JOURNAL OF ECONOMIC THEORY **73**, 438–452 (1997)
ARTICLE NO. ET962245

Belief Affirming in Learning Processes*

Dov Monderer†

Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, Haifa, Israel

Dov Samet‡

Faculty of Management, Tel Aviv University, Tel Aviv, Israel

and

Aner Sela

Department of Economics, Ben-Gurion University, Beer Sheva, Israel

Received January 31, 1995; revised July 6, 1996

A learning process is belief affirming if the difference between a player's expected payoff in the next period, and the average of his or her past payoffs converges to zero. We show that every smooth discrete fictitious play and every continuous fictitious play is belief affirming. We also provide conditions under which general averaging processes are belief affirming. *Journal of Economic Literature* Classification Numbers: C72, C73, D83. © 1997 Academic Press

1. INTRODUCTION

Consider players engaged in the repeated play of a finite game in strategic form. In each period, each player, on the basis of the history of past moves, forms a belief about the next joint move of the other players. She then chooses an action that is a myopic pure best reply, that is, an

* We thank Drew Fudenberg for very helpful discussions and suggestions. This work was supported by the Fund for the Promotion of Research in the Technion.

† E-mail: dov@ie.technion.ac.il.
‡ E-mail: samet@vm.tau.ac.il.

0022-0531/97 $25.00

action that maximizes her payoff in the next period according to her belief. We call such a process a *learning process*.[1]

Learning processes are models of bounded rationality. The participants in such processes think myopically and nonstrategically. Moreover, full statistical analysis would prove their beliefs wrong, if they were to use one. Yet blatant statistical inconsistency of beliefs should not be expected even from players who are rationally bounded. Several conditions can be imposed on learning processes that prevent them from being overly inconsistent. The oldest consistency condition is the convergence of the process, namely the convergence of players' beliefs concerning their rivals' strategies. Foster and Vohra [4] suggested calibration of the forecasts, in a learning process, as a consistency condition (see Dawid [2]) and showed that in calibrated processes the joint distribution of players' actions converge to the set of correlated equilibria. Another consistency condition was suggested by Milgrom and Roberts [9]. They defined and studied adaptive learning process in which each player assigns a low probability to strategy profiles that have not been used by the other players for many periods. All these conditions are defined purely in terms of strategies.

In this paper we study a consistency condition, called *belief affirming*, which relates beliefs about strategies to payoffs. The essence of this condition is as follows. At each point in time, t, a player's belief and action determine her expected payoff, $E(t)$, in the next period. At the same time, the player also observes her payoff history. If in the long run, these histories seem to contradict her expected payoff then her confidence in her beliefs will be shaken. If on the contrary, past and future payoffs fit, then the process affirms the player's beliefs. This belief affirming, in terms of payoffs, is the condition we study here.

The simplest way to examine the expected payoff, $E(t)$, in light of payoff history is to compute the average payoff, $A(t)$, up to time t and compare it with $E(t)$. We say that the learning process is *belief affirming* (with respect to average payoff) if $\lim_{t\to\infty}(E(t)-A(t))=0$, for each player.[2]

The results of this paper indicate that belief affirming in fictitious play and even more general learning processes is the rule rather than the exception. We consider first, in Section 2, the classical fictitious play in which time is discrete. Most of our results are obtained for a player who is engaged individually in a fictitious play. That is, the player chooses in each

[1] Such processes are to be distinguished from evolutionary and psychological models of learning in which agents do not form beliefs about future moves, but rather choose their actions according to previous payoffs associated with the actions. See e.g., Roth and Erev [14], and the references listed there, for a discussion of a variety of such processes.

[2] Fudenberg and Levine [6] define "universal consistency," which is the same as belief affirming but applies to a slightly more general environment.

period a best response against the average action of the past, while other players are not restricted in the way they play.

Our first key result, The Rosy Theorem, is a simple and potent observation. It states that a player engaged in a fictitious play is always optimistic and expects to be paid no less than what he has been paid on the average in the past. That is, for each t, $E(t) \geqslant A(t)$. This simple theorem is enough to guarantee that, in zero-sum two-player games, fictitious play is always belief affirming. The convergence of fictitious play in such games was proved by Robinson [13]. We show, moreover, that the average payoff converges to the value of the game.

