Discussion Paper No. 926

# PRIVATE-BELIEFS EQUILIBRIUM

by

Ehud Kalai*

and

Ehud Lehrer*

January 1991

---

*Department of Managerial Economics and Decision Sciences, J. L. Kellogg Graduate School of Management and Department of Mathematics, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208. This research was partly supported by grants nos. SES-9011790 and SES-9022305 from the National Science Foundation, Economics program.

[L6-6]

# PRIVATE-BELIEFS EQUILIBRIUM

by

Ehud Kalai and Ehud Lehrer

## Abstract

At a private-beliefs equilibrium of an n-person infinitely repeated game with discounting, each player maximizes his expected payoff relative to some private, possibly false, belief regarding the strategies chosen by his opponents. Moreover, the probability distribution induced over the observed play paths of the game according to his belief coincides with the one actually played. Thus, any statistical updating can only reinforce the beliefs. It is shown that if the game is played with perfect monitoring, then the joint behavior induced by a private-beliefs equilibrium coincides with a behavior induced by a Nash equilibrium even when perturbations are allowed.

## 1. Introduction

At a private-beliefs equilibrium of an infinitely repeated game, each player has a belief, possibly a false one, regarding the strategies chosen by his opponents. His own strategy is a best response to his belief. Moreover, the probability distribution, induced on the play paths of the game, by his own strategy and his belief coincides with the distribution induced by the real strategies chosen by the players. Thus, his belief and the real strategies are realization equivalent and any statistical learning can only reinforce his belief.

While this notion of equilibrium is significantly weaker than the one of Nash equilibrium, where the belief must coincide with the true strategies of the opponents, we show that the actual behavior induced by any private-beliefs equilibrium must coincide with a behavior induced by some Nash equilibrium, even when we allow perturbations of both.

Despite being weaker than Nash equilibrium, private-beliefs equilibrium still captures an important aspect of stability. A player can learn nothing about the correctness of conjectures that relate to histories that have zero probability of being reached. On the other hand, on positive probability histories, statistical learning can serve only to reinforce his private beliefs, even if opportunities for learning are very readily available. Suppose, as an extreme example, that the infinite game with a given private-beliefs equilibrium is played repeatedly. As more sequences of realized infinite play paths are made available to a player, his confidence in his possibly false beliefs will only increase. This is due to the fact that, on the positive probability histories, his beliefs are statistically correct, while zero probability histories, where he may be wrong, will never be reached.

Indeed, the emergence of private-beliefs equilibrium and a related notion of self enforcing equilibrium were recently discovered in two different models of learning. Fudenberg and Levine (1990) studied steady state learning in an overlapping generations model where players were randomly matched to play a fixed finite extensive form game. They showed that, as the life length of the players becomes longer, the average strategy played converges to self confirming equilibrium of the extensive form game. Kalai and Lehrer (1990a) studied an n-person infinitely repeated game with discounting. They showed that if the players optimally respond to initial private beliefs about opponents' strategies, they must converge with time to play according to an $\epsilon$-perturbed private-beliefs equilibrium with arbitrarily small $\epsilon$.

Learning, as discussed above, for example, does not lead to a full equilibrium but only converges to one. This is due to the fact that Bayesian updating in general can only lead to approximately correct posterior probabilities of future events (see, for example, Blackwell and Dubins, 1962, and Kalai and Lehrer, 1990b). Thus, we can only hope to obtain equilibria with approximately correct private beliefs. For this reason private-beliefs equilibria and their perturbed versions are the subject of this paper. Because these are new concepts, one would like to know the type of behaviors they induce. The main message of this paper--that they induce the same behaviors as their Nash counterparts--is surprising and not trivial. This fact, however, relies on the perfect monitoring property of the games considered here, and the fact that players know their own payoff matrices.

While the notion of private-beliefs equilibrium is only defined here for the simple model of infinitely repeated games, it seems to be applicable to other models of learning, be it strategic, scientific, or others. One may think of them as a type of informationally local maxima (possibly by choice) which the players assume to be global and would never learn the truth. With this interpretation, the social implications of playing them seem important.

