Discussion Paper No. 597

BELIEFS IN REPEATED GAMES

John H. Nachbar

September 2003

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Beliefs in Repeated Games

John H. Nachbar*

September 23, 2003

Abstract

Consider a two-player discounted infinitely repeated game. A player's *belief* is a probability distribution over the opponent's repeated game strategies. This paper shows that, for a large class of repeated games, there are no beliefs that satisfy three conditions, *learnability*, *consistency*, and a diversity condition, *CS*. This impossibility theorem generalizes results in Nachbar (1997).

KEYWORDS: Bayesian learning, repeated games.

1 Introduction.

1.1 A statement of the result.

Consider an infinitely repeated game with discounting and perfect monitoring based on a finite stage game. For simplicity, I restrict attention to two-player games, but the results extend to games with any finite number of players. A *belief* is a probability distribution over the opponent's repeated game strategies. This paper establishes an impossibility theorem for beliefs.

By way of motivation and illustration, suppose that each player has only two stage game actions, H and T. Suppose further that each player is certain that the opponent's repeated game behavior strategy is i.i.d. – in each period, the opponent plays H with some probability q, independent of history – but is uncertain as to the value of $q \in [0, 1]$. Each player's belief can then be represented as a probability over $[0, 1]$. Suppose that each player's belief has a continuous and strictly positive density. In the special case in which each player's belief is a Beta distribution, and

---

*Department of Economics, Box 1208; Washington University; One Brookings Drive; St. Louis, MO 63130. Phone: (314) 935-5612; e-mail: nachbar@wueconc.wustl.edu. I wish to thank the Institute for Social and Economic Research at Osaka University and the Center in Political Economy at Washington University for their support. I wish to thank the editor and three anonymous referees for their comments. And I would like to thank Larry Blume, Yaw Nyarko, Alvaro Sandroni, Rann Smorodinsky, and Bill Zame for helpful conversations. Finally, and as will be obvious to readers who know this literature, both this paper and its predecessor, Nachbar(1997), owe a great debt to Kalai and Lehrer (1993).

players best respond, it is a folk theorem that, for any discount factor, this Bayesian model is equivalent to fictitious play.[1]

In this model, player beliefs satisfy three conditions. The belief supports are *learnable*, meaning that if player i's opponent were to play an i.i.d. strategy then (one can show) player i's one-period-ahead forecasts along the path of play would become accurate asymptotically. Beliefs are *weakly cautious* because the support of each player's belief over q is the entire interval $[0, 1]$; no i.i.d. strategy is excluded (as defined, weak caution actually requires less than this). And beliefs are *symmetric* in that they have the same supports, reflecting the symmetry of the strategy sets in this repeated game. Symmetry does not require that probabilities over q have the same densities. For general definitions (not restricted to i.i.d. strategies) and additional motivation, see Section 4.1 (learnability) and Section 4.3 (weak caution and symmetry).

But player beliefs are not *consistent*. Consistency requires that each player have a best response in the support of his opponent's belief; the formal definition is in Section 4.2. It is not hard to see that, with any beliefs of the sort described, no i.i.d. strategy will be optimal, except in the trivial case that one of the stage game actions is weakly dominant. Informally, each player is certain the other is playing an i.i.d. strategy even though, if players optimize, neither player is. It is as though each player were certain that he were more sophisticated than his opponent.

The main result of this paper, Theorem 1, is that this generalized, Bayesian version of fictitious play exhibits a phenomenon that is common in repeated game models: for a large class of repeated games, there are no beliefs that simultaneously satisfy learnability, consistency, and *CS*, a diversity condition that is weaker than weak caution and symmetry.

The proof of Theorem 1 rests on the following observation. For a large class of repeated games, for any strategy σ_1 that player 1 might play, there is a strategy σ_2, an *evil twin* of σ_1, such that, were player 2 actually to play σ_2 and were player 1's one-period-ahead forecasts along the path of play to become accurate asymptotically, then player 1 would eventually learn that continued play of σ_1 was suboptimal. Because a player can hold erroneous beliefs about his opponent's behavior off the path of play even if his forecasts along the path of play are accurate, the statement that σ_2 is σ_1's evil twin is strictly stronger than the statement that σ_1 is not a best response to σ_2.

It follows from the definitions that if player 1's belief satisfies learnability and if σ_1's evil twin is in the support of player 1's belief then σ_1 is not optimal. Say that beliefs have the *evil twin property* if *every* strategy in the support of player 2's belief about player 1 has an evil twin in the support of player 1's belief about player 2, and vice versa. If learnability is satisfied and if the evil twin property holds then consistency fails. This observation, which is simply a marshalling of definitions, is recorded in the paper as Theorem 2. The bulk of the proof of Theorem 1 is devoted

[1] See the appendix to Chapter 2 in Fudenberg and Levine (1998) and also Lehrer (1996).

to showing that CS implies the evil twin property.

The above discussion glossed over an ambiguity in the definition of "support" stemming from the fact that there are typically many equivalent ways to represent the same belief. For example, in the two action case, the degenerate belief that assigns probability one to the i.i.d. behavior strategy with $q = 1/2$ is outcome equivalent to the belief that is uniform over pure strategies that specify actions as a function of time but that are otherwise independent of history (for example, the strategy "play H in period $1, T$ in period 2, etc."). This paper finesses this ambiguity by stating results in a form that is independent of the belief representation. A more accurate statement of Theorem 1 is, *For a large class of repeated games, for any beliefs, there are no strategy subsets that are simultaneously learnable, consistent, and satisfy CS.* Learnability of a given strategy subset holds for one belief if and only if it holds for all outcome equivalent beliefs, and similarly for consistency. The definition of CS does not depend on beliefs at all.

In the remainder of this introductory section, I discuss the application of Theorem 1 to models of out of equilibrium learning and I review some of the other relevant literature. Section 2 provides basic notation and definitions for repeated games. Section 3 deals with beliefs and learning. Section 4 provides formal definitions of learnability, consistency, and CS. Section 5 states and proves Theorem 1. Section 6 contains some concluding remarks.

1.2 An application to learning in repeated games.

Suppose that two players are engaged in a repeated game, with neither player knowing the strategy of her opponent. One would like to argue that, via Bayesian updating of their beliefs, players learn to make increasingly accurate forecasts, and that, as a result, play asymptotically resembles that of a Nash equilibrium of the repeated game. This is the class of learning models examined in Kalai and Lehrer (1993). A natural way to proceed in constructing such a model is to start by identifying strategy subsets, to be interpreted as the supports of player beliefs, with the property that there are, in fact, beliefs for which these sets are both learnable and consistent. For such beliefs, consistency guarantees that an optimizing player might play a strategy in the support of his opponent's belief, and, assuming such a strategy is played, learnability then guarantees that each player will indeed learn to make increasingly accurate forecasts. Theorem 1 says that, for a large class of repeated games, any such strategy sets will violate CS. Put differently, Theorem 1 says that, given learnability, consistency requires that a certain kind of equilibrium assumption be built into the supports of player beliefs. Consistency doesn't require that the supports exclude everything except i's actual strategy, as in a Nash equilibrium. But consistency goes a non-trivial distance in that direction.

To head off confusion, let me make two additional comments. First, there is a well developed literature that examines within equilibrium (rather than out of equilibrium) learning in repeated games with a type space structure. In particular, each player may privately learn his stage game payoff function, drawn from a commonly known distribution, before engaging in the repeated game. A strategy in the type space game is a map from types to strategies in the repeated game. By "within equilibrium," I mean that players play a Nash equilibrium of the type space game. The classic cite is Jordan (1991).

The assumption that players are in a Nash equilibrium means that, in effect, each player knows the other's type space strategies. If one relaxes the Nash equilibrium assumption, the analog of the results in this paper hold, now for type space strategies rather than for strategies in the realized repeated game; see Nachbar (2001) and also Section 4.1.

Second, there is a large literature in which players use *prediction rules*, which are functions from histories to one period ahead forecasts of the opponent's action. A prediction rule is formally equivalent to a behavior strategy, and so a model in which players use prediction rules is formally equivalent to a Bayesian model in which each player's belief places probability 1 on a behavior strategy.[2] There will, typically, be many other, less trivial, belief representations generating the same prediction rule. Theorem 1 states that, for a large class of repeated games, for any learning model using prediction rules, there is no equivalent Bayesian model in which the support of player beliefs are learnable, consistent, and satisfy CS.[3]

1.3 Other relevant literature.

Theorem 1 generalizes the results of Nachbar (1997). Most notably, Nachbar (1997) had left largely unresolved whether an impossibility result continues to hold if players simultaneously (a) ε optimize rather than optimize and (b) randomize.[4] If the answer were "no" then the impossibility results in Nachbar (1997) could be avoided by appealing to a form of bounded rationality. The present paper shows that the impossibility results of Nachbar (1997) are robust in this respect. The present paper also provides a reformulation that is, I think, much more transparent.

