k \subseteq A$,

$$P(T_1 \times S_1 \times \dots \times T_n \times S_n \times T_{n+1} \times A_{n+1} \times \Theta_{n+2} \times \dots) = \int_{T_1 \times S_1 \times \dots \times T_n \times S_n \times \Theta_{n+1} \times A_{n+1} \times \Theta_{n+2} \times \dots} P_{\theta^1, a^1, \dots, \theta^n, a^n}(T_{n+1}) dP, \quad (14)$$

and similarly for rectangles whose last restricted factor is A .

Step 4: Since by construction $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}, \theta^t} \left(\prod_{i=1}^n b(\theta_i^t(A_{-i})) \right) = 1$ and $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}(H) = 1$,

it is clear that P satisfies ASSUMPTION 0.

Step 5: To prove that P satisfies ASSUMPTION 1 it suffices to show that $\left\{ \left(\frac{1}{2} \right)^n \right\}$ is a bounding sequence (as in that assumption) for all $E \subseteq A$. Take any $t \geq 2$, $n \geq 1$ and any $E \subseteq A$. For notational convenience set $F = \left\{ \{a^t\} \in A^\infty \mid a^{t-n-1} \notin A, a^{t-n} \in A, \dots, a^{t-1} \in A \right\}$. Then by (14),

$$P(\theta^t \in S(A), a^{t-n-1} \notin A, a^{t-n} \in A, \dots, a^{t-1} \in A) = \int_{F \times \Theta^\infty} P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}(S(A)) dP.$$

But if a partial history $(\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})$ is in F , the actions $(a^{t-n}, \dots, a^{t-1})$ therein must be contained in A and so $S(a^{t-n}, \dots, a^{t-1}) \subseteq S(A)$. Therefore

$$\int_{F \times \Theta^\infty} P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}(S(A)) dP \geq \int_{F \times \Theta^{st}} P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}(S(a^{t-n}, \dots, a^{t-1})) dP.$$

Further, by construction of $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}$ (setting $s = t - n$),

$$\int_{F \times \Theta^\infty} P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}(S(a^{t-n}, \dots, a^{t-1})) dP \geq \int_{F \times \Theta^\infty} \left(1 - \left(\frac{1}{2} \right)^n \right) dP = \left(1 - \left(\frac{1}{2} \right)^n \right) P(F).$$

Therefore, substituting for F and combining the last three expressions,

$$P(\theta' \in S(A), a^{t-n-1} \notin A, a^{t-n} \in A, \dots, a^{t-1} \in A) \geq \left(1 - \left(\frac{1}{2}\right)^n\right) P(a^{t-n-1} \notin A, a^{t-n} \in A, \dots, a^{t-1} \in A),$$

as was to be proved.

Step 6: Let m be the length of recent history. To show that P satisfies ASSUMPTION 2, I will show the stronger result that if a is best response (for each player) to some beliefs on the last m periods of play in the history $(\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})$, then $P(a^t = a | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})) \geq \frac{1}{2m^2}$. For

notation convenience, set $D = (a^{t-m}, \dots, a^{t-1})$. By construction of $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}}$,

$$P(\theta' \in S(D) | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})) \geq 1 - \left(\frac{1}{2}\right)^m \geq \frac{1}{2}. \text{ Now since } a_i \in b_i \circ \Delta(D_{-i}), \text{ all } i, \text{ there exists}$$

$\theta \in S(D)$ (indeed there exists $\theta \in S(1)(D)$) such that $a_i \in b_i(\theta_i(A_{-i}))$, all i . Hence, the cell

$b^{-1}(a) \cap S(D)$ is nonempty and so receives at least $\frac{1}{m} P(\theta' \in S(D) | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})) \geq \frac{1}{2m}$

weight in $P(\cdot | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}))$. Now the construction of $P_{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}, \theta'}$ insures that for each

$\theta \in b^{-1}(a) \cap S(D)$, $P(a^t = a | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}, \theta')) \geq \frac{1}{m}$. Hence,

$$\begin{aligned} & P(a^t = a | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})) \\ & \geq \int_{b^{-1}(a) \cap S(D)} P(a^t = a | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}, \underline{\theta}')) dP \\ & \geq \int_{b^{-1}(a) \cap S(D)} \frac{1}{m} dP \\ & \geq \frac{1}{m} P(\theta' \in b^{-1}(a) \cap S(D) | (\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1})) \\ & \geq \frac{1}{m} \frac{1}{2m}, \end{aligned}$$

as was to be proved.

11. REFERENCES

Ash, R. (1972), *Real Analysis and Probability*, New York, Academic Press.

Basu, K. and J. W. Weibull (1991), "Strategy Subsets Closed Under Rational Behavior,"
Economics Letters, 36:141-146.