A well-known example, where a loss of welfare is associated with a private-beliefs equilibrium, is the multi-armed bandit game. There, a single player with pessimistic beliefs about the payoffs of alternative activities uses repeatedly the activity whose payoff is actually the lowest. Nevertheless, acting optimally under his highly false prior beliefs, he concludes that experimenting is wasteful, and continues to use the inferior activity. This phenomenon contradicts our main result which states that acting optimally under private beliefs must coincide with a behavior of a real payoff maximizer. The difference lies in the fact that, unlike an assumption in our model, the multi-armed bandit player does not know his own payoff matrix.

From a game theoretic viewpoint, private-beliefs equilibria are interesting because of weaker informational requirements they impose on the players. They allow for any private beliefs about opponents' strategies and do not require the use of any common prior probability distributions on the players' payoff matrices, their strategies, or other such parameters. The only stability they impose is with regard to statistical learning of opponents' strategies. We find it interesting that, despite this significant relaxation, they end up yielding the same behaviors as Nash equilibria.

## 2. The Repeated Game

First, we briefly review the standard model of n-player discounted repeated game with perfect monitoring. An n-person stage game is described by a set of action combination $\Sigma = \times_{i=1}^{n} \Sigma_i$ with $\Sigma_i$ denoting a finite action set of player i. Functions $u_i : \Sigma \to \mathbb{R}$ describe the stage game payoffs of the players.

The set of histories of length t, $H_t$, is defined to be the set of all t tuples of elements of $\Sigma$, i.e., $\Sigma^t$. There is a singleton set $H_0$ containing the "empty history" and $H = \cup_t H_t$ is the set of all finite length histories.

An infinite play path is an element of $\Sigma^\infty$. We assume that each player has a discount parameter $\lambda_i$, $0 < \lambda_i < 1$, by which he evaluates the payoff received along play paths. Thus, if $z = (z^1, z^2, \ldots)$ is a play path, we define

$$u_i(z) = (1 - \lambda_i) \, \Sigma_t \, \lambda_i^{t-1} u_i(z^t).$$

A (behavior) strategy of player i is a function $f_i : H \to \Delta(\Sigma_i)$ with $\Delta(\Sigma_i)$ denoting the set of probability distributions on $\Sigma_i$. Thus, in the definition of a strategy, we are implicitly assuming that the game is played with perfect monitoring.

A strategy vector $f = (f_1, f_2, \ldots, f_n)$ is a vector consisting of individual strategies. Such a strategy vector induces a probability distribution over the play paths of the game, yielding expected utilities described as follows.

With every history h we associate a cylinder set c(h) consisting of all the play paths going through h, i.e., having h as a prefix. We will use h to denote the history and also to denote the event c(h) when we think of the space of play paths. The $\sigma$-algebra defined over the set of all play paths is defined to be the smallest one to contain all the cylinder sets, c(h). Following this standard probability formulation it suffices to assign probabilities to all the cylinder sets in order to obtain a probability distribution over the set of all play paths. We follow this procedure in defining a probability distribution $\mu_f$ on the set of play paths for every strategy vector f.

We define $\mu_f$ of the empty history to be one and proceed inductively. For a history h and an action combination a, we define $\mu_f(ha) = \mu_f(h)\ \Pi_i\ f_i(h)(a_i)$. Thus, the probability of the history consisting of h followed by the action combination a is the probability of h times the product of the conditional probabilities of each player taking his action $a_i$ given the history h.

Now we complete the definition of the repeated game by defining individual payoffs for each strategy vector f,

$$U_i(f) = Eu_i(z) = \int u_i(z)d\mu_f(z).$$

Equivalently, one can define the expected stage payoffs and take the discounted sum of these.

## 3. Description of the Main Result

As usual, we say that a strategy $f_i$ is a best response to $g_{-i}$ if $U_i(g_1,\dots,h_i,\dots,g_n) - U_i(g_1,\dots,f_i,\dots,g_n) \le 0$ for every strategy $h_i$. If the right side 0 is replaced by $\varepsilon$ we say that $f_i$ is an $\varepsilon$-best response to $g_{-i}$.