For non-zero-sum games belief affirming of fictitious play does not generally hold. We prove, though, that *any smooth fictitious play is belief affirming*. A path is smooth for a player if the proportion of time, up to time t, she switches from one action to another, converges to zero when t approaches infinity.[3] This theorem fails without the smoothness condition, as the first example in Section 4 demonstrates. Though we do not have a formal characterization of smooth processes, it seems that generically fictitious play processes are smooth. This can be easily shown for 2×2 games.

The Rosy Theorem turns out to be an efficient tool in proving non-convergence of fictitious play. We demonstrate it by providing a very short proof to the non-convergence of the fictitious play in Shapley's [15] famous example.

In Section 3 we introduce continuous fictitious play in which time flows continuously. For these processes, we have an unqualified result. *Continuous fictitious plays are always belief affirming*. Continuous fictitious plays share many properties with discrete ones. The main difference is that in the continuous process, all switches are on a tie, that is, a player who switches is indifferent between the two actions. If it happens in a discrete fictitious play that all switches are on ties then there is a corresponding continuous process that is completely analogous to the discrete one.[4]

In Section 4 we study the relation between belief affirming and the convergence of fictitious play, i.e., convergence of beliefs. We show that neither condition implies the other. A very simple two-person game demonstrates how a converging process fails to be belief affirming. Shapley's example of fictitious play that does not converge shows the opposite, since the process in his example is smooth and therefore, by our result, is belief affirming. If however, a two player fictitious play is belief affirming and it does converge, then average payoffs converge to the equilibrium payoffs.

[3] This theorem was independently proved by Fudenberg and Levine [14].

[4] See Monderer and Sela [11] for a discussion of the sensitivity of fictitious play to tie breaking rules.

In Section 5 we explore generalizations of continuous fictitious play. We allow players to average action history nonuniformly using general averaging functions. Similarly, we generalize the notion of belief affirming such that expected payoff is compared with average payoff where the average is computed using the same averaging function with which beliefs about actions were formed. We then show that if the following two conditions hold then the process is belief affirming. First, enough weight must be given to recent history (we call such a process non-nostalgic). Second, averaging in each point of time is derived from an all-times-embracing averaging function (a condition we call time-consistency).

In this paper we adopt the approach that player i's belief is a probability distribution over the next joint move of the other players. In particular, our definition of fictitious play for more than two players coincides with the one given in Fudenberg and Kreps [5], and not with Monderer and Shapley [12], where player's beliefs about different players are independent. Our theorems do not seem to apply to the latter process.

2. BELIEF AFFIRMING IN DISCRETE PROCESSES

Let $N = \{1, 2, ..., n\}$ be the set of players. For each $i \in N$, S^i is a finite strategy set of Player i. Let $S = \times_{i \in N} S^i$, and for each i, denote $S^{-i} = \times_{j \in N \setminus \{i\}} S^j$. Let $u^i \colon S \to R$ be i's payoff function. For each finite set A we denote by $\Delta(A)$ the set of probability measures over A. The set of Player i's mixed strategies $\Delta(S^i)$ is denoted by Δ^i, and the set of i's beliefs, concerning her rivals' joint moves, $\Delta(S^{-i})$, is denoted by Δ^{-i}. For each $s^i \in S^i$ and $b^{-i} \in \Delta^{-i}$ we denote by $U^i(s^i, b^{-i})$ Player i's expected payoff, according to her belief b^{-i}, when she plays s^i.

A *path* in S is a sequence $\sigma = \{\sigma(t)\}$, for $t = 0, 1, 2, ...$ of elements in S. We think of $\sigma^i(t)$ and $\sigma^{-i}(t)$ as being (extreme) points in Δ^i and Δ^{-i} correspondingly. Thus, these points being elements of a linear space, we can average pure strategies or beliefs of the same player.

A *belief sequence for player i* is a sequence, $b^{-i} = (b^{-i}(t))$, for $t \geqslant 1$, of elements in Δ^{-i}.

A *belief sequence*, is a vector $b = (b^{-1}, b^{-2}, ..., b^{-n})$ of belief sequences for all players.

A *myopic learning process for player i* (in short, *a learning process for player i*) is a pair (σ, b^{-i}) where σ is a path in S, and b^{-i} is a belief sequence for player i such that for every $t \geqslant 1$, the strategy $\sigma^i(t)$ is a best reply according to $b^{-i}(t)$. A *myopic learning process* (in short a *learning process*) is a pair (σ, b) such that (σ, b^{-i}) is a learning process for player i for every i.