Bernheim, B. D. (1984), "Rationalizable Strategic Behavior," *Econometrica*, 52:1007-1028.

Fudenberg, D. and D. Kreps, (1988), "A Theory of Learning, Experimentation and Equilibrium
in Games," mimeo, M.I.T. and Stanford Graduate School of Business.

----- (1993), "Learning Mixed Equilibria," *Games and Economic Behavior*, 5:320-367.

Gul, F. (1991), "Rationality and Coherent Theories of Strategic Behavior," mimeo, Stanford
Graduate School of Business.

Hurkens, S. (1994), "Learning by Forgetful Players: From Primitive Formations to Persistent
Retracts," CentER for Economic Research, Tilburg University, The Netherlands, Discussion
Paper No. 9437.

Kalai, E. and E. Lehrer (1993), "Rational Learning Leads to Nash Equilibrium," *Econometrica*,
61: 1019-1045.

Kandori, M and R. Rob (1992), "Bandwagon Effects and Long Run Technological Choice,"
mimeo University of Tokyo and University of Pennsylvania.

Knopp, K. (1971), *Theory and Application of Infinite Series*, New York, Hafner Publishing.

Krishna, V. (1991), "Learning in Games with Strategic Complementarities," mimeo Harvard
Business School.

- Milgrom, P. and J. Roberts (1991), "Adaptive and Sophisticated Learning in Normal Form Games," *Games and Economic Behavior*, 3: 82-100
- Milgrom, P. and C. Shannon (1994), "Monotone Comparative Statics," *Econometrica* 62: 157-180.
- Monderer, D. and Shapley, L. S.(1993a), "Fictitious Play Property for Games with Identical Interests," mimeo, The Technion, Haifa and the University of California, Los Angeles.
- (1993b), "Potential Games," mimeo, The Technion, Haifa and the University of California, Los Angeles.
- Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52:1029-1050
- Roth, D. and E. Stacchetti (1993), "A Folk Theorem for Rationalizability without the Grain of Truth," mimeo, University of Michigan.
- Sanchirico, C. W. (1996), "A Probabilistic Model of Learning in Games," mimeo, Columbia University, Economics Department, *Forthcoming in Econometrica*. (A shorter version of this paper.)
- (1996b), "Minimal Inclusive Sets in Special Classes of Games," Columbia University, Department of Economics, Discussion Paper No. 9596-23.
- Sonsino, D. (1994), "Learning to Learn, Pattern Recognition, and Nash Equilibrium," mimeo, Stanford Graduate School of Business.
- Shapley, L. S. (1964), "Some Topics in Two Person Games," *Advances in Game Theory*, 26:1-28

Tan, T. and S. Werlang, (1988), "The Bayesian Foundations of Solution Concepts of Games,"

Journal of Economic Theory, 45: 370-391.

