Econometrica, Vol. 61, No. 3 (May, 1993), 523–545

SELF-CONFIRMING EQUILIBRIUM

BY DREW FUDENBERG AND DAVID K. LEVINE[1]

In a self-confirming equilibrium, each player's strategy is a best response to his beliefs about the play of his opponents, and each player's beliefs are correct along the equilibrium path of play. Thus, if a self-confirming equilibrium occurs repeatedly, no player ever observes play that contradicts his beliefs, even though beliefs about play at off-path information sets need not be correct. We characterize the ways in which self-confirming equilibria and Nash equilibria can differ, and provide conditions under which self-confirming equilibria correspond to standard solution concepts.

KEYWORDS: Nash equilibrium, self-confirming equilibrium, correlated equilibrium, extensive-form games, learning in games.

1. INTRODUCTION

NASH EQUILIBRIUM AND ITS REFINEMENTS describe a situation in which (i) each player's strategy is a best response to his beliefs about the play of his opponents, and (ii) each player's beliefs about the opponents' play are exactly correct. We propose a new equilibrium concept, self-confirming equilibrium, that weakens condition (ii) by requiring only that players' beliefs are correct along the equilibrium path of play. Thus, each player may have incorrect beliefs about how his opponents would play in contingencies that do not arise when play follows the equilibrium, and moreover the beliefs of different players may be wrong in different ways.

The concept of self-confirming equilibrium is motivated by the idea that noncooperative equilibria should be interpreted as the outcome of a learning process, in which players revise their beliefs using their observations of previous play. Suppose that each time the game is played, the players observe the actions chosen by their opponents, but that players do not observe the actions their opponents would have played at the information sets that were not reached along the path of play. Then, if a self-confirming equilibrium occurs repeatedly, no player ever observes play that contradicts his beliefs, so the equilibrium is "self-confirming" in the weak sense of not being inconsistent with the evidence. By analogy with the literature on the bandit problem (e.g., Rothschild (1974)) one might expect that a non-Nash self-confirming equilibrium can be the outcome of plausible learning processes. This point was made by Fudenberg and Kreps (1988), who gave an example of a game in which a simple learning process converges to a non-Nash outcome unless the players engage in a sufficient amount of "experimentation" with actions that do not maximize the

[1] This paper contains the sections on self-confirming equilibrium from the earlier paper "Steady State Learning and Self-Confirming Equilibrium." We thank David Kreps for many helpful conversations. Pierpaolo Battigalli, Martin Hellwig, three referees, and conference participants at Luminy, Ohio State, and Stonybrook made useful comments. National Science Foundation Grants 89-07999-SES, 90-23697-SES, 88-08204-SES, and 90-9008770-SES, and the John Simon Guggenheim Foundation provided financial support. We thank the IDEI, Toulouse, for its hospitality.

current period's expected payoff. Our notion of self-confirming equilibrium was developed to capture the implications of learning when players do little or none of this experimentation.

The idea of self-confirming equilibrium is implicit in the example of Fudenberg and Kreps. Our contribution here is to give a formal definition of the concept, and to explore its properties more fully. We show that there are three reasons that a self-confirming equilibrium outcome may not be the outcome of a Nash equilibrium. These are: (i) two players can have inconsistent beliefs about the play of a third at an information set that is relevant to both of them; (ii) different pure strategies that a player assigns positive probability may be best responses to different beliefs about his opponents' play; and (iii) a player's subjective uncertainty about the play of two opponents may be correlated. If all three effects are absent, the equilibrium outcome can be supported by a Nash equilibrium. When the first effect is absent, we say that the self-confirming equilibrium is *consistent*; every outcome of a consistent self-confirming equilibrium is also the outcome of one of Forges' (1985) extensive-form correlated equilibria. Moreover, all self-confirming equilibria are consistent in games with *observed deviators*. This condition, which means that deviations by different players cannot lead to the same information set, is satisfied in all two-player games of perfect recall, and also in all multistage games with observed actions.

Although self-confirming equilibrium is motivated by the idea of a dynamic learning process, such processes are not explicitly modelled in this paper. One learning-theoretic foundation for self-confirming equilibrium is presented in our (1993) paper "Steady State Learning and Nash Equilibria," which considers the steady states of a system in which a fixed stage game is played repeatedly by a large population of players who are randomly matched with one another. Players learn about their opponents' strategies by observing the actions played in their own matches; they do not observe the intended off-path play of their opponents or the actions chosen in other matches. Individual players remain in the population a fixed number of periods; new players enter each period to keep the total population size constant. Entering players believe that the opponents' play corresponds to a fixed but unknown steady-state distribution; players update their exogenous priors over this distribution using Bayes' rule. Players choose their strategies in each period to maximize their expected present value given their beliefs. If the lifetimes are long, then steady state distributions approximate those of self-confirming equilibria. Moreover, if the discount factors are low, so that players optimally choose to do little experimentation, then non-Nash outcomes can be steady states. (If the players are sufficiently patient, they will do enough experimentation to learn the relevant aspects of off-path play, and steady states approximate Nash equilibria.)