A learning process for player i, (σ, b^{-i}) is a *fictitious play for i*, if

$$b^{-i}(t) = \frac{1}{t} \sum_{\tau=0}^{t-1} \sigma^{-i}(\tau).$$

A learning process is a *fictitious play*, if it is a fictitious play for every player.

Given any learning process for player i, (σ, b^{-i}), we denote by $E^i(t)$, for $t \geqslant 1$, i's expected payoff at period t. That is,

$$E^i(t) = U^i(\sigma^i(t), b^{-i}(t)).$$

Note that if (σ, b^{-i}) is a fictitious play for i then by the linearity of expectation

$$E^i(t) = U^i\left(\sigma^i(t), \frac{1}{t} \sum_{\tau=0}^{t-1} \sigma^{-i}(\tau)\right)$$
$$= \frac{1}{t} \sum_{\tau=0}^{t-1} u^i(\sigma^i(t), \sigma^{-i}(\tau)).$$

We denote by $A^i(t)$, i's average payoff up to time t. That is,

$$A^i(t) = \frac{1}{t} \sum_{\tau=0}^{t-1} u^i(\sigma(\tau)).$$

We say that the learning process for i, (σ, b^{-i}), is *belief affirming for i* (with respect to average payoff) if

$$\lim_{t \to \infty} (E^i(t) - A^i(t)) = 0. \tag{1}$$

We say that a learning process is *belief affirming* if it is belief affirming for every player.

To study belief affirming in fictitious plays we show first that "fictitious players" are chronically optimistic and always expect to receive in the next period more than experience has taught them. We use this property in studying belief affirming in Theorems B and C, as well as in providing a straightforward proof for the nonconvergence of the fictitious play in Shapley's example.

THEOREM A (The Rosy Theorem). *For each player i and each fictitious play for i*

$$E^i(t) \geqslant A^i(t) \qquad \textit{for every} \quad t \geqslant 1. \tag{2}$$

Proof. Note that for every $t \geqslant 1$, and $s^i \in S^i$

$$\sum_{\tau=0}^{t-1} u^i(\sigma^i(t), \sigma^{-i}(\tau)) \geqslant \sum_{\tau=0}^{t-1} u^i(s^i, \sigma^{-i}(\tau)). \tag{3}$$

Denote $a_{t,\tau} = u^i(\sigma^i(t), \sigma^{-i}(\tau))$. Denote also by α_t the left hand side of (3). Substituting $\sigma^i(t-1)$ for s^i in the right-hand side of (3), we can write it as $\alpha_{t-1} + a_{t-1,t-1}$. Thus (3) is: $\alpha_t \geqslant \alpha_{t-1} + a_{t-1,t-1}$, for each $t \geqslant 1$. Therefore, $\alpha_t \geqslant \sum_{\tau=0}^{t-1} a_{\tau,\tau}$. Dividing this inequality by t yields,

$$E^i(t) \geqslant A^i(t). \quad \blacksquare$$

While optimism is a good recipe for life, it is unrealistic when exaggerated. Someone who expects every day to gain millions of dollars, and finds at the end of the day that he has got nothing, is certainly unrealistic and can hardly be thought of as someone who learns from past experience. Belief affirming is the right balance for the ever-optimistic fictitious player. Such a player when engaged in a belief affirming play, is in the long run realistic.

In Theorem C we give a condition under which fictitious players are indeed realistic. In the next section we show that without this condition fictitious players can be unrealistically optimistic and the play may fail to be belief affirming. However, using the Rosy Theorem, we show now that for two-person zero-sum games no condition is required—fictitious plays in such games are always belief affirming. The intuition here is quite simple. By the Rosy Theorem both players are optimistic. But in a zero-sum game, the optimism of one player is the pessimism of the other. Hence both players are pessimistic as well. Being both optimistic and pessimistic boils down to being realistic.

In particular, this result implies a hitherto unknown fact, that in a two-person zero-sum games the average payoff of player 1 converges to the value of the game.

Theorem B. *Every fictitious play in a two-person zero-sum game is belief affirming. Moreover,* $\lim_{t\to\infty} A(t) = v$, *where* $A(t)$ *is the average payoff to the first player and* v *is the value of the game.*

Proof. Let $N = \{1, 2\}$ be the set of players. By Theorem A, $E_1(t) \geqslant A_1(t)$ and $E_2(t) \geqslant A_2(t) = -A_1(t)$ for each t. By Robinson's Theorem, $\lim_{t\to\infty} E_1(t) = v$ and $\lim_{t\to\infty} E_2(t) = -v$. Thus $\limsup_{t\to\infty} A_1(t) \leqslant v$ and $\limsup_{t\to\infty} A_2(t) \leqslant -v$. As $\liminf_{t\to\infty} A_1(t) = -\limsup_{t\to\infty} A_2(t)$, it follows that $\lim_{t\to\infty} A_1(t) = v$. $\blacksquare$

We next give a sufficient condition that a fictitious play for a player is belief affirming for her. The same condition was independently proved by Fudenberg and Levine (1994).