Propositions 4 and 5 of Nachbar (1997) are immediate corollaries of Theorem 1 of this paper. The Theorem and Proposition 3 of Nachbar (1997) are corollaries as well, with the minor qualification that weak caution condition (1) is stronger than its counterpart in Nachbar (1997). See also the discussion of neutrality in Remark 4 in Section 4.3 of this paper. Theorem 1 is also stronger than its analogs in Nachbar (1997) in that it uses a weaker definition of prediction, and in that it applies to a

[2] In many such models, players maximize, or ε maximize, payoffs period-by-period. Myopic ε optimization corresponds to $\delta = 0$ uniform ε optimization in the present paper; see Section 3.3.

[3] I have implicitly assumed that prediction rules are deterministic. There is also a literature (e.g. Young (1993)) in which players use prediction rules that are probabilistic. Probabilistic prediction rules are not equivalent to beliefs over strategies and hence the players using them are not Bayesian in the standard sense. Theorem 1 has no implications for such learning models.

[4] Proposition 5 in Nachbar (1997) showed, using a computability argument, that the inconsistency arises if the $\hat{\Sigma}_i$ comprise the behavior strategies that can be implemented by Turing machines with access to randomization devices. Proposition 5 in Nachbar (1997) was restricted to δ close to zero.

somewhat larger class of repeated games (condition MM in this paper is somewhat weaker than its analog, assumption M, in Nachbar (1997)). The basic argument, however, is essentially the same as in Nachbar (1997); in particular the proof of Lemma 3 of this paper is a modification of the proof of Proposition 1 in Nachbar (1997).

Both this paper and Nachbar (1997) are close relatives of Dawid (1985). Roughly put, Dawid (1985) points out that if a Bayesian, looking at the data generated by an unknown stochastic process, thinks that the set of possible stochastic processes satisfies a condition similar to weak caution then the Bayesian's prediction rule, which is itself a stochastic process, will be, loosely speaking, more complicated than any of the processes that he thinks are possible. Both Nachbar (1997) and the present paper can be viewed as exploring the implications of this observation for game theory.

Finally, as noted in Nachbar (1997), these impossibility results are reminiscent of difficulties that arise in static (rather than repeated) environments in which each player chooses a decision procedure that takes his opponent's decision procedure as input. The classic cite is Binmore (1987). I refer the reader to Nachbar (1997) for a discussion.

2 Repeated Games.

As in Nachbar (1997), I focus on two-player infinitely repeated games of perfect monitoring. The analysis extends easily to games with more than two players, but adding additional players introduces the possibility that correlation may be built into beliefs.

2.1 The stage game.

Let A_i denote the set of *actions* available to player i in the stage game. Let a_i denote an element of A_i. I assume that $|A_i| < \infty$, where $|A_i|$ is the cardinality of A_i. Let $A = A_1 \times A_2$. An element $a = (a_1, a_2) \in A$ is an *action profile.*

Let $\Delta(A_i)$ denote the set of probability mixtures over A_i. $\Delta(A_i)$ can be identified with the unit simplex in $\mathbb{R}^{|A_i|}$.

Each player has a stage game *payoff function*, $u_i : A \to \mathbb{R}$. I also let u_i denote the mixed extension of u_i. Thus $u_i(\alpha) = \mathbb{E}_\alpha u_i(a)$, where $\mathbb{E}_\alpha$ denotes expectation with respect to the measure over A defined by the mixed action profile $\alpha \in \Delta(A_1) \times \Delta(A_2)$.

2.2 Histories and paths of play.

An n period *history*, denoted h, is an element of A^n, the n-fold Cartesian product of A. Let h^0 denote the null history, the history that obtains before play begins. Let $\mathcal{H}$ denote the set of all finite histories, including h^0.

A *path of play*, denoted z, is an infinite history, an element of A^∞. Let $\mathcal{Z} = A^\infty$ denote the set of paths of play. z_n denotes the action profile played at date n under the path of play z. $\pi(z, n) \in A^n$ denotes the projection of z onto its first n coordinates, giving the initial n period history determined by z.

Make $\mathcal{Z}$ measurable by giving it the σ-algebra generated by the cylinders $C(h) \subset \mathcal{Z}$, where $C(h)$ is the set of paths of play with initial segment h. Let $\Delta(\mathcal{Z})$ denote the set of probability measures over $\mathcal{Z}$.

2.3 Behavior strategies and mixed strategies.

Players at date $n + 1$ know the realized n period history (perfect monitoring). I adopt the convention that the term *action* refers to the stage game while the term *strategy* refers to the repeated game. A *behavior strategy* for i is a function of the form

$$\sigma_i : \mathcal{H} \to \Delta(A_i).$$

Given a behavior strategy σ_i and a history h, the probability that player i chooses action a_i in the period following h is $\sigma_i(h)(a_i)$. Let Σ_i denote the set of i's behavior strategies. Since $\mathcal{H}$ is countable, Σ_i can be identified with $[\mathbb{R}^{|A_i|}]^\infty$. Let $\Sigma = \Sigma_1 \times \Sigma_2$.

I use s_i to denote a *pure strategy*, which is a behavior strategy that, for each h, assigns probability 1 to an element of A_i. Following history h, if the pure strategy s_i assigns probability 1 to a_i then I write $s_i(h) = a_i$. $S_i \subset \Sigma_i$ denotes the set of player i's pure strategies.

Make Σ_i measurable by giving $[\mathbb{R}^{|A_i|}]^\infty$ the product Borel σ algebra. Let $\Delta(\Sigma_i)$ denote the set of probability measures over Σ_i.

2.4 Induced measures over $\mathcal{Z}$.

A behavior strategy profile σ induces a probability measure $\mu_\sigma \in \Delta(\mathcal{Z})$. Similarly, a mixed strategy profile $\beta \in \Delta(\Sigma_1) \times \Delta(\Sigma_2)$ induces a probability measure $\mu_\beta \in \Delta(\mathcal{Z})$, and so on for combinations of mixed and behavior strategies.

Say that history h is *reachable* under the behavior strategy profile σ iff $\mu_\sigma(C(h)) > 0$. Say that history h is reachable under σ_1 iff there exists some pure strategy s_2 for which $\mu_{(\sigma_1, s_2)}(C(h)) > 0$. Analogous definitions hold for mixed strategies.

2.5 Outcome equivalence.

Two behavior strategies σ_1 and σ_1^* are *outcome equivalent* iff for any behavior strategy $\sigma_2 \in \Sigma_2$,

$$\mu_{(\sigma_1, \sigma_2)} = \mu_{(\sigma_1^*, \sigma_2)}.$$

One can verify that σ_i and σ_i^* are outcome equivalent iff $\mu_{(\sigma_1, \beta_2)} = \mu_{(\sigma_1^*, \beta_2)}$ for any β_2. The definition of outcome equivalence for mixed strategies is analogous.

One can prove the following variant of Kuhn's theorem, establishing outcome equivalence between mixed and behavior strategies (in the standard version of Kuhn's Theorem, mixtures are over S_i rather than Σ_i).

Lemma 1 (Kuhn's Theorem). *For any mixed strategy β_i there is a behavior strategy σ_i such that β_i is outcome equivalent to σ_i, and conversely.*

If σ_i is outcome equivalent to the mixed strategy β_i then say that σ_i is a *reduced form* of β_i.

Much of this paper implicitly assumes that players adopt behavior strategies rather than mixed strategies. By Kuhn's Theorem, this is largely a matter of interpretation. If one wants, for reasons of interpretation, to retain mixed strategies, one way to do so is to introduce types. The results of this paper carry over into type space models; see Section 1.2.

2.6 Repeated game payoffs.

Player i's payoff in the repeated game is the expectation of the discounted value of the payoffs he receives along the realized path of play. Formally, given a behavior strategy profile σ, the expected payoff to player i is

$$V_i(\sigma) = \mathbb{E}_{\mu_\sigma}\left(\sum_{n=1}^{\infty} \delta^{n-1} u_i(z_n)\right),$$

where $\delta \in [0, 1)$. One can analogously define $V_i(\beta)$ and $V_i(\sigma_1, \beta_2)$.