12. FIGURES

	Heads	Tails	Out
Heads	1, -1	-1, 1	1, 2
Tails	-1, 1	1, -1	-1, -4

Figure 1: Game 1

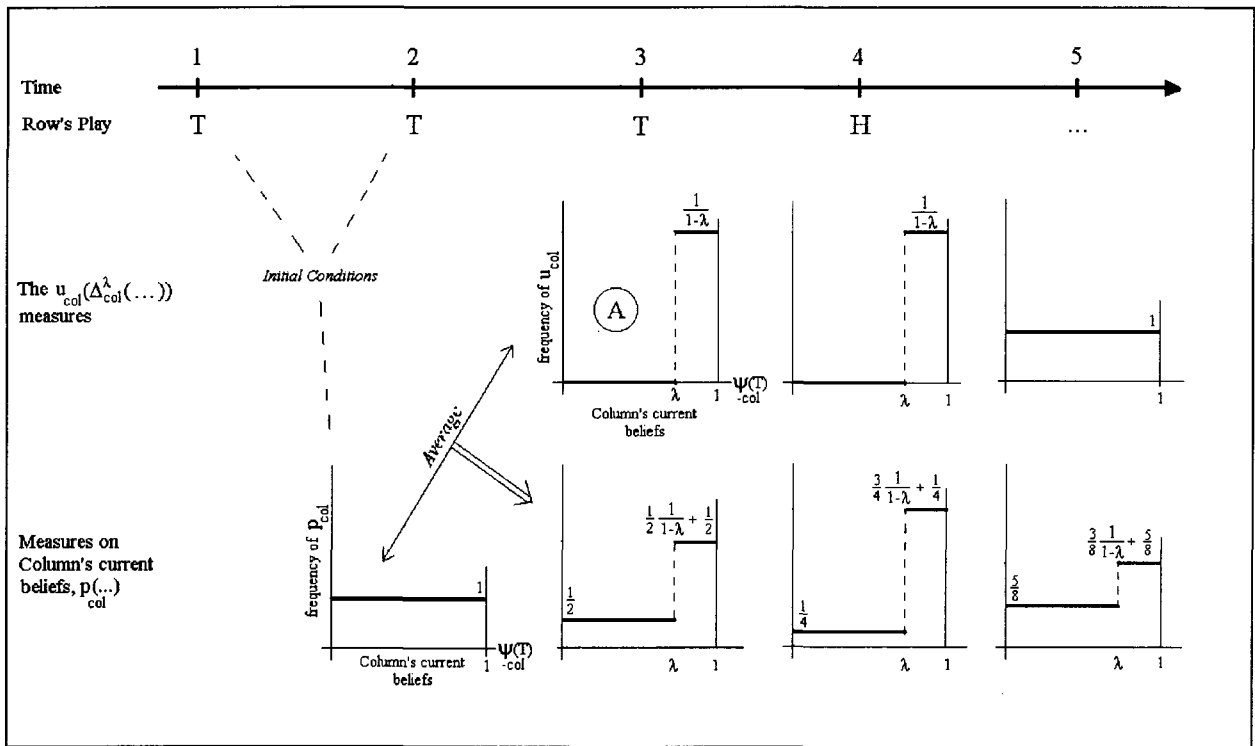


Figure 2

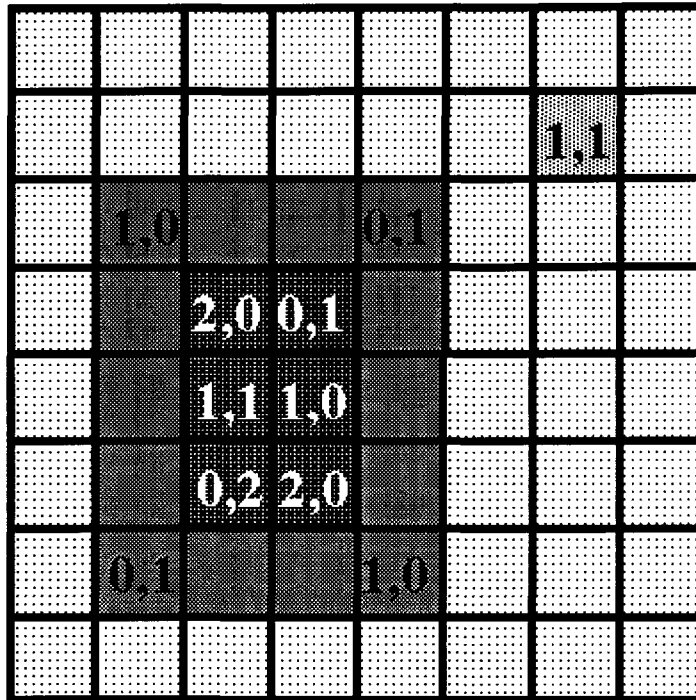


Figure 3

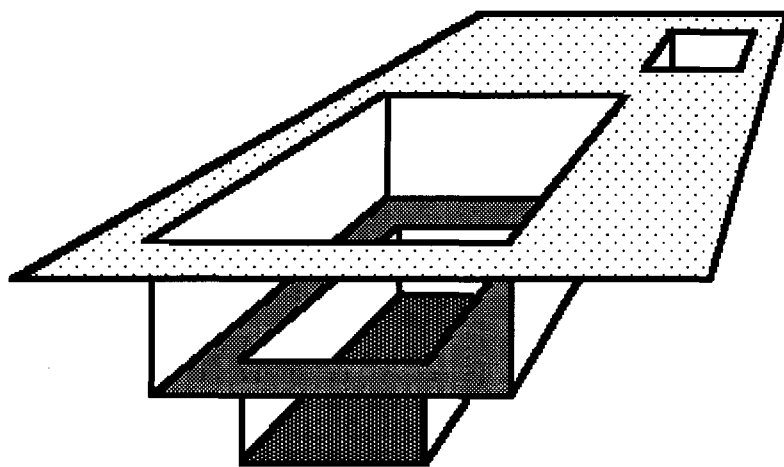


Figure 4

	Heads	Tails	Out
Heads	1, -1	-1, 1	$-\frac{1}{2}, \frac{1}{2}$
Tails	-1, 1	1, -1	$-\frac{1}{2}, \frac{1}{2}$
Out	$\frac{1}{2}, -\frac{1}{2}$	$\frac{1}{2}, -\frac{1}{2}$	0,0

Figure 5: Game 2

¹ I wish to thank Truman Bewley, Margaret Bray, Joseph Chang, John Geanakoplos, Roger Myerson, Phil Reny, Rafael Rob, Susan Rose-Ackerman, the members of Yale's game theory reading group, and seminar participants at the Econometric Society's 1994 North American summer meetings, SITE's 1994 summer workshop, U.C. Berkeley Economics, Yale Law School, Yale Economics and the Universidad Carlos III de Madrid's 1995 summer conference in economic theory. I especially wish to thank David Pearce for his advice and encouragement.