A path σ is (*asymptotically*) *smooth for player i* if $\lim_{t\to\infty}(1/t)\, M^i(t)=0$, where $M^i(t)$ is the number of periods k, such that $1\leqslant k\leqslant t$, and $\sigma^i(k)\neq\sigma^i(k-1)$. A path is *smooth* if it is smooth for every player.

THEOREM C. *Every smooth fictitious play for a player is belief affirming for the player.*

Proof. Substitute $\sigma^i(t+1)$ for s^i in (3). Then the right-hand side of (3) is $\alpha_{t+1}-a_{t+1,t}$. Hence (3) is: $\alpha_t\geqslant\alpha_{t+1}-a_{t+1,t}$, or $\alpha_{t+1}\leqslant\alpha_t+a_{t+1,t}$ for each $t\geqslant 1$. Therefore,

$$\begin{aligned}\alpha_t &\leqslant \sum_{\tau=0}^{t-1} a_{\tau+1,\tau}\\ &= \sum_{\tau=0}^{t-1} a_{\tau,\tau}+\sum_{\tau=0}^{t-1}(a_{\tau+1,\tau}-a_{\tau,\tau}).\end{aligned}$$

Dividing this inequality by t yields,

$$E^i(t)\leqslant A^i(t)+\frac{1}{t}\sum_{\tau=0}^{t-1}(a_{\tau+1,\tau}-a_{\tau,\tau}) \tag{4}$$

but the last term is just $(1/t)\sum_{\tau=0}^{t-1}(u^i(\sigma^i(\tau+1),\sigma^{-i}(\tau))-u^i(\sigma(\tau)))$, which by the smoothness assumption converges to 0. Hence,the required equality follows from Theorem A and (4). ∎

Finally, using the Rosy Theorem, we give a simple proof for the nonconvergence of the belief sequence in Shapley's example [15].[5]

Consider a typical symmetric Shapley's example

$$\begin{pmatrix}(0,0) & (x,y) & (y,x)\\ (y,x) & (0,0) & (x,y)\\ (x,y) & (y,x) & (0,0)\end{pmatrix}$$

where $x>y>0$. According to the improvement principle of Monderer and Sela [10], if the initial move of the players is one of the non-zero entries of the bimatrix, then the process runs through the unique better reply cycle defined by Shapley and in particular, the players' payoff vector at each t is either (x,y) or (y,x). If the process converges, then it must converge to

[5] Other proofs, based on different ideas are given by Deschamps [3] and by Monderer and Sela [10].

the unique equilibrium of this game where each player uses each pure strategy with equal probability of $\frac{1}{3}$. Hence, the expected payoff $E^i(t)$ of each player i must converge to the equilibrium payoff $(x+y)/3$. By Theorem A, the average payoff of each player i, $A^i(t)$, cannot exceed her expected payoff $E^i(t)$. Hence for $i=1, 2$:

$$\limsup_{t\to\infty} A^i(t) \leqslant \frac{x+y}{3}.$$

But for each t, $A^1(t)+A^2(t)=x+y$. Thus we obtain the contradictory inequality:

$$x+y \leqslant \limsup_{t\to\infty} A^1(t) + \limsup_{t\to\infty} A^2(t) \leqslant \frac{2(x+y)}{3}.$$

3. BELIEF AFFIRMING IN CONTINUOUS PROCESSES

Lately the literature on fictitious play focuses on the continuous process that was defined by Brown [1]. In some cases such processes are more "natural." For example, the convergence of fictitious play in zero-sum games, which requires an intricate proof for the discrete process (Robinson [13]) can be proved easily for the continuous case (Hofbauer [7]). Also, results of Krishna and Sjöström [8] concerning the cyclical structure of a converging continuous fictitious play seem to be difficult to prove for the discrete process. There are no formal results relating discrete and continuous processes, though it seems that whenever the continuous fictitious play exists, the discrete process path behaves similarly to the continuous one.