2.7 Continuation games.

An n period history h defines a *continuation game*, the repeated game starting in period $n + 1$. In the continuation game following h, a behavior strategy σ_i induces a continuation behavior strategy σ_{ih} defined by

$$\sigma_{ih}(h') = \sigma_i(h \cdot h')$$

for any history h', where $h \cdot h'$ denotes the concatenation of h and h'. Given a profile $\sigma = (\sigma_1, \sigma_2)$, let $\sigma_h = (\sigma_{1h}, \sigma_{2h})$.

The profile σ_h induces the distribution μ_{σ_h} over the set of continuation paths, which is simply $\mathcal{Z}$. Given a history h and the profile σ, the expected continuation payoff to player i is

$$V_i(\sigma_h) = \mathbb{E}_{\mu_{\sigma_h}}\left(\sum_{n=1}^{\infty} \delta^{n-1} u_i(z_n)\right).$$

Note that, in this definition, payoffs are discounted to the start of the period following h, rather than back to the first period.

3 Beliefs, Optimization, and Learning.

3.1 Beliefs.

A *belief* about player i is a probability distribution over player i's behavior strategies. A belief about player i is thus formally equivalent to a mixed strategy for player i. By Kuhn's Theorem, for any belief β_i there is a reduced form σ_i^β, not uniquely defined, that is outcome equivalent to β_i, and conversely. It is often more convenient to work with a reduced form σ_i^β than with β_i.

3.2 Bayesian updating.

As the game proceeds, each player learns by Bayesian updating of her prior. Given a prior β_i, any history h that is reachable under β_i, and any reduced form σ_i^β, a reduced form of the posterior over continuation game strategies is simply σ_{ih}^β.

3.3 Optimization.

Given $\varepsilon \geq 0$, σ_1^* is an *ex ante* ε best response to β_2 (or σ_1^* is *ex ante* ε optimal) iff

$$V_1(\sigma_1^*, \beta_2) + \varepsilon \geq \max_{\sigma_1 \in \Sigma_1} V_1(\sigma_1, \beta_2).$$

Even if a behavior strategy is *ex ante* ε optimal, the induced continuation strategy may be ε *sub*optimal in subgames that are far in the future or that the player views as unlikely. Should such a subgame be reached, the player would, presumably, deviate from her behavior strategy. The following stronger version of ε optimization eliminates this problem for subgames that are reachable given the player's belief (see also Lehrer and Sorin (1998)).

Definition 1. *Given $\varepsilon \geq 0$, σ_1^* is a* uniform ε best response *to β_2 (or σ_1^* is* uniformly ε optimal*) iff, for any reduced form σ_2^β of β_2 and any h that is reachable under $(\sigma_1^*, \sigma_2^\beta)$,*

$$V_1(\sigma_{1h}^*, \sigma_{2h}^\beta) + \varepsilon \geq \max_{\sigma_1 \in \Sigma_1} V_1(\sigma_1, \sigma_{2h}^\beta).$$

If σ_1^* is a uniform ε best response to β_2, write $\sigma_1^* \in \mathrm{BR}_1^\varepsilon(\beta_2)$. The definition for player 2 is analogous.

Note that if $\delta = 0$ then uniform ε optimization corresponds to myopic (period-by-period) ε optimization.

3.4 Learning to predict the path of play.

Informally, player 1 learns to predict the path of play generated by σ_2 (and player 1's own strategy) if her one period ahead forecasts along the path of play eventually become almost as accurate as if she knew σ_2. This is not the same thing as saying that player 1 learns σ_2. Player 1 could learn to predict the path of play generated by σ_2 and still hold erroneous beliefs about what player 2 would do off the path of play. Note also that if $\sigma_i(h) = \alpha_i$ is not pure then accurate prediction means only that the player predicts approximately α_i, not that she predicts the realized action.

Recall that $\pi(z, n)$ is the history corresponding to the n period initial segment of z and that $C(h)$ is the cylinder of all paths of play with initial segment h.

Definition 2. *Fix a belief β_2. Player 1* learns to predict the path of play *generated by the behavior strategy profile $\sigma = (\sigma_1, \sigma_2)$ iff the following conditions hold.*

1. *$\mu_\sigma(C(h)) > 0$ implies $\mu_{(\sigma_1,\beta_2)}(C(h)) > 0$ for any finite history h.*
2. *For any real number $\eta > 0$ and μ_σ almost any path of play z, there is a period $n(\eta, z)$ such that, for any $n > n(\eta, z)$ and any $a_2 \in A_2$, letting $h = \pi(z, n)$,*

$$|\sigma_2(h)(a_2) - \sigma_2^\beta(h)(a_2)| < \eta.$$

A similar definition holds for player 2.

This definition is equivalent to the one given in Nachbar (1997). Prediction corresponds to the more general concept of *weak merging*; see Kalai and Lehrer (1994).

As defined, prediction requires accurate forecasts in every period following period $n(\eta, z)$. To be able to apply results from Lehrer and Smorodinsky (1996) (see also Section 4.1), I consider a weaker form of prediction that requires forecasts to be accurate only on a set of dates of density 1. Formally, let $\mathbb{N}$ be the set of natural numbers (excluding zero) and consider a set $\mathbb{N}^\diamond \subset \mathbb{N}$. Say that $\mathbb{N}^\diamond$ has *density 1* iff

$$\lim_{n\to\infty} \frac{|\{1, 2, \ldots, n\} \cap \mathbb{N}^\diamond|}{n} = 1.$$

Definition 3. *Fix a belief β_2. Player 1* weakly learns to predict the path of play *generated by the behavior strategy profile $\sigma = (\sigma_1, \sigma_2)$ iff the following conditions hold.*

1. *$\mu_\sigma(C(h)) > 0$ implies $\mu_{(\sigma_1,\beta_2)}(C(h)) > 0$, for any finite history h.*
2. *For any real number $\eta > 0$ and μ_σ almost any path of play z, there is a set $\mathbb{N}^P(\eta, z) \subset \mathbb{N}$ of density 1 such that, for any $n \in \mathbb{N}^P(\eta, z)$ and any $a_2 \in A_2$, letting $h = \pi(z, n)$,*

$$|\sigma_2(h)(a_2) - \sigma_2^\beta(h)(a_2)| < \eta.$$

A similar definition holds for player 2.

Weak prediction corresponds to the more general concept of *almost weak merging* introduced in Lehrer and Smorodinsky (1996). Prediction implies weak prediction but not conversely.

4 Consistency, Learnability, and CS

4.1 Learnability.

Here and throughout, $\hat{\Sigma}_i \subset \Sigma_i$ and $\hat{\Sigma} = \hat{\Sigma}_1 \times \hat{\Sigma}_2$.

Definition 4. *Given $\hat{\Sigma} \subset \Sigma$ and belief β_2, $\hat{\Sigma}_2$ is* learnable *iff for any $\sigma_1 \in \hat{\Sigma}_1$ and any $\sigma_2 \in \hat{\Sigma}_2$, player 1 weakly learns to predict the path of play. An analogous definition holds for learnability of $\hat{\Sigma}_1$. $\hat{\Sigma}$ is learnable iff $\hat{\Sigma}_i$ is learnable for each i.*

There are product sets that are not learnable for any beliefs. In particular, Theorem 1 implies that if either player has at least two actions in the stage game then there are no beliefs for which Σ, the full product set of behavior strategies, is learnable; this is a game theoretic version of an observation about stochastic processes made in Dawid (1985). Thus, $\hat{\Sigma} \subset \Sigma$ can be learnable only if it is sufficiently small in some sense.

A well known sufficient condition for player 1 to learn to predict the path of play is for player 1's belief to assign positive probability to player 2's actual strategy. Even if the probability assigned to player 2's strategy is extremely low, a strong form of prediction will be satisfied: player 1 learns to accurately forecast the entire infinite tail of the game, rather than merely the next period. Since every element of a countable set can be given positive probability, it follows that if $\hat{\Sigma}$ is countable there are beliefs for which that set will be learnable.