Definition 1: A private-beliefs equilibrium (PB eq.) is a strategy vector g with a beliefs matrix $(g^i_j)_{1\le i,j\le n}$ satisfying for each player i:

(0) $g^i_i = g_i$;

(1) $g_i$ is best response to $g^i_{-i}$; and

(2) $\mu_g = \mu_{g^i}$.

The idea is that the i-th row, $g^i$, represents the belief of player i about the strategy vector that is played. Condition (0) is required since we assume that every player knows his own strategy. Condition (1) expresses the usual utility maximization assumption where $g^i_{-i}$ is the (n - 1) vector consisting of all the entries of $g^i$ excluding the i-th. Condition (2) states that g plays like $g^i$. It expresses the idea that any statistical study will serve to only strengthen player i's belief that $g^i$ is being played.

Obviously, in the above definition, if we let $g^i = g$ for all i we have a Nash equilibrium as a special case.

It is important to note that, while player i's belief about j's strategy is described here by a single behavior strategy, $g^i_j$, it is not a serious restriction. By a well-known theorem of Kuhn (1953) (see also Aumann, 1964, and Selten, 1975), a mixed strategy of player j, i.e., a probability distribution over his strategies, can be replaced by a single realization equivalent behavior one. Thus, a belief of player i given by a probability distribution over j's strategies can be replaced by an equivalent belief consisting of a single behavior strategy.

We wish to allow for perturbations in the accuracy of the beliefs of such an equilibrium.

Definition 2: Let $\varepsilon > 0$ and let $\mu$ and $\tilde{\mu}$ be two probability measures defined on the same space. We say that $\mu$ is $\varepsilon$-close to $\tilde{\mu}$ if there is a measurable set Q satisfying:

(i) $\mu(Q)$ and $\tilde{\mu}(Q)$ are greater than $1 - \varepsilon$; and

(ii) for every measurable set $A \subseteq Q$

$$(1 - \varepsilon)\tilde{\mu}(A) \leq \mu(A) \leq (1 + \varepsilon)\tilde{\mu}(A)$$

Notice that this notion of $\varepsilon$-closeness is strong since it imposes a tough restriction even on small probability events.

Definition 3: Let f and g be two joint strategies, and let $\varepsilon > 0$. We say that f plays $\varepsilon$-like g if $\mu_f$ is $\varepsilon$-close to $\mu_g$.

The above notion of playing $\varepsilon$-like is strong due to the strength of the measures closeness notion. It guarantees that even for very long histories, ones that have small probability of being reached, f and g will assign close probabilities, with ratio close to one. In other words, such f,g are very likely to play almost the same throughout an infinite game (additional discussion of this point is given in Sections 5 and 6).

Definition 4: An n-vector of strategies, g, is a private-beliefs ε-equilibrium (PB ε-eq) if there is a matrix of n-vector of strategies $(g_j^i)_{1\le i,j\le n}$ such that for every player i:

(0) $g_i^i = g_i$;

(1) $g_i$ is a best response to $g_{-i}^i$; and

(2) g plays ε-like $g^i$.

Our main result is:

Theorem 1: For every ε > 0 there is $\bar{\eta} > 0$ s.t. for all $\eta \le \bar{\eta}$ if g is a private-beliefs η-equilibrium then there exists f, s.t.

(i) g plays ε-like f, and

(ii) f is ε-Nash equilibrium.

Theorem 1 states that behaviors induced by private-beliefs equilibria are the same as behaviors induced by Nash equilibria. The theorem also holds when perturbations are allowed and then it has further implications. Recall that ε-Nash equilibrium requires that each player chooses a strategy that is ε-optimal against the precise strategies used by his opponent, i.e., his payoff should be within ε of the optimally possible against theirs. On the other hand, the ε-perturbed private-beliefs equilibrium requires precise optimization but against beliefs that are almost accurate. Relating these two notions of perturbations, as done in Theorem 1, turns out to be nontrivial.