In this section we show that in the continuous framework, fictitious plays are always belief affirming and, unlike Theorem C, the smoothness assumption is not required. We also show that players cannot deviate from optimism by more than c/t for a positive constant c.

In defining the continuous version of learning process and fictitious play, we skip, for simplicity, the individual versions ("... for player i") and define it directly for all the players. However, all theorems can be stated and proved in individualistic terms as in Section 2.

A *path* σ in S is a right continuous function σ: $[0, \infty) \to S$ such that the discontinuity points of σ have no (finite) accumulation points. A *belief path* is a pair (b, t_1) where $t_1 > 0$, and b is a function b: $[t_1, \infty) \to \times_{i\in N} \Delta^{-i}$. That is, $b(t) = (b^{-1}(t), b^{-2}(t), ..., b^{-n}(t))$, for $t \geqslant t_1$.

A *continuous learning process* is a pair $(\sigma, (b, t_1))$ where σ is a path in S and (b, t_1) is a belief path such that $\sigma^i(t)$ is a best reply according to $b^{-i}(t)$ for each $t \geqslant t_1$ and for every player i.

A continuous learning process $(\sigma, (b, t_1))$ is a *continuous fictitious play* if for every player i

$$b^{-i}(t) = \frac{1}{t} \int_0^t \sigma^{-i}(\tau)\, d\tau,$$

for $t \geqslant t_1$. Note that $b^{-i}(t)$ satisfies the differential inclusion

$$\frac{\partial b^{-i}(t)}{\partial t} \in \frac{1}{t} (\times_{j \neq i} BR^j(b^{-j}(t)) - b^{-i}(t)),$$

where BR^j is j's best response correspondence, and $\partial/\partial t$ is the right derivative. This differential inclusion can be used to define continuous fictitious play.

Given any continuous learning process $(\sigma, (b, t_1))$, we denote by $E^i(t)$, for $t \geqslant t_1$, i's expected payoff at period t. That is,

$$E^i(t) = U^i(\sigma^i(t), b^{-i}(t)).$$

We denote by $A^i(t)$, i's average payoff up to time t. That is,

$$A^i(t) = \frac{1}{t} \int_0^t u^i(\sigma(\tau))\, d\tau.$$

Note that if $(\sigma, (b, t_1))$ is a fictitious play, then by the linearity of expectation,

$$E^i(t) = U^i\left(\sigma^i(t), \frac{1}{t} \int_0^t \sigma^{-i}(\tau)\, d\tau\right)$$

$$= \frac{1}{t} \int_0^t u^i(\sigma^i(t), \sigma^{-i}(\tau))\, d\tau.$$

As in the discrete case, we say that a continuous learning path $(\sigma, (b, t_1))$ is *belief affirming* if for every i

$$\lim_{t \to \infty} (E^i(t) - A^i(t)) = 0.$$

THEOREM D. *Every continuous fictitious play is belief affirming.*

Proof. First we note that the function $tE^i(t)$ is continuous in the interval $[t_1, \infty]$. To see this, define for each $s^i \in S^i$ a function $u^{s^i}(t) = \int_0^t u^i(s^i, \sigma^{-i}(\tau))\, d\tau$. Clearly the functions u^{s^i} are continuous, and for each $t \geqslant t_1$, $tE^i(t) = \max\{u^{s^i}(t) \mid s^i \in S^i\}$. Hence $tE^i(t)$ is continuous in $[t_1, \infty]$.

Moreover, if σ is continuous at t then $\sigma^i(t)$ is fixed in a neighborhood of t and thus the derivative of $tE^i(t)$ at t is just the integrand at this point, namely, $u^i(\sigma(t))$. But this is precisely the derivative of $tA^i(t)$ at t. Thus $tE^i(t) - tA^i(t)$ is constant in each interval in which σ is continuous. But $tE^i(t) - tA^i(t)$ is continuous and therefore

$$tE^i(t) - tA^i(t) = C,$$

for all $t \geqslant t_1$, where $C = t_1 E^i(t_1) - t_1 A^i(t_1)$. Dividing by t yields the desired result. ∎

COROLLARY 1 (The Rosy Theorem for Continuous Processes). *Let* $(\sigma, (b, t_1))$ *be a continuous fictitious play. Then, there exists a constant* $c > 0$ *such that for every player* i

$$E^i(t) - A^i(t) \geqslant -\frac{c}{t} \qquad \textit{for every} \quad t \geqslant t_1.$$

Proof. By the proof of Theorem D, this corollary holds with

$$c = \max\{t_1 \,|(E^i(t_1) - A^i(t_1))|: 1 \leqslant i \leqslant n\}. \quad ∎$$

4. CONVERGENCE OF BELIEFS AND BELIEF AFFIRMING

Neither of the two properties of learning processes, convergence of beliefs and belief affirming, implies the other. First we show a converging discrete fictitious play which is not belief affirming.