Uncountable $\hat{\Sigma}$ can also be learnable. Example 2 in Lehrer and Smorodinsky (1996) implies that if the $\hat{\Sigma}_2$ has a "nice" finite dimensional parameterization then there are beliefs for which $\hat{\Sigma}_2$ is learnable. By "nice," I mean that if two parameters are close in the standard Euclidean topology then the associated strategies are likewise close under a particular metric on Σ_2.[5] The i.i.d. example of Section 1.1 is an example of a "nice" finite dimensional parameterization.[6]

Theorem 1 can be viewed as a purely positive statement: if the strategy sets happen to be learnable then either consistency or CS must fail. But one can argue that learnability *should* be "built in" by choosing belief representations in which, as in the i.i.d. example, beliefs are probabilities over learnable strategies. Such learnable representations are always possible: given any belief β_i, take the alternative

[5] Adapting the definition of neighborhood in Lehrer and Smorodinsky (1996) to the present context (strategies in repeated games rather than stochastic processes), σ' is in the ε neighborhood of σ iff, for any history h that is reachable under σ_2 and any $a_2 \in A_2$ for which $\sigma(h)(a_2) > 0$,

$$|\sigma'_2(h)(a_2) - \sigma_2(h)(a_2)| < \varepsilon\sigma_2(h)(a_2).$$

This definition of neighborhood is more restrictive than that implied by the sup norm. It is not known whether the sup norm will suffice.

[6] One can show that beliefs in the i.i.d. example allow prediction rather than merely weak prediction. A natural conjecture is that, more generally, beliefs allow prediction rather than merely weak prediction if, as in the i.i.d. example, the parameter set is closed and the distribution over parameters admits a strictly positive density. But at this writing this conjecture has not been established.

representation that places probability 1 on the reduced form σ_i^{β}. This representation is trivial, but non-trivial learnable representations are typically also available; see, in particular, the representation proposed by Jackson, Kalai, and Smorodinsky (1999).

An argument for adopting learnable representations is the following. Suppose that the beliefs used in Bayesian models are a construction by boundedly rational players who consider the game, decide what strategies their opponents might play, and choose probability distributions over those strategies. If a player really thinks that an opposing strategy σ_i is possible then his belief should give every neighborhood of σ_i positive probability. If one defines neighborhood with a sufficiently strong metric, a metric that takes into account behavior in the tails, then σ_i will be learnable (exactly what metrics will serve remains an open question; see also footnote 5).

Adoption of a strong metric implies that players care about the tail of the game, which conflicts with modeling them as discounting ε optimizers. One, informal, justification for nevertheless focusing on learnable representations mirrors the justification offered for uniform, as opposed to *ex ante*, ε optimization. Upon reaching the subgame defined by history h, a player wants to get a high average payoff from that point forward. So his continuation strategy should be ε optimal with respect to his posterior, and his posterior should itself be "good." The player should not, on reaching h, want to violate Bayes's rule by reconsidering the game and constructing a new posterior to permit forecasting of strategies effectively ruled out by his prior. To avoid this kind of dynamic problem, the priors in a Bayesian model should be modeled *as if* players cared about the tail of the game.[7] Whether one can build a coherent model of belief construction along the lines sketched remains a topic for future work.

Remark 1. Weak prediction is stronger than necessary for the main results of this paper. As will be clear from the proofs, a weaker condition that will work just as well is that, for any positive integer ℓ, with high probability, there is a period after which the player's forecast of his opponent's play over the next ℓ periods is highly accurate. Even one such period will serve; the proof does not require that the set of such periods have density 1.

Also, for the sake of generality, the definition of learnability, like the definition of conventional prediction in Nachbar (1997), requires prediction by player 1 only if she herself plays a strategy in $\hat{\Sigma}_1$. The sufficient conditions for prediction given above do not, in fact, require that player 1's strategy be restricted to $\hat{\Sigma}_1$. The results hold *a fortiori* if the definition of learnability is strengthened to require prediction regardless of one's own strategy. □

Remark 2. In type space models (see Section 1.2), the requirement that players learn to predict the path of play generated by the opponent's realized repeated

[7] For a model, necessarily non-Bayesian, in which players do jettison their Bayes's rule posteriors, see Foster and Young (2003).

game strategy can be too strong, and must be replaced by the appropriate analog, namely that players learn to predict the path of play generated by the opponent's type space strategy (a map from types to repeated game strategies). See Nachbar (2001). Thus defined, learnability is compatible with purification in the spirit of Harsanyi (1973), a phenomenon that has been emphasized in the context of type space learning models by Jackson and Kalai (1997) (although, strictly speaking, this is for recurring rather than repeated games), Jordan (1995), and Nyarko (1998). □

4.2 Consistency.

Definition 5. *Given beliefs β_1 and β_2, $\hat{\Sigma} \subset \Sigma$ is ε* consistent *iff for any $\varepsilon > 0$, player 1 has a uniform ε best response in $\hat{\Sigma}_1$ and player 2 has a uniform ε best response in $\hat{\Sigma}_2$: $\mathrm{BR}_1^{\varepsilon}(\beta_2) \cap \hat{\Sigma}_1 \neq \emptyset$ and $\mathrm{BR}_2^{\varepsilon}(\beta_1) \cap \hat{\Sigma}_2 \neq \emptyset$. $\hat{\Sigma}$ is* consistent *iff $\hat{\Sigma}$ is ε consistent for every $\varepsilon > 0$.*

A trivial learnable and consistent strategy set is a Nash equilibrium strategy profile with associated beliefs (probability 1 on the opponent's actual strategy). A Nash equilibrium profile, however, violates CS, defined below.

4.3 CS, weak caution, and symmetry.

Definition 6. $\hat{\Sigma} \subset \Sigma$ satisfies CS *iff the following conditions hold.*

1. *There is an $\xi \in (0,1)$ such that, for each i, the following is true. Consider any $\sigma_i \in \hat{\Sigma}_i$. There is a pure strategy $s_i \in \hat{\Sigma}_i$ such that, for any history h, if $s_i(h) = a_i$ then $\sigma_i(h)(a_i) > \xi$.*

2. *Consider any pure strategy $s_1 \in S_1$. For any function $\gamma_{12} : A_1 \to A_2$ there is a pure strategy $s_2 \in \hat{\Sigma}_2$ such that the following is true. Let z be the path of play generated by (s_1, s_2). There is a set $\mathbb{N}^{\gamma}(z) \subset \mathbb{N}$ of density 1 such that for any $n \in \mathbb{N}^{\gamma}(z)$, letting $h = \pi(z, n)$,*

   $$s_2(h) = \gamma_{12}(s_1(h)).$$

   And an analogous statement holds for $s_2 \in \hat{\Sigma}_2$ and $\gamma_{21} : A_2 \to A_1$.

The motivation for condition (1) is that if a behavior strategy σ_2 is in $\hat{\Sigma}_2$ (to be thought of as the support of player 1's belief) then a prudent player 1 should think that at least one comparatively simple, nonrandomizing variation on σ_2 is contained in $\hat{\Sigma}_2$ as well. The leading candidate for a nonrandomizing variation on a behavior strategy σ_2 is the pure strategy, call it s_2^*, that, for any history h, plays the action to which σ_2 assigns highest probability. s_2^* has the property that, for any h, if $s_2^*(h) = a_2$ then

$$\sigma_2(h)(a_2) \geq \frac{1}{|A_2|}.$$

CS condition (1) is in the spirit of requiring that s_2^* be contained in $\hat{\Sigma}_2$, but it is weaker. CS condition (1) requires only that *one* such pure strategy be contained in $\hat{\Sigma}_2$, not that *every* such pure strategy be contained in $\hat{\Sigma}_2$.

CS condition (2) is implied by two stronger conditions, weak caution and symmetry (the name CS is derived from Caution and Symmetry) that are easier to motivate.

Definition 7. *$\hat{\Sigma}_i \subset \Sigma_i$ is* weakly cautious *iff the following is true.*

1. *CS Condition 1 holds.*
2. *Consider any strategy $\sigma_i \in \hat{\Sigma}_i$ and any function $g : A_i \to A_i$ (not necessarily 1-1 or onto). There is a strategy $\sigma_i^g \in \hat{\Sigma}_i$ such that, for any h and any $a_i^* \in g(A_i)$,*

$$\sigma_i^g(h)(a_i^*) = \sum_{a_i \in g^{-1}(a_i^*)} \sigma_i(h)(a_i).$$

$\hat{\Sigma}$ is weakly cautious iff $\hat{\Sigma}_i$ is weakly cautious for each i.