## 4. Proof of the Main Result

We assume that g is PB $\eta$-eq, where $\eta$ is to be determined later. Thus, there exists $(g_j^i)_{i,j}$ satisfying the conditions of Definition 4. We show that if $\eta$ is small enough then there exists an $\varepsilon$-Nash equilibrium f s.t. g plays $\varepsilon$-like it. The strategy f will be defined by using all $(g_j^i)$. In constructing $f_i$ we describe the individual strategy of player i. However, part (2) of Definition 4 is expressed in terms of histories consisting of joint actions. Thus, we need a lemma that connects between the closeness notion of the aggregate ("plays like") and some closeness notion of the individuals' actions.

In the following lemma we confine ourselves to finitely repeated games of length s + 1. Any vector of n individual strategies, f, in the infinitely repeated game induces a joint strategy in the finitely repeated game. Moreover, the notion "play $\varepsilon$-like" extends naturally to finite games.

Lemma 1: In a finitely repeated game of length s + 1, for every $\delta > 0$ there is $\bar{\eta} > 0$ s.t. if k plays $\eta$-like k' (k and k' are joint strategies) with $\eta \leq \bar{\eta}$, then there exists a set Q of play paths (of the finitely repeated game), s.t.

(i) $\mu_k(Q)$ and $\mu_{k'}(Q)$ are greater than $1 - \delta$;

(ii) after every history h satisfying $c(h) \cap Q \neq \emptyset$ and for every player i, the probability according to both k and k', to play an action $a_i$ that satisfies $|k_i(h)(a_i)/k_i'(h)(a_i) - 1| \leq \delta$ is greater than or equal to $1 - \delta$; and

(iii) for every h satisfying $c(h) \cap Q \neq \emptyset$ $|\mu_k(h)/\mu_{k'}(h) - 1| \leq \delta$.

<u>Proof</u>: Otherwise there exists $\delta > 0$ s.t. for every $\eta > 0$ there are $k_\eta$ and $k'_\eta$ for which the lemma does not hold and $k_\eta$ play $\eta$-like $k'_\eta$.

By letting $\eta$ tend to zero along a sequence and by taking a converging subsequence of the corresponding $k_\eta$, $k'_\eta$ one gets two limits k and k'. These limits do not satisfy the conclusion of the lemma with $\delta$. However, the strategy k plays 0-like k'. Thus, k and k' are realization equivalents. This means that $k_i(h)$ is identical to $k'_i(h)$ for any history reached (according to either one) with positive probability. Hence, k and k' satisfy the conclusion of the lemma with $\delta = 0$, which is a contradiction. //

If an action $a_i \in \Sigma_i$ satisfies the inequality stated in (ii) we say that $a_i$ is <u>$(k,k',h)$-$\delta$-good</u>. If $a_i \in \Sigma_i$ is $(g,g^j,h)$-$\delta$-good for all j, we say that $a_i$ is <u>h-$\delta$-good</u>. Due to the discounting of future rewards there exists a time s after which all the payoffs cannot account for more than $\varepsilon/3$ of player i's total payoff, $i = 1,\dots,n$. Denote by $|h|$ the length of the history h (i.e., $h \in H_t$ implies that $|h| = t$), and let $\delta$ be a small number to be determined later.

By Lemma 1 applied to s, just described, and to $\delta$, there is $\eta > 0$ so that if g plays $\eta$-like $g^j$ then there is a set $Q^j$ of play paths (in the finite repeated game of length s + 1) for which (i), (ii), and (iii) of the lemma are satisfied with g replacing k and $g^j$ replacing k'. Since g is PB $\eta$-eq the hypothesis is automatically satisfied for all j.

Taking an intersection of all $Q^i$'s one obtains a set Q of play paths whose probability (w.r.t. g) is at least $1 - n\delta$. The factor n appears due to the number of different selections of a player j above. The probability w.r.t. to g, of all the h-$\delta$-good actions is at least $1 - (n - 1)\delta$. The factor $(n - 1)$ appears because when dealing with the player i we looked at all $g_i^j$, $j \neq i$.

Before proceeding with the construction of f we need the following lemma.

Lemma 2: In a finitely repeated game of length $s + 1$, for every $\gamma > 0$ there is $\bar{\delta} > 0$ s.t. for any strategy vector g, if $\bar{Q}$ is a set of play paths whose probability w.r.t. g is $\geq 1 - n\delta$ with $\delta \leq \bar{\delta}$ then there is a set $Q' \subseteq \bar{Q}$ s.t.