Consider the two-person game where each player can choose α or β. If they coordinate on the same action then each gets 1, if they fail to coordinate, each gets 0. If players fail to coordinate in the first period and play (α, β) then a fictitious play path is:

$$(\alpha, \beta)(\beta, \alpha)(\beta, \alpha)(\alpha, \beta)(\alpha, \beta)(\beta, \alpha)(\beta, \alpha) \cdots.$$

Players' beliefs, $(b^{-1}(t), b^{-2}(t))$, converge in this process to the mixed strategy equilibrium $(0.5, 0.5)$, $(0.5, 0.5)$. Thus each player i's expected payoff, $E^i(t)$, converges to her equilibrium payoff 0.5. But the players fail to ever coordinate and their average payoff $A^i(t)$ is 0 for all t.

Observe that this process is not smooth which shows that without the smoothness condition, Theorem C is not true. Note also that there is no continuous fictitious play in which players play (α, β) at time 0. The anomaly of this example can be avoided if we amend slightly the definition of fictitious play to allow players beliefs to be constructed as if at time 0 players played a mixed strategy. It can easily be seen that in such processes, for almost all initial beliefs, the resulting discrete fictitious play is belief affirming and continuous fictitious play (which is always belief affirming) does exist.

Shapley's famous example (see Section 3) of nonconverging fictitious play demonstrates that the condition of belief affirming does not imply convergence of the process. The process in this example is smooth since the time between consecutive changes of strategies grows exponentially. Hence the process is belief affirming by Theorem C.

When a process is both belief affirming and converging then it also enjoys the convergence properties of payoffs. Recall that if the process of a two-person game converges, then there exists an equilibrium s such that for each player i, her beliefs, $b^{-i}(t)$, converge to s^i. Thus, $E^i(t)$ converges to $U^i(s)$, which is i's equilibrium payoff. Hence, by belief affirming, $A^i(t)$ converges to i's equilibrium payoff. Thus it follows immediately from Theorem C.

COROLLARY 2. *If a smooth fictitious play of two-person game converges, then for each player i, the average payoff $A^i(t)$ converges to i's payoff in the equilibrium to which the process converges.*

Remark. In a two-person game, a fictitious play *approaches equilibrium*, if every limit point of the belief sequence b, is an equilibrium. It is not known whether Corollary 2 continues to hold if the convergence assumption is replaced by the assumption of approaching equilibria.

5. GENERAL AVERAGING PROCESSES

In this section we consider continuous learning processes which we call averaging processes. As in fictitious play, players' beliefs are generated by averaging history, except that now we allow very general schemes of averaging, other than the uniform one used in fictitious play. We study conditions on general averaging processes that guarantee belief affirming. As in Section 4, we (unnecessarily) assume that *all* playersare learning.

A function f: $\{(t, \tau) | 0 \leqslant \tau \leqslant t, t > 0\} \to R_+$ is an *averaging function* if it satisfies the following conditions for all t and τ in its domain

(1) $f(t, \tau) \geqslant 0$.

(2) for each $t > 0$, $f(t, \cdot)$ is continuous in the interval $[0, t]$, and $\int_0^t f(t, \tau)\, d\tau = 1$.

That is, for each t, $f(t, \cdot)$ is a continuous density function on the interval $[0, t]$.

An *averaging process* is a triple $(\sigma, (b, t_1), \phi)$, where $(\sigma, (b, t_1))$ is a continuous learning process, $\phi = (f^1, ..., f^n)$ is a vector of averaging functions—one for each player—and for each player i and time $t \geqslant t_1$,

$$b^{-i}(t) = \int_0^t \sigma^{-i}(\tau)\, f^i(t, \tau)\, d\tau.$$

Clearly a continuous fictitious play is an averaging process where $f(t, \tau) = 1/t$ for $0 \leqslant \tau \leqslant t$.