The motivation for weak caution condition (2) is that if a strategy σ_2 is in $\hat{\Sigma}_2$ (still to be thought of as the support of player 1's belief) then a prudent player 1 should include all computationally trivial variants of σ_2 in $\hat{\Sigma}_2$ as well. More explicitly, view a behavior strategy as a black box that takes histories as input and yields stage game actions as output. Condition (2) states that if σ_i is in $\hat{\Sigma}_i$ then so is any behavior strategy generated by using a function g to relabel outputs, leaving the black box, which is where all of the strategic complexity resides, unchanged. For any σ_i, the set of such computationally trivial variants is finite, since I have assumed that the action set is finite.

Naively, one might demand, instead of weak caution, full caution: $\hat{\Sigma}_i = \Sigma_i$. But, as noted in Section 4.1, if either player has at least two stage game actions then $\hat{\Sigma}$ is not learnable for any beliefs. Thus, to avoid triviality when dealing with learnable sets, one must restrict attention to some weakened form of caution. The definition offered here is arguably too strong in that it is payoff independent, but it is not clear how much payoff dependence would be incorporated into the support of a prudent player's belief. The payoff independent formulation is, at a minimum, a natural and useful benchmark.

Definition 8. *$\hat{\Sigma} \subset \Sigma$ is* symmetric *iff the following holds. Consider any behavior strategy $\sigma_1 \in \hat{\Sigma}_1$. Let $A_1^* \subset A_1$ be the set of actions such that if $a_1 \in A_1^*$ then there exists a history h for which $\sigma_1(h)(a_1) > 0$. Let $r = \min\{|A_1^*|, |A_2|\}$. Then there is a set $A_2^* \subset A_2$ with $|A_2^*| = r$, an onto map $\gamma : A_1^* \to A_2^*$, and a behavior strategy $\sigma_2^\gamma \in \hat{\Sigma}_2$ such that, for any $a_2^* \in A_2^*$,*

$$\sigma_2^\gamma(h)(a_2^*) = \sum_{a_1 \in \gamma^{-1}(a_2^*)} \sigma_1(h)(a_1).$$

And a similar condition holds for any behavior strategy $\sigma_2 \in \hat{\Sigma}_2$.

Symmetry is motivated by the idea that, aside from differences in the composition and cardinality of the stage game action sets, the mechanics of constructing strategies are the same for the two players. More explicitly, as in the discussion of weak caution condition (2), view a strategy as a black box that takes histories as inputs and produces stage game actions as outputs. Then symmetry requires that for every $\sigma_1 \in \hat{\Sigma}_1$, $\hat{\Sigma}_2$ contains a strategy σ_2^γ constructed from σ_1 by relabeling outputs (using the function γ) while leaving the black box unchanged. σ_2^γ is the same as σ_1, except that whenever σ_1 chooses, say, *up*, σ_2 chooses, say, H. The definition of symmetry is complicated by the desire to ensure that the relabeling function γ is, loosely speaking, complexity preserving.

Lemma 2. *If* $\hat{\Sigma}$ *is weakly cautious and symmetric then* $\hat{\Sigma}$ *satisfies CS.*

Proof. CS condition (1) is immediate. For CS condition (2), it is easy to verify that weak caution condition (2) and symmetry imply that, for any $\gamma_{12} : A_1 \to A_2$ and any pure strategy $s_1 \in \hat{\Sigma}_1$, there is a pure strategy $s_2 \in \hat{\Sigma}_2$ such that, for any history h (not just histories along the path of play generated by (s_1, s_2)),

$$s_2(h) = \gamma(s_1(h)).$$

That is, CS condition (2) holds for every z (not just z generated by (s_1, s_2)), and $\mathbb{N}^\gamma(z) = \mathbb{N}$. ■

Example 1. Consider any $\hat{\Sigma}$ where, for each i, $\hat{\Sigma}_i$ is defined as the set of strategies that satisfy some bound on strategic complexity. For example, $\hat{\Sigma}_i$ might be the set of strategies that have one period memory, or k-period memory, or are implementable by automata or, or by Turing machines, possibly with access to randomization devices. Any such set is weakly cautious and symmetric, and hence satisfies CS. □

Example 2. Suppose that $\hat{\Sigma}_1$ is the set of pure strategies for player 1 that are implementable by automata and $\hat{\Sigma}_2$ is the set of pure strategies for player 2 that are implementable by Turing machines. Then $\hat{\Sigma}$ violates CS and, in particular, violates symmetry. It is not hard to see, however, that $\hat{\Sigma}$ is nevertheless inconsistent for any beliefs for which $\hat{\Sigma}$ is learnable. □

Example 3. Each player's action set in the stage game is $A_i = \{H, T\}$.

1. Suppose that $\hat{\Sigma}_i$ consists of the two elements $\overline{H}$ and $\overline{T}$, where $\overline{H}$ is the pure strategy "play H in every period" and $\overline{T}$ is the pure strategy "play T in every period." Then $\hat{\Sigma}$ satisfies weak caution and symmetry and hence CS. This $\hat{\Sigma}_i$ is sparse, which underscores the fact that weak caution is, as a characterization of prudence, weak.

2. Suppose that, for each i, $\hat{\Sigma}_i$ consists of the single element "randomize 50:50 in every period, regardless of history," hereafter 50:50. This $\hat{\Sigma}$ violates weak caution and CS. In particular, it is not hard to see that CS requires here that if 50:50 is in $\hat{\Sigma}_i$ then so are $\overline{H}$ and $\overline{T}$.

□

Remark 3. Trembles can be accommodated by the following slight weakening of weak caution condition (1). Fix $k \in (0, 1)$. There is a $\xi > 0$ such that, for any $\sigma_i \in \hat{\Sigma}_i$, there is a behavior strategy $\sigma_i^k \in \hat{\Sigma}_i$ such that, for any history h, there is an $a_i \in A_i$ for which

1. $\sigma_i^k(h)(a_i) > 1 - k$, and
2. $\sigma_i(h)(a_i) > \xi$.

One can verify that the main results continue to hold with this modification, for any k sufficiently small. See also the discussion of trembling in Nachbar (1997). □

Remark 4. Symmetry and weak caution are, collectively, analogous to what Nachbar (1997) called *neutrality*. Symmetry and weak caution condition (2) are weaker than the first four neutrality conditions in Nachbar (1997). Weak caution condition (1) is stronger than neutrality condition (5) in Nachbar (1997). Neutrality condition (5) in Nachbar (1997) in effect allowed the bound ξ to depend both on the history and on the behavior strategy. Weak caution (1) requires ξ to be uniform across histories and behavior strategies. Although, as a technical exercise, one can construct examples in which neutrality condition (5) holds but weak caution condition (1) fails, I am not aware of any interesting examples where this occurs. □

5 Impossibility.

5.1 The main result and the evil twin property.

An action $a_1^* \in A_1$ is *weakly dominant* iff, for any $a_2 \in A_2$,

$$u_1(a_1^*, a_2) \geq \max_{a_1 \in A_1} u_1(a_1, a_2).$$

This definition is somewhat weaker than the standard one in that I do not require strict inequality for any a_2. The definition for player 2 is similar.

Definition 9. *The stage game satisfies* No Weak Dominance (NWD) *iff neither player has a weakly dominant action.*

Player 1's *minmax* payoff is given by

$$m_1 = \min_{\alpha_2 \in \Delta(A_2)} \max_{\alpha_1 \in \Delta(A_1)} u_1(\alpha_1, \alpha_2).$$

Player 1's *pure action maxmin* payoff is given by

$$M_1 = \max_{a_1 \in A_1} \min_{a_2 \in A_2} u_1(a_1, a_2).$$

The definitions for player 2 are analogous.

Definition 10. *The stage game satisfies* MM *iff, for each player i, the pure action maxmin payoff is strictly less than the minmax payoff,*

$$M_i < m_i.$$

Examples of stage games that satisfy MM are matching pennies, rock-scissors-paper, battle of the sexes, and many coordination games.

The main result of the paper is the following.

Theorem 1. *Fix beliefs β_1 and β_2. Consider any $\hat{\Sigma} \subset \Sigma$.*

1. *If NWD holds then there is a $\bar{\delta} \in (0, 1]$ such that, for any $\delta \in [0, \bar{\delta})$, if $\hat{\Sigma}$ is learnable and satisfies CS then $\hat{\Sigma}$ is not consistent.*

2. *If MM holds then, for any $\delta \in [0, 1)$, if $\hat{\Sigma}$ is learnable and satisfies CS and symmetric then $\hat{\Sigma}$ is not consistent.*

The intended interpretation is that, for a large set of repeated games, if beliefs have supports that are both learnable and sufficiently diverse then the beliefs are inconsistent.