(i) the probability of $Q'$ (w.r.t. g) is at least $1 - \gamma$;

(ii) for every h satisfying $c(h) \cap Q' \neq \emptyset$ and i there is a subset of player i's actions $A_i(h) \subseteq \Sigma_i$ s.t. the probability of $A_i(h)$ according to $g_i(h)$ is at least $1 - \gamma$; and

(iii) if $c(h) \cap Q' \neq \emptyset$ and $|h| < s$ then the continuation ha satisfies $c(ha) \cap Q' \neq \emptyset$ for all $a \in \times A_i(h)$.

Proof: By a continuity argument similar to the one employed in the proof of Lemma 1. The details are omitted. //

Using Lemma 2 for $\gamma > 0$ to be determined and $\bar{Q} = Q$ we find a set $Q'$ of play paths (in the finitely repeated game of length $s + 1$) having the properties specified as in the lemma. Together with the other h-good actions of other players, $a_i$ cannot form a continuation of h which is disjoint to $Q'$. Recall that $Q' \subseteq Q$. Thus, at any h satisfying $c(h) \cap Q' \neq \emptyset$ the probability of the actions in $\Sigma_i$ that are h-$\delta$-good is at least

$1 - (n - 1)\delta$. To sum up, the probability of an action $a_i \in \Sigma_i$, w.r.t. $g_i(h)$, to be in $A_i(h)$ and to be h-$\delta$-good is at least $1 - \gamma - (\eta - 1)\delta$.

The strategy f will be defined first on histories reached (by f itself) with positive probability. Therefore, it should be done inductively. The histories h, $|h| \le s$, reached with positive probability (later we will refer to those as <u>positive histories</u>), will be histories whose corresponding cylinder sets intersect Q' (not every h intersecting Q' will be a positive one).

Suppose that h, $|h| \le s$, is a positive history. The mixed action $f_i(h)$ will be defined according to the following procedure. Delete from the support of $g_i(h)$ all the actions that are not in $A_i(h)$ or not h-$\delta$-good. The remainder has probability (w.r.t. g) at least $1 - \gamma - (n - 1)\delta$. Normalize the probability of all the remaining actions. If, on the other hand, h is a positive history (i.e., reached with positive probability according to f) and $|h| > s$, then define $f_i(h) = g_i(h)$. In other words, $f_i$ differs from $g_i$ only on positive histories h of length less than or equal to s.

So far, we have defined f on histories reached with positive probability. In case h is not a positive history, then consider the shortest prefix of h, $\bar{h}$, which has a zero probability and distinguishes two cases. (i) If the last action vector of $\bar{h}$ corresponds to a deviation of only <u>one</u> player, say, j, from a positive history, f(h) is defined as $g^j(h)$. In other words, once j deviates, all other players follow his belief $g_i^j$ forever. (ii) If $\bar{h}$ does not correspond to a <u>one-player</u> deviation, define f(h) arbitrarily. Since Nash equilibrium deals solely with unilateral deviations from positive histories, multi-player deviations will not matter.

To see that g plays $\varepsilon$-like f, notice the following. f coincides with g on positive histories of length greater than s. Thus, it suffices to show that g plays $\varepsilon$-like f in the s-truncated repeated game. The set Q' of play paths has probability (w.r.t. g) of at least $1 - \gamma$. At any positive history the probability of the actions in the support of $f_i(h)$ were obtained by dividing by at most $(1 - \gamma - (n - 1)\delta)$ the corresponding probability assigned by $g_i(h)$. Thus, the probability of the union of the histories h intersecting Q' that are positive is at least $(1 - \gamma - (n - 1)\delta)^{ns}$ times the probability of Q'. In other words, on a set of histories with length less than or equal to s, whose union's probability is at least $(1 - \gamma)(1 - \gamma - (n - 1)\delta)^{ns}$ the ratio of the probabilities (assigned by f and by g) is closed to 1 up to $(1/(1 - \gamma - (n - 1)\delta)^{ns}) - 1$. If $\gamma$ and $\delta$ are small then g plays $\varepsilon$-like f. By taking $\gamma$, $\delta$ even smaller, we may assume that g plays $\varepsilon/10$-like f.