Player i's expected payoff at time t is

$$\begin{aligned} E^i(t) &= U^i(\sigma^i(t), b^{-i}(t)) \\ &= \int_0^t u^i(\sigma^i(t), \sigma^{-i}(\tau))\, f^i(t, \tau)\, d\tau. \end{aligned}$$

In a general averaging process the "right" way for a player to summarize her payoff history is by using her averaging function. Thus we define Player i's *average payoff* at time t as

$$W^i(t) = \int_0^t u^i(\sigma(\tau))\, f^i(t, \tau)\, d\tau.$$

An averaging process is *belief affirming* (with respect to the given averaging functions) if $\lim_{t \to \infty} (E^i(t) - W^i(t)) = 0$.

Continuous fictitious play, viewed as an averaging process, are indeed belief affirming, according to this definition, since for fictitious play $W^i = A^i$. But all general averaging processes are not belief affirming as the examples, following Theorem C, show.

We study now properties of averaging processes that guarantee that the process is belief affirming. The value $f(t, t)$ of an averaging function reflects the weight given to recent history at time t. (Here we use the continuity of $f(t, \cdot)$ in $[0, t]$. If f were only an L_1 function, then its value at one point would not be well defined.) We refer to putting much weight on past history and little weight on recent history as nostalgia. Therefore a non-nostalgic player gives a considerable weight to recent history, which reflects in high values of $f(t, t)$. We define an averaging function f as *non-nostalgic* if $\int_d^\infty f(\tau, \tau)\, d\tau = \infty$, for some $d > 0$ (or equivalently for all $d > 0$).[6]

Time consistency of averaging function requires that averages of history made at different points in time are all tied together by an overall view of time. Formally, we say that an averaging function f is *time-consistent* if

[6] A condition in this spirit was used by Thorlund-Petersen [16].

there exists a continuous function $F\colon R_+ \to R_+$, such that $\int_0^t F(x)\,dx > 0$, for each $t > 0$, and for $0 \leqslant \tau \leqslant t$,

$$f(t, \tau) = \frac{F(\tau)}{\int_0^t F(x)\,dx}.$$

An averaging process $(\sigma, (b, t_1), \phi)$, is time consistent and non-nostalgic if for each i, f^i has these properties.

THEOREM D. *A non-nostalgic, time-consistent averaging process is belief affirming.*

Proof. Observe that by definition, if an averaging function f is time consistent, then $f(\cdot, \tau)$ is continuous in $[\tau, \infty)$. It follows that the average payoff function W^i is continuous. The same proof as in Theorem C shows that the expected payoff function E^i is also continuous for $t \geqslant t_1$. Note that for each i, $t > 0$ and $0 \leqslant \tau \leqslant t$, the function f^i is differentiable at (t, τ) with respect to t, and

$$\frac{\partial f^i}{\partial t}(t, \tau) = -f^i(t, t)\, f^i(t, \tau).$$

Using this equality and differentiating W^i we obtain

$$\begin{aligned} \frac{dW^i}{dt}(t) &= u^i(\sigma(t))\, f^i(t, t) - \int_0^t u^i(\sigma^i(\tau), \sigma^{-i}(\tau))\, \frac{\partial f^i}{\partial t}(t, \tau)\, d\tau \\ &= u^i(\sigma(t))\, f^i(t, t) - f^i(t, t)\, W^i(t). \end{aligned}$$

Similarly, differentiating E^i at a point t in which σ is continuous (and hence σ^i is constant in a neighborhood of t), we obtain

$$\begin{aligned} \frac{dE^i}{dt}(t) &= u^i(\sigma(t))\, f^i(t, t) - \int_0^t u^i(\sigma^i(t), \sigma^{-i}(\tau))\, \frac{\partial f^i}{\partial t}(t, \tau)\, d\tau \\ &= u^i(\sigma(t))\, f^i(t, t) - f^i(t, t)\, E^i(t). \end{aligned}$$

Thus for each interval in which σ is continuous

$$\frac{d(E^i - W^i)}{dt}(t) = -f^i(t, t)(E^i - W^i)(t).$$

Choose $d > 0$. Then, for each such an interval there exists a constant C such that

$$E^i(t) - W^i(t) = C \exp\left(-\int_d^t f^i(\tau, \tau)\, d\tau \right). \tag{5}$$

Since $E^i - W^i$ is continuous for all $t \geqslant t_1$, then (5) holds with the same constant C for all $t \geqslant t_1$. The theorem follows since the process is non-nostalgic. ▮

In the following two examples we show that neither of the two conditions in Theorem D, can be omitted. Note that the averaging function $f(t, \tau)$ in both example is not continuous in one point in the interval $[0, t]$ and hence does not satisfy our definition of averaging function. It is easy, though , to see how these functions can be slightly changed to make them continuous while keeping the main feature of the examples, namely the lack of belief affirming.