NWD cannot be relaxed. If δ is low and NWD fails then at least one of the players will be able to satisfy consistency because, no matter what his belief, it will be a best response for him to play the constant strategy in which he repeatedly plays his weakly dominant stage game action. The repeated prisoner's dilemma, for example, is not vulnerable to this form of inconsistency.

It is not clear whether MM is necessary, although it is used in an important way in the proof of Theorem 3, which is the key technical result.

For more on the hypotheses in Theorem 1, see the discussion following Lemma 3 in Section 5.2.

The proof of Theorem 1 relies on the following concept.

Definition 11. *Fix δ. A behavior strategy $\sigma_2 \in \Sigma_2$ is an ε* evil twin *of a behavior strategy $\sigma_1 \in \Sigma_1$ iff σ_1 is not a uniform ε best response to any belief β_2 for which player 1 weakly learns to predict the path of play generated by (σ_1, σ_2). A similar definition holds for player 2.*

In the coordination game below, it is easy to show that

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

an ε evil twin of any pure strategy s_1, for any ε, is simply s_1 itself, that is, the strategy s_2 defined by, for any h,

$$s_2(h) = s_1(h).$$

In this case, an evil twin is an identical twin. In repeated matching pennies, an ε evil twin of any pure strategy s_1, for any ε, is any best response to s_1. I take up the general construction of evil twins in Section 5.2.

Definition 12. *Fix δ. $\hat{\Sigma} \subset \Sigma$ has the* evil twin property *iff there is an $\varepsilon > 0$ such that, for any $\sigma_1 \in \hat{\Sigma}_1$, there is a strategy $\sigma_2 \in \hat{\Sigma}_2$ that is an ε evil twin of σ_1, and similarly for player 2.*

Marshaling definitions yields the following observation.

Theorem 2. *Fix δ and fix beliefs β_1 and β_2. If $\hat{\Sigma} \subset \Sigma$ is learnable and has the evil twin property then it is not consistent.*

The key technical fact underlying Theorem 1 is that the evil twin property follows from CS.

Theorem 3. *Consider any $\hat{\Sigma} \subset \Sigma$.*

1. *If NWD holds then there is a $\bar{\delta} \in (0, 1]$ such that, for any $\delta \in [0, \bar{\delta})$, if $\hat{\Sigma}$ satisfies CS then $\hat{\Sigma}$ has the evil twin property.*
2. *If MM holds then, for any $\delta \in [0, \bar{\delta})$, if $\hat{\Sigma}$ satisfies CS then $\hat{\Sigma}$ has the evil twin property.*

The proof of Theorem 1 follows immediately from Theorems 2 and 3. It remains to prove Theorem 3.

5.2 The proof of Theorem 3.

The proof of Theorem 3 follows from a sequence of lemmas. I begin by showing how to construct evil twins for pure strategies.

Define $a_2^M : A_1 \to A_2$ by, for any pure action $a_1 \in A_1$,

$$a_2^M(a_1) = \arg \min_{a_2 \in A_2} u_1(a_1, a_2).$$

If the right-hand side is not single valued, arbitrarily pick one of the values to be $a_2^M(a_1)$. The function a_1^M is defined similarly.

Recall that $\pi(z, n)$ is the n period initial segment of the infinite path of play z. Given s_1, define $S_2^M(s_1) \subset S_2$ to be the set consisting of all s_2 for which there exists a set $\mathbb{N}^\diamond \subset \mathbb{N}$ of density 1 such that, for all $n \in \mathbb{N}^\diamond$, letting z denote the path of play generated by (s_1, s_2) and letting $h = \pi(z, n)$,

$$s_2(h) = a_2^M(s_1(h)).$$

The definition of $S_1^M(s_2)$ is analogous.

Define the function $\tilde{a}_2 : A_1 \to A_2$ by, for any pure action $a_1 \in A_1$,

$$\tilde{a}_2(a_1) = \arg\max_{a_2 \in A_2} \left[\max_{a_1' \in A_1} u_1(a_1', a_2) - u_1(a_1, a_2) \right].$$

If the right-hand side is not single valued, arbitrarily pick one of the values to be $\tilde{a}_2(a_1)$. The function $\tilde{a}_1$ is defined similarly. Loosely, $\tilde{a}_2(a_1)$ is the action that gives player 1 maximal one-period incentive *not* to play a_1. $\tilde{a}_2(a_1)$ does not necessarily minimize player 1's payoff from a_1. That is, it is not necessarily true that $\tilde{a}_2(a_1) = a_2^M(a_1)$.

Given s_1, define $\tilde{S}_2(s_1) \subset S_2$ to be the set consisting of all s_2 for which there exists a set $\mathbb{N}^\diamond \subset \mathbb{N}$ of density 1 such that, for all $n \in \mathbb{N}^\diamond$, letting z denote the path of play generated by (s_1, s_2) and letting $h = \pi(z, n)$,

$$s_2(h) = \tilde{a}_2(s_1(h))$$

The definition of $\tilde{S}_1(s_2)$ is analogous.

Lemma 3 confirms that, for a large set of games, elements of either $\tilde{S}_1(s_2)$ or $S_1^M(s_2)$ are indeed evil twins of s_1. The proof is in the appendix.

Lemma 3.

1. *Suppose that NWD holds. Then there is a $\bar{\delta} \in (0, 1)$ and an $\varepsilon' > 0$ such that, for any $\delta \in [0, \bar{\delta})$, any $\varepsilon \in [0, \varepsilon')$, and any pure strategy $s_1 \in S_1$, if $s_2 \in \tilde{S}_2(s_1)$ then s_2 is an ε evil twin of s_1. And an analogous statement holds for any $s_2 \in S_2$.*

2. *Suppose that MM holds. Then there is an $\varepsilon' > 0$ such that, for any $\delta \in [0, 1)$, any $\varepsilon \in [0, \varepsilon')$, and any pure strategy $s_1 \in S_1$, if $s_2 \in S_2^M(s_1)$ then s_2 is an ε evil twin of s_1. And an analogous statement holds for any $s_2 \in S_2$.*

Informally, the logic of the argument is as follows. If NWD holds, then $s_1(h)$ is not a stage game ε best response to $\tilde{a}_2(s_1(h))$ for any h, for ε sufficiently small. If δ is close to zero, therefore, and if player 1 learns to predict the path of play, then player 1 will learn to predict that the continuation strategy induced by s_1 is not ε optimal on a set of periods of density 1, which means that s_1 is not uniformly ε optimal.

On the other hand, for any δ, if MM holds, then player 1 will learn to predict that his payoff average is at most M_1, which, by assumption, is strictly less than the minmax payoff m_1. But a uniform ε optimizing player always expects to earn at least $m_1 - \varepsilon$, on average, in any continuation game he thinks is reachable.

If MM fails then it is possible that player 1 never learns to expect to receive less than m_1, on average. Although meager, a payoff average of m_1 can be consistent with ε optimization, depending on what player 1 expects to happen should she deviate. In this regard, recall that the Folk Theorem states that player 1 can receive a payoff average arbitrarily close to m_1 in equilibrium (indeed, in subgame perfect equilibrium). This said, I do not know of any examples in which (a) MM is violated but NWD is satisfied, and (b) learnability, consistency, and CS are all satisfied.

The next lemma records that evil twins of the sort just constructed will be included in any $\hat{\Sigma}$ that satisfies CS.

Lemma 4. *Suppose that $\hat{\Sigma}$ satisfies CS. Consider any $s_1 \in \hat{\Sigma}_1$.*

1. *There is an $s_2 \in \hat{\Sigma}_2 \cap \tilde{S}_2(s_1)$.*
2. *There is an $s_2 \in \hat{\Sigma}_2 \cap S_2^M(s_1)$.*

And analogous statements hold with the roles of the players reversed.