It remains to show that f is an $\varepsilon$-Nash equilibrium. It suffices to show that by deviating to any pure strategy k player 1, for example, can gain no more than $\varepsilon$.

If all the histories in the support of $\mu_{(k,f_{-1})}$ are positive histories, then

(7) $$U_1(k,f_{-1}) \leq \varepsilon/10 + U_1(k,g_{-1})/(1 - \varepsilon/10).$$

The fact that $\mu_g$ is $\eta$-close to $\mu_{g^1}$ implies that any player 1's strategy (say, k) satisfies

(8) $$1 - \eta \leq \mu_{(k,g_{-1})}(h)/\mu_{(k,g^1_{-1})}(h) \leq 1 + \eta$$

for every h assigned a positive probability by g (recall (iii) of Lemma 1).

Inequalities (7) and (8) imply

$$(9) \qquad \begin{aligned} U_1(k,f_{-1}) &\le \varepsilon/10 + U_1(k,g_{-1})/(1-\varepsilon/10) \\ &\le \varepsilon/10 + U_1(k,g^1_{-1})/(1-\eta)(1-\varepsilon/10) \text{ [if } \eta \text{ is small enough]} \\ &\le U_1(k,g^1_{-1}) + \varepsilon/6. \end{aligned}$$

If, however, the support of $\mu_{(k,f_{-1})}$ does not consist only of positive histories, then there is a positive history h s.t. k(h) is not in the support for $f_1(h)$. Suppose first that $|h| \le s$. At this point all the players follow $g^1_{-1}$ because player 1 is responsible for the deviation. However, $g_1$ is a best response to $g^1_{-1}$. Thus, by altering k(h) for $|h| \le s$ to be within what is expected from player 1 (i.e., k(h') is in the support of $f_1(h')$ for every continuation h' of h), one can only improve upon k (in the sense that the altered strategy yields at least as much as k). The same argument holds for every non-positive history reached due to k. Therefore, one may find another pure strategy of player 1, say, $k^1$, which satisfies: (i) $U_1(k,f_{-1}) \le U_1(k^1,f_{-1})$, and (ii) every history h, $|h| \le s$, played with a positive probability according to $(k^1,f_{-1})$ is a positive history. Since (ii) holds only for histories with length less than s and since, after time s, the deviation k may possibly contribute at most $\varepsilon/3$, one obtains (similarly to (9) with the additional $\varepsilon/3$):

$$(10) \qquad U_1(k,f_{-1}) \le U_1(k,g^1_{-1}) + \varepsilon/6 + \varepsilon/3 \text{ for every } k.$$

Since $g_1$ is a best response to $g^1_{-1}$,

$$(11)\qquad U_1(k,g^1_{-1}) + \varepsilon/2 \le U_1(g^1) + \varepsilon/2 \le \qquad \text{[since g } \eta\text{-plays like } g^1]$$

$$U_1(g)/(1-\eta) + \eta + \varepsilon/2 \le \qquad \text{[since g } \varepsilon/10 \text{ plays like f]}$$

$$U_1(f)/(1-\eta)(1-\varepsilon/10) + \eta + \varepsilon/10 + \varepsilon/2 \qquad \text{[if } \eta \text{ is small enough]}$$

$$\le U_1(f) + \varepsilon.$$

(10) and (11) imply

$$U_1(k,f_{-1}) \le U_1(f) + \varepsilon.$$

Thus, f is an $\varepsilon$-Nash equilibrium as desired.

## 5. An Asymptotically Equivalent Definition

The notion of "g plays $\varepsilon$-like f" used in this paper is strong, since it applies to the ratios of the probabilities induced by the two strategies. So even for histories that have very small probabilities of being played, $\mu_f(h)/\mu_g(h)$ being close to one is meaningful. If we used the differences of the measures $|\mu_f(h) - \mu_g(h)| \le \varepsilon$ for every h to describe closeness, then we could have small probability histories that are twice as likely according to f as with g and still declare that f plays close to g. It turns out, however, that the event consisting of those histories has a small probability and we actually obtain equivalence of the two notions of closeness as we proceed to show.