Example 1. Consider the zero-sum game "matching pennies." Each player has two strategies H and T. Players' payoffs are: $u^1(H, H) = u^1(T, T) = 1$, $u^1(H, T) = u^1(T, H) = -1$, and $u^2 = -u^1$. Let σ be the path in which players play successively

$$(H, H), (H, T), (T, T), (T, H), (H, H) \cdots,$$

where each pair of of strategies is played for interval of time of length 1, which we call a round. That is, (H, H) is played in $[0, 1)$, (H, T) in $[1, 2)$ and so on. Both players use the same averaging function f, according to which they average history uniformly, starting with the beginning of the previous round. Thus, for $n = 1, 2, \ldots$, and $n \leqslant t < n + 1$,

$$f(t, \tau) = \frac{1}{t - (n-1)}$$

for $n - 1 \leqslant \tau \leqslant t$.

It is easy to see that when players form their beliefs according to f, starting at $t_1 = 1$, then the process is a learning process, that is, the strategies they choose in σ are indeed best responses according to their beliefs.

If a player changes her strategy at $t = n$, then her rival keeps the same strategy, say X, in $[n-1, n+1)$, and hence the player believes with probability 1 that X is her rival's strategy. Therefore her expected payoff is 1 at all times in $[n, n+1)$. But this player's payoff drops from 1 at $t = n$ to 0, when t approaches $n+1$. Therefore the difference between her expected payoff and her average payoff has no limit, and the process is not belief affirming. The averaging function f is non-nostalgic but is not time-consistent.

EXAMPLE 2. Consider again the game "matching pennies" with the path

$$(H, H), (H, T), (H, T), (H, T), \ldots,$$

where each pair of strategies is played in a round of length 1. The players use the same averaging function $f(t, \tau) = 1$ for $0 \leqslant \tau \leqslant 1$ when $t \geqslant 1$. With $t_1 = 1$, the described process is a learning one. At each time $t \geqslant 1$, player 2 expects to receive 1, while her average payoff is -1. Hence the process is not belief affirming. The averaging function f is time consistent but is not non-nostalgic.

REFERENCES

1. G. W. Brown, "Iterative Solution of Games by Fictitious Play," Activity analysis of production and allocation, Wiley, New York, 1951.
2. A. P. Dawid, The well calibrated Bayesian, *J. Amer. Statist. Assoc.* **77** (1982), 605–613.
3. R. Deschamps, Ph.D. thesis, University of Louvain, 1973.
4. D. P. Foster and R. V. Vohra, Calibrated learning and correlated equilibrium, mimeo, 1993.
5. D. Fudenberg and D. Kreps, Learning mixed equilibria, *Games Econ. Behav.* **5** (1993), 320–367.
6. D. Fudenberg and D. Levine, Consistency and cautious fictitious play, mimeo, 1994; *J. Econ. Dynam. Control*, in press.
7. J. Hofbauer, Stability for the best response dynamics, mimeo, 1994.
8. V. Krishna and T. Sjöström, On the convergence of fictitious play, mimeo, 1995.
9. P. Milgrom and J. Roberts, Adaptive and sophisticated learning in normal form games, *Games Econ. Behav.* **3** (1991), 82–100.
10. D. Monderer and A. Sela, Fictitious play and no-cycling conditions, mimeo, 1993.
11. D. Monderer and A. Sela, A 2×2 game without the fictitious play property, *Games Econ. Behav.* **14** (1996), 144–148.
12. D. Monderer and L. S. Shapley, Fictitious play property for games with identical interests, *J. Econ. Theory* **1** (1996), 258–265.
13. J. Robinson, An iterative method of solving a game, *Ann. Math.* **54** (1951), 296–301.
14. A. E. Roth and I. Erev, Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term, *Games Econ. Behav.* **8** (1995), 164–212.
15. L. S. Shapley, Some topics in two–person games, *in* "Advances in Game Theory" (M. Dresher, L. S. Shapley, and A. W. Tucker, Eds.), pp. 1–29, Princeton Univ. Press, Princeton, NJ, 1964.
16. L. Thorlund-Petersen, Iterative computation of Cournot equilibrium, *Games Econ. Behav.* **2** (1990), 61–95.