Proof. Given s_1 as in the statement of the lemma, define $\gamma_{12} : A_1 \to A_2$ by $\gamma_{12}(a_1) = \tilde{a}_2(a_1)$. In view of the definition of $\tilde{S}_2(s_1)$, the conclusion then follows from CS condition (2), with $\mathbb{N}^\diamond = \mathbb{N}^\gamma(z)$. The proof for $S_2^M(s_1)$ is similar, as is the proof for the analogous claims with the roles of the players reversed. ■

These lemmas establish Theorem 3, and hence Theorem 1, for the special case in which $\hat{\Sigma}$ is pure. It remains to extend the argument to $\hat{\Sigma}$ that include randomizing behavior strategies. I do so by means of the following fact about uniform ε optimization. The proof is in the appendix.

Lemma 5. *Consider any belief β_2, any $\varepsilon \geq 0$, any behavior strategy σ_1 that is a uniform ε best response to β_2, and any $\xi > 0$. Consider any pure strategy s_1 such that, for any history h, if $s_1(h) = a_1 \in A_1$ then*

$$\sigma_1(h)(a_1) > \xi.$$

Then s_1 is uniformly

$$\frac{\varepsilon}{(1-\delta)\xi}$$

optimal. A similar statement holds for player 2.

Proof of Theorem 3. Consider any belief β_2 and any behavior strategy $\sigma_1 \in \hat{\Sigma}_1$. By CS condition (1), there is a pure strategy $s_1 \in \hat{\Sigma}_1$ and a $\xi > 0$ such that, for any history h, if $s_1(h) = a_1 \in A_1$, then

$$\sigma_1(h)(a_1) > \xi.$$

By Lemma 5, if σ_1 is a uniform ε best response to player 1's belief then s_1 is a uniform $\varepsilon/[(1-\delta)\xi]$ best response. By Lemma 3 and Lemma 4, it follows by contraposition that s_1 cannot be a uniform $\varepsilon/[(1-\delta)\xi]$ best response for $\varepsilon/[(1-\delta)\xi]$ less than ε', where ε' is as in the statement of Lemma 3, hence σ_1 cannot be a uniform ε best response for ε less than $\varepsilon'(1-\delta)\xi$. The proof follows. ■

6 Conclusion

The paper's main result, Theorem 1, states that, for a large class of repeated games, if beliefs have supports that are learnable and sufficiently diverse, in the sense of satisfying CS, then the supports are inconsistent.

Although one can work with an inconsistent model, like Bayesian fictitious play, consistency seems to me to be so attractive property that I would prefer to retain it, if possible. And, for reasons discussed in Section 4.1, I would also like to represent beliefs as probabilities over learnable strategies, in which case learnability is built in. This leaves CS.

I have chosen definitions of CS, and of weak caution and symmetry, that are based exclusively on complexity considerations, independent of payoffs. This formulation is, I think, a natural benchmark, but a compelling theory of belief formation would also take account of payoffs. It is possible that, because of this, a prudent player, according to some more fundamental definition of prudence, would nevertheless have beliefs that violate CS. But developing such a theory remains a topic for future work.

Appendix

Proof of Lemma 3. The proof is a modification of the proof of Proposition 1 in Nachbar (1997).

Let

$$\underline{w}_1 = \min_{a_1 \in A_1} \left[\max_{a_1' \in A_1} u_1(a_1', \tilde{a}_2(a_1)) - u_1(a_1, \tilde{a}_2(a_1)) \right].$$

By NWD, $\underline{w}_1 > 0$. Let

$$\bar{u}_1 = \max_{a \in A} u_1(a),$$
$$\underline{u}_1 = \min_{a \in A} u_1(a),$$

By NWD, $\bar{u}_1 > \underline{u}_1$.

To prove the first part of the lemma, choose $\bar{\delta}$ sufficiently small that, under uniform ε optimization, player 1 acts to maximize his current period payoff (i.e. he is effectively myopic). In particular, it will turn out that the argument below goes through for $\varepsilon' > 0$ and $\bar{\delta} \in (0, 1]$ such that, for any $\varepsilon \in [0, \varepsilon')$ and any $\delta \in [0, \bar{\delta})$,

$$\varepsilon < \underline{w}_1 - \frac{\delta}{1-\delta} \left[\overline{u}_1 - \underline{u}_1 \right].$$

Note that such ε' and $\bar{\delta}$ do exist.

Consider any pure strategy $s_1 \in S_1$ and any $s_2 \in \tilde{S}_2(s_1)$. Temporarily fix $\eta \in (0, 1)$. Let z be the path of play generated by (s_1, s_2) and suppose that player 1 weakly learns to predict the path of play. Then there is a set $\mathbb{N}^P(\eta, z) \subset \mathbb{N}$ of density 1 such that, for any $n \in \mathbb{N}^P(\eta, z)$, in the continuation game beginning in period $n + 1$, player 1 assigns some probability greater than $1 - \eta$ to the actual action chosen by player 2 in period $n + 1$. Let $\mathbb{N}^\circ(\eta, z) = \mathbb{N}^P(\eta, z) \cap \mathbb{N}^\diamond$, where $\mathbb{N}^\diamond$ is from the definition of $\tilde{S}_2(s_1)$. Note that $\mathbb{N}^\circ(\eta, z)$, as the intersection of two sets of density 1, has density 1.

Choose any $n \in \mathbb{N}^\circ(\eta, z)$. Recalling that $\pi(z, n)$ is the history corresponding to the n period initial segment of z, let $s_1(\pi(z, n)) = a_1^*$ be the action chosen by player 1 in period $n + 1$. Then player 2 chooses action $\tilde{a}_2(a_1^*)$. Discounting payoffs to period $n + 1$, player 1's expected payoff in the continuation game is *at most*

$$(1-\eta) u_1(a_1^*, \tilde{a}_2(a_1^*)) + \eta \overline{u}_1 + \frac{\delta}{1-\delta} \overline{u}_1.$$

If player 1 were instead to choose an action a_1 in period $n+1$ to maximize $u_1(a_1, \tilde{a}_2(a_1^*))$, his expected payoff in the continuation game would be *at least*

$$(1-\eta) \max_{a_1 \in A_1} u_1(a_1, \tilde{a}_2(a_1^*)) + \eta \underline{u}_1 + \frac{\delta}{1-\delta} \underline{u}_1.$$

Thus, uniform ε optimization requires

$$\varepsilon + (1-\eta)u_1(a_1^*, \tilde{a}_2(a_1^*)) + \eta\overline{u}_1 + \frac{\delta}{1-\delta}\overline{u}_1$$
$$\geq (1-\eta)\max_{a_1 \in A_1} u_1(a_1, \tilde{a}_2(a_1^*)) + \eta\underline{u}_1 + \frac{\delta}{1-\delta}\underline{u}_1$$

or

$$\varepsilon + \eta(\overline{u}_1 - \underline{u}_1) \geq \underline{w}_1 - \frac{\delta}{1-\delta}[\overline{u}_1 - \underline{u}_1].$$

By the construction of ε' and $\bar{\delta}$, there is an η sufficiently small such that this inequality cannot hold for any $\varepsilon \in [0, \varepsilon')$ and $\delta \in [0, \bar{\delta})$. This establishes the first part of the lemma.

As for the second part of the lemma, suppose that Assumption MM holds. Fix any $\delta \in [0, 1)$ and choose $\varepsilon' > 0$ such that, for any $\varepsilon \in [0, \varepsilon')$,

$$\varepsilon < \frac{1}{1-\delta}\left[m_1 - M_1\right].$$

By Assumption MM, $m_1 > M_1$, hence such ε' exist.

Once again, consider any pure strategy $s_1 \in S_1$ and any $s_2 \in S_2^M(s_1)$. Let z be the path of play generated by (s_1, s_2). Temporarily fix $\eta > 0$ and an integer $\ell > 0$. Again, suppose that player 1 weakly learns to predict the continuation path of play. Then there is a set $\mathbb{N}^P(\eta, z) \subset \mathbb{N}$ of density 1 such that in the continuation game beginning in period $n+1$, player 1 assigns a probability greater than $1-\eta$ to the actual action chosen by player 2 in period $n+1$. Let $\mathbb{N}^\circ(\eta, z) = \mathbb{N}^P(\eta, z) \cap \mathbb{N}^\diamond$, where $N^\diamond$ is from the definition of $S_2^M(s_1)$. Note that $\mathbb{N}^\circ(\eta, z)$, as the intersection of two sets of density 1, has density 1.