<u>Proposition 1</u>: For every $\varepsilon > 0$ there exists $\delta > 0$ s.t. if $\mu$ and $\tilde{\mu}$ are two measures defined on the same measure space and if $|\mu(A) - \tilde{\mu}(A)| < \delta$ for

every event A, then $\mu$ is $\varepsilon$-close to $\tilde{\mu}$.

Proof: Let $1/4 > \delta > 0$, to be specified later, and let $\mu,\tilde{\mu}$ be two arbitrary measures satisfying $|\mu(A) - \tilde{\mu}(A)| < \delta$ for every event A.

Define[1] $S = \mathrm{supp}(\mu) \cap \mathrm{supp}(\tilde{\mu})$, and let $S^c$ be its complement. It is clear that $\mu(S^c)$ and $\tilde{\mu}(S^c)$ are both smaller than $\delta$.

On S $\mu$ is absolutely continuous w.r.t. $\tilde{\mu}$ (i.e., for every $D \subseteq S$ $\mu(D) > 0$ implies $\tilde{\mu}(D) > 0$). By the Radon-Nikodym theorem there exists a measurable function f satisfying

$$\mu(D) = \int_D f d\tilde{\mu} \text{ for all events } D \subseteq S.$$

Define $\bar{B} = \{\omega | f(\omega) - 1 > \sqrt{\delta}\}$, and

$\underline{B} = \{\omega | f(\omega) - 1 < -\sqrt{\delta}\}$.

Observe that $\mu(\bar{B}) - \tilde{\mu}(\bar{B}) > \sqrt{\delta}\tilde{\mu}(\bar{B})$. Therefore, $\tilde{\mu}(\bar{B}) < \sqrt{\delta}$. For a similar reason, $\tilde{\mu}(\underline{B}) < \sqrt{\delta}$. Defining $Q = S - (\bar{B} \cup \underline{B})$, we get $\tilde{\mu}(Q) \geq 1 - \delta - 2\sqrt{\delta}$. Moreover, for every event $D \subseteq Q$ one gets $|\mu(D)/\tilde{\mu}(D) - 1| < 2\sqrt{\delta}$. Therefore, if $\delta + 2\sqrt{\delta} < \varepsilon$ then $\mu$ is $\varepsilon$-close to $\tilde{\mu}$. //

In view of this discussion, we can introduce a seemingly different notion of private-beliefs equilibrium which asymptotically behaves like the one defined above.

Definition 5: A vector $g = (g_1,\ldots,g_n)$ of individual strategies is a private-beliefs $\varepsilon$-equilibrium in the norm sense if there is a matrix $(g_j^i)$

---

[1] $\mathrm{Supp}(\mu)$ and $\mathrm{supp}(\tilde{\mu})$ are the supports of $\mu$ and $\tilde{\mu}$, respectively.

with:

(0) $g^i_i = g_i$;

(1) $g_i$ is a best response to $(g^i_{-i})$; and

(2) for every measurable set of play paths, A, and every i, the probability of A, w.r.t. to g, and w.r.t. $g^i$, are $\varepsilon$-close to each other. I.e.,

$$(*) \qquad |\mu_g(A) - \mu_{g^i}(A)| \leq \varepsilon \text{ for every event } A.$$

It is easy to see that for arbitrarily small $\delta$, there is $\eta > 0$ such that all private-beliefs $\eta$-equilibria are private-beliefs $\delta$-equilibrium in the norm sense. Proposition 1 demonstrates that the converse is also true and that limit results, obtained for the two different notions, coincide. In particular, one obtains:

Theorem 2: In an infinite repeated game with discounting, for every $\varepsilon > 0$ there is a $\eta > 0$ s.t. if g is a private-beliefs $\eta$-equilibrium in the norm sense then there is a strategy f s.t.

(1) g plays $\varepsilon$-like f; and

(2) f is a $\varepsilon$-Nash equilibrium.