² GAME 1 is essentially the same as an example in Fudenberg and Kreps (1993).

³ All symbols used here are defined formally in the first paragraph of SECTION 2.

⁴ Where there will be no confusion, I use sequential notation for both the sequence and its range.

⁵ The parameter r should not be confused with the parameter ρ in ASSUMPTION 2.

⁶ All statements made about this system of measures applies to any system which approximates it in a particular sense: the approximating measures always assign the same weight as (1) to each cell in a canonical simplicial subdivision of the simplex (of i 's beliefs), which subdivision's cells have sides with length no greater than $1 - \lambda$.

⁷ Define tie-breaking rules $q_i(\psi_i)$: for each ψ_i , a measure on A_i supported on the set of best responses to ψ_i . Then apply Ash (1972, Theorem 2.7.2) to the product

$$\Delta(A_{Col}) \times \Delta(A_{Row}) \times A_{Col} \times A_{Row} \times \Delta(A_{Col}) \times \Delta(A_{Row}) \times A_{Col} \times A_{Row} \times \dots$$

⁸ To "rationalize" $\{a'_{Row}\}$, for example, construct the history-independent strategy s_{Row} prescribing a'_{Row} at all time- t information sets/stage games. Since all Row's actions are stage-game rationalizable, a'_{Row} is stage game best response to some stage game beliefs ψ'_{-Row} on

Column's rationalizable actions. By the classic product measure theorem (See, e.g., Ash 1972, Corollary 2.7.3), we can then find a repeated game prior μ_{-Row} that is supported on similarly history-independent, stage game rationalizable strategies for Column, and induces ψ'_{-Row} at each time- t information set. Since Row believes his current actions do not affect Column's future play, s_{Row} is a perfect best response to μ_{-Row} , whatever Row's discount factor $\delta \in [0,1)$. Thus, the sequence $\{a'_{Row}\}$ is generated by a perfect best response to a prior that is supported on similarly structured strategies for Column, which may be similarly rationalized.

⁹ Convergence is no easier to prove for this example than for the general case and so, to save space, the following analysis is confined to explanation rather than proof.

¹⁰ To be sure Kalai and Lehrer (1993) allow for behavioral strategies, in which case there is no one true path of play. This makes their theorems more interesting and difficult than the results reported here, but it does not defeat the basic criticism.

¹¹ Let $\{a'\}$ be the true path of play induced by the players' repeated game strategy profile. Let p_i be the probability that i places on the true action profile a' in the stage game following the true history $\{a^1, \dots, a^{t-1}\}$. (We may derive this from i 's strategy and prior.) The probability that i places on the true path is just $\prod_{t=1}^{\infty} p_t$. Now a basic result on infinite products says that

$\prod_{t=\tau}^{\infty} p_t > 0$ for some $\tau \geq 1$ is equivalent to $\lim_{\tau \rightarrow \infty} \prod_{t=\tau}^{\infty} p_t = 1$. But $\prod_{t=\tau}^{\infty} p_t$ is just the probability that i 's prior places on the continuation $\{a^\tau, a^{\tau+1}, \dots\}$ of the true path from time τ on.

¹²I use the following conventional notation for products. A product set $X_1 \times \dots \times X_n$ is denoted interchangeably as X . Given any subset S of a product $X_1 \times \dots \times X_n$ (whether or not the subset is

itself a product), S_i denotes the projection of S onto the i th factor and S_{-i} denotes the projection of S onto the product of all factors except the i th.

¹³ Two technical notes about this probability space: First, since each pair of profiles of repeated game strategies and beliefs (types) induces a unique sequence $\{(\theta^t, a^t)\}$, specifying P is the same as specifying a measure on these repeated game objects. Second, specifying one “big” P over all sequences of beliefs and actions is equivalent to specifying a separate probability measure over beliefs and actions at each “node,” $\{\theta^1, a^1, \dots, \theta^{t-1}, a^{t-1}\}$.

¹⁴ ASSUMPTION 0 is less general than it might be. First, all results hold with a sufficiently small, non-zero discount rate. Second, (as with all the assumptions) it need only hold for almost all t . Third, it suffices that players’ play best response to some belief that is “almost” supported on their current beliefs’ support and that this is “almost” common knowledge.

¹⁵ Clearly some genericity requirement (such as “all pure equilibria are strict”) will typically be necessary, since these classes of games typically include those with constant payoffs.

¹⁶ Again I abuse notation by using (a^s, \dots, a^{t-1}) to denote both the sequence and the subset of A which is the range of that sequence.