Let $\mathbb{N}^\ell(\eta, z) \subset \mathbb{N}^\circ(\eta, z)$ be such that for any $n \in \mathbb{N}^\ell(\eta, z)$ and any $k \in \{1, \ldots, \ell\}$, player 1 assigns probability greater than $1-\eta$ to the actual action chosen by player 2 in period $n+k$. Because $\mathbb{N}^\circ(\eta, z)$ has density 1, $\mathbb{N}^\ell(\eta, z)$ has density 1 (and, in particular, is not empty). Thus, for any $n \in \mathbb{N}^\ell(\eta, z)$, player 1 assigns a probability greater than $(1-\eta)^\ell$ to the actual ℓ-period continuation history, beginning in period $n+1$. For ease of notation, define $\lambda \in (0, 1)$ by

$$1 - \lambda = (1-\eta)^\ell.$$

Note that λ decreases as either η decreases or ℓ increases.

Consider any $n \in \mathbb{N}^\ell(\eta, z)$. In the ℓ-period continuation history beginning in period $n+1$, player 1 receives at most M_1 per period. On the other hand, player 1 believes that there is a probability of at most λ that the continuation history might be something else. In an alternate ℓ-period continuation history, player 1 receives at most $\overline{u}_1$ per period. Finally, from period $n+\ell+1$ onwards, player 1 receives at most $\overline{u}_1$ per period. Thus, beginning in period $n+1$, player 1 expects to earn *at most*

$$\frac{1-\delta^\ell}{1-\delta}\left[(1-\lambda)M_1+\lambda\overline{u}_1\right]+\frac{\delta^\ell}{1-\delta}\overline{u}_1.$$

In contrast, any best response must expect to earn at least m_1, on average, following any history given positive probability by $\mu_{(s_1,\sigma_2^\beta)}$. (And, by the assumption that player 1 learns to predict the path of play, any history along z is given positive probability.) Thus, under a true best response, player 1 expects to earn *at least*

$$\frac{m_1}{1-\delta}.$$

Thus ε optimization requires

$$\varepsilon+\frac{1-\delta^\ell}{1-\delta}\left[(1-\lambda)M_1+\lambda\overline{u}_1\right]+\frac{\delta^\ell}{1-\delta}\overline{u}_1\geq\frac{m_1}{1-\delta}$$

or

$$\varepsilon\geq(1-\lambda)\frac{1-\delta^\ell}{1-\delta}\left[m_1-M_1\right]-\frac{\delta^\ell+\lambda(1-\delta^\ell)}{1-\delta}[\overline{u}_1-m_1].$$

By the construction of ε', one can find η sufficiently small and ℓ sufficiently large that this inequality cannot hold for any $\varepsilon\in[0,\varepsilon')$. This establishes the second part of the lemma. ■

Proof of Lemma 5. Fix any history h that is reachable under $(\sigma_1,\sigma_2^\beta)$. For concreteness, let the length of h be n. I must show that s_{1h} is $\frac{1}{(1-\delta)\xi}\varepsilon$ optimal.

Following *any* history h, let $J(a_1)$ denote the expected discounted payoff to player 1 from playing action a_1 in the next period and thereafter playing according to σ_1. Let $s_1(h)=a_1^*$. Then, by the construction of s_1, for any h,

$$V_1(\sigma_{1h},\sigma_{2h}^\beta)\leq\xi J(a_1^*)+(1-\xi)\max_{a_1\in A_1}J(a_1).$$

On the other hand, since σ_1 is uniformly ε optimal, for any h that is reachable under $(\sigma_1,\sigma_2^\beta)$,

$$V_1(\sigma_{1h},\sigma_{2h}^\beta)+\varepsilon\geq\max_{\sigma_1'\in\Sigma_1}V_1(\sigma_1',\sigma_{2h}^\beta).$$

Combining these inequalities with the fact that

$$\max_{\sigma_1'\in\Sigma_1}V_1(\sigma_1',\sigma_{2h}^\beta)\geq\max_{a_1\in A_1}J(a_1),$$

one concludes that playing a_1^* next period and thereafter playing according to σ_1 loses, relative to the optimal continuation strategy, at most ε/ξ:

$$\max_{\sigma_1' \in \Sigma_1} V_1(\sigma_1', \sigma_{2h}^{\beta}) - J(a_1^*) \leq \frac{\varepsilon}{\xi}.$$

Fix any integer $N > 0$. Let s_1^N denote the pure strategy that is identical to s_1 on every history of length $n + N - 1$ or less, and thereafter is identical to σ_1.

Consider *any* continuation history h' of length $N - 1$ that is reachable under $(\sigma_{1h}, \sigma_{2h}^{\beta})$. From the argument above, the continuation payoff from playing s_1^N, starting at history $h \cdot h'$, is within ε/ξ of the maximum continuation payoff.

Similarly, for any continuation history h'' of length $N-2$ that is reachable under $(\sigma_{1h}, \sigma_{2h}^{\beta})$, the continuation payoff from playing s_1^{N-1}, starting at history $h \cdot h''$, is within ε/ξ of the maximum continuation payoff. If, instead of playing s_1^{N-1}, one plays s_1^N, as in the previous paragraph, then one loses, at most, an additional $\delta\varepsilon/\xi$, discounted back to date $n + N - 1$. Therefore, for any continuation history h'' of length $N-2$ that is reachable under $(\sigma_{1h}, \sigma_{2h}^{\beta})$, the continuation payoff from playing s_1^N, starting at history $h \cdot h''$, is within

$$\frac{\varepsilon}{\xi} + \delta\frac{\varepsilon}{\xi}$$

of the maximum continuation payoff.

Continuing in this way, it follows that the continuation payoff from playing s_1^N, starting at history h, is within

$$\frac{\varepsilon}{\xi} + \cdots + \delta^{N-1}\frac{\varepsilon}{\xi} = \left(\frac{1-\delta^N}{1-\delta}\right)\frac{\varepsilon}{\xi}$$

of the maximum continuation payoff. This holds for any integer $N > 1$. Therefore, taking the limit as $N \to \infty$ (and noting that s_1^N converges to s_1 in the product topology, and that V_1 is continuous in the product topology), the continuation payoff from playing s_1, starting at history h, is within

$$\frac{\varepsilon}{(1-\delta)\xi}$$

of the maximum continuation payoff, as was to be shown. ■

References

Binmore, K. G. (1987): "Modeling Rational Players, Part I," *Economics and Philosophy*, 3, 179–214.

Dawid, A. P. (1985): "The Impossibility of Inductive Inference," *Journal of the American Statistical Association*, 80(390), 340–341.

Foster, D., and P. Young (2003): "Learning, Hypothesis Testing, and Nash Equilibrium," *Games and Economic Behavior*, 45, 73–96.

Fudenberg, D., and D. Levine (1998): *Theory of Learning in Games.* MIT Press, Cambridge, MA.

Harsanyi, J. (1973): "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Eequilibrium Points," *International Journal of Game Theory*, 2, 1–23.

Jackson, M., and E. Kalai (1997): "Social Learning in Recurring Games," *Games and Economic Behavior*, 21, 102–134.

Jackson, M., E. Kalai, and R. Smorodinsky (1999): "Bayesian Representation of Stochastic Processes under Learning: de Finetti Revisited," *Econometrica*, 67(4), 875–893.

Jordan, J. S. (1991): "Bayesian Learning in Normal Form Games," *Games and Economic Behavior*, 3, 60–81.

——— (1995): "Bayesian Learning in Repeated Games," *Games and Economic Behavior*, 9, 8–20.

Kalai, E., and E. Lehrer (1993): "Subjective Equilibrium in Repeated Games," *Econometrica*, 61(5), 1231–1240.

——— (1994): "Weak and Strong Merging of Opinions," *Journal of Mathematical Economics*, 23, 73–86.

Lehrer, E. (1996): "Fictitious Play Fits the Subjective Learning Approach," Tel Aviv.

Lehrer, E., and R. Smorodinsky (1996): "Compatible Measures and Merging," *Mathematics of Operations Research*, 21(3), 697–706.

Lehrer, E., and S. Sorin (1998): "ε-Consistent Equilibrium In Repeated Games," *International Journal of Game Theory*, 27, 231–244.

Nachbar, J. H. (1997): "Prediction, Optimization, and Learning in Repeated Games," *Econometrica*, 65, 275–309.

——— (2001): "Bayesian Learning in Repeated Games of Incomplete Information," *Social Choice and Welfare*, 18(2), 303–326.

Nyarko, Y. (1998): "Bayesian Learning and Convergence to Nash Equilibria Without Common Priors," *Economic Theory*, 11(3), 643–655.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.