Proof: Fix $\varepsilon > 0$. By Theorem 1 there exists $\delta > 0$ s.t. if g is PB $\delta$-eq there exists f satisfying conclusions (1) and (2) of Theorem 2. In view of Proposition 1 there exists $\eta > 0$ such that if g is a private-beliefs equilibrium in the norm sense, then g is PB $\delta$-eq. This concludes the proof. //

## 6. The Strong Private-beliefs Equilibrium

In Definition 4 we required that for every i g plays $\varepsilon$-like $g^i$. Thus, there exists a big set of histories, say, Q, such that for every $h \in Q$, the ratio between the probabilities assigned to h by g and by $g^i$ is close to 1. Thus, this definition allows the ratio to be far from 1 on a set of histories whose union has a positive probability.

The strong version of the private-beliefs equilibrium requires that the ratio between the probabilities, assigned by g and by $g^i$ to every history reached with a positive probability is close to 1. In case a strong private-belief $\varepsilon$-equilibrium is played there will be no dramatic changes in any player's beliefs, while if the equilibrium is of the weaker type there are chances (although slim) that players will have to alter their beliefs drastically as a result of new incoming observations. The strong concept is defined here:

Definition 6: A vector $g = (g_1,\ldots,g_n)$ of individual strategies is a strong private-beliefs $\varepsilon$-equilibrium (SPB $\varepsilon$-eq) if there is a matrix $(g^i_j)$ with

(0) $g^i_i = g_i$;

(1) $g_i$ is a best response to $(g^i_{-i})$; and

(2) every history h with $\mu_g(h) > 0$ satisfies

$$|\mu_g(h)/\mu_{g^i}(h) - 1| < \varepsilon \text{ for every } i.$$

In other words, when a SPB $\varepsilon$-eq is played: (a) each player plays a best response to his beliefs and (b) each player knows, before the game starts (up to $\varepsilon$), anything that he can possibly learn during the course of the game. Thus, players play a best response to their beliefs not only at the beginning of the game but also at any stage and after every likely history. This suggests that a SPB $\varepsilon$-eq, g, is realization equivalent to some approximate Nash equilibrium. That is, g plays 0-like some $\varepsilon$-Nash equilibrium as shown in what follows.

Theorem 3: For $\varepsilon > 0$ there exists $\delta > 0$ s.t. if g is SPB $\delta$-eq then there exists an $\varepsilon$-Nash equilibrium, f, which plays 0-like g.

Proof: Let g be a SPB $\delta$-eq. Define f to be identical to g on histories h which satisfy $\mu_g(h) > 0$. If $\mu_g(h) = 0$ and h corresponds to a unilateral deviation of player i from $g_i$, define f to be $g^i$ on h and on all its continuations. On all other histories define f arbitrarily.

The proof that f is $\varepsilon$-Nash equilibrium (where $\varepsilon$ depends on $\delta$) follows some of the arguments of Theorem 1's proof and is therefore omitted. //

## References

Aumann, R. J. (1964), "Mixed and Behaviour Strategies in Infinite Extensive Games," in M. Dresher, L. S. Shapley and A. W. Tucker (eds.), Advances in Game Theory, pp. 627-650, Annals of Mathematics Studies, 52, Princeton University Press.

Blackwell, D. and L. Dubins (1962), "Merging of Opinions with Increasing Information," Annals of Mathematical Statistics, 38, 882-886.

Fudenberg, D. and D. Levine (1990), "Steady State Learning and Self-Confirming Equilibria," Massachusetts Institute of Technology.

Kalai, E. and E. Lehrer (1990a), "Rational Learning Leads to Nash Equilibrium," Discussion Paper No. 895, CMSEMS, Northwestern University.

Kalai, E. and E. Lehrer (1990b), "Merging of Opinions Revisited," Northwestern University Discussion Paper.

Kuhn, H. W. (1953), "Extensive Games and the Problem of Information," in H. W. Kuhn and A. W. Tucker (eds.), Contributions to the Theory of Games, Vol. II, pp. 193-216, Annals of Mathematics Studies, 28, Princeton University Press.

Selten, R. (1975), "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," International Journal of Game Theory, 4, 25-55.