Kalai and Lehrer (1991b) consider an alternative model of learning, in which a single player 1, player 2, etc. play a repeated game. They show that, if the priors and the true play jointly satisfy an "absolute continuity" condition, then play will eventually resemble that of a "private-beliefs equilibrium." This concept refines self-confirming equilibrium by requiring that (1) each player's subjective uncertainty about his opponents corresponds to the product of independent marginal distributions, and (2) each player has "unitary beliefs," meaning that each strategy a player assigns positive probability is a best response to the same (possibly incorrect) beliefs about his opponents.

Kalai and Lehrer (1991a) show that private-beliefs equilibria are outcome-equivalent to Nash equilibria. Our results also yield this conclusion and show that it is due to the fact that the extensive form they consider has observed deviators. Theorem 1 shows that, in such games, every self-confirming equilibrium is consistent, and Theorem 3 shows that consistent self-confirming equilibria with independent, unitary beliefs are outcome-equivalent to Nash equilibria.

The Kalai-Lehrer learning model, like ours, supposes that players learn about their opponents' strategies by observing the actions that are played. Since the idea of the self-confirming equilibrium concept is to model the possibility that certain mistakes in beliefs can persist if not contradicted by the evidence, different assumptions about what the players observe would lead to alternative notions of equilibrium. The "conjectural equilibrium" of Battigalli (1987) takes as data an "observation function" specifying what players see when each profile is played; specifying that players see the corresponding terminal node yields our concept as a special case.[2] The generality of conjectural equilibrium makes it difficult to characterize. For this reason, we prefer to focus our attention on the case where the players observe one another's actions, which we believe corresponds to many situations of interest.

To illustrate the relationship between Nash and self-confirming equilibrium, note first that in a one-shot simultaneous-move game, every information set is reached along every path, so that self-confirming equilibrium reduces to the Nash condition that beliefs are correct at every information set. Somewhat less obvious is the fact that self-confirming equilibrium must have Nash outcomes in any two-player game, so long as each player has "unitary beliefs" in the sense described above.

Unitary beliefs seem natural if we think of equilibrium as corresponding to the outcome of a learning model with a single player 1 and a single player 2, and so forth, as in Fudenberg and Kreps. We were led to consider the alternative of heterogeneous beliefs, which allows each strategy a player uses with positive probability to be a best response to a different belief about his opponents, by our study of learning in models where a large number of individual players are randomly matched each period. In such models, heterogeneous beliefs can arise because different individuals have different learning experiences or different prior beliefs. When heterogeneous beliefs are allowed, the self-confirming

[2] In our terminology, Battigalli's concept imposes unitary beliefs. More strongly, he also supposes that beliefs are point-valued, that is that each player's beliefs correspond to a point mass on a particular strategy profile for his opponents. Rubinstein and Wolinsky (1990) refine this concept by adding a condition in the spirit of rationalizability. We should also mention the "conjectural equilibrium" of Hahn (1977), in which firms misperceive the market excess demand function, but correctly predict the prevailing price. Hahn shows that such misperceptions can lead to non-Walrasian outcomes, and that these outcomes can persist even if the firms' conjectures are "locally correct."

equilibria of two-player games need not be Nash equilibria, but rather are Nash equilibria of an extended version of the game in which players can observe the outcome of certain correlating devices, as in Forges' (1985) extensive-form correlated equilibria.

Since self-confirming equilibrium requires that beliefs be correct along the equilibrium path of play, it is inherently an extensive-form solution concept, in contrast to Nash equilibrium, which can be defined on the strategic form of the game. Indeed, two extensive-form games with the same strategic form can have different sets of self-confirming equilibria. This conclusion runs counter to the argument, recently popularized by Kohlberg and Mertens (1986), that the strategic form encodes all strategically relevant information, and two extensive forms with the same strategic form will be played in the same way. However, dependence on the extensive form is natural when equilibrium is interpreted as the result of learning, as the strategic form does not pin down how much of the opponents' strategies each player will observe when the game is played. In our view, the contrast between our approach and that of Kohlberg and Mertens shows that it is better to specify the process that leads to equilibrium play before deciding which games are equivalent or which equilibria are most reasonable.

2. THE EXTENSIVE FORM GAME

Consider an I-player extensive-form game of perfect recall. The game tree X, with nodes $x \in X$, is finite. The initial node, 0, corresponds to Nature's moves, if any; the terminal nodes are $z \in Z \subset X$. Information sets, denoted by $h \in H$, are a partition of $X \setminus (Z \cup 0)$. The information sets where player i has the move are $H_i \subset H$, and $H_{-i} = H \setminus H_i$ are information sets for other players. The feasible actions at information set $h_i \in H$ are denoted $A(h_i)$.

A pure strategy for player i, s_i, is a map from information sets in H_i to actions satisfying $s_i(h_i) \in A(h_i)$; S_i is the set of all such strategies. We let $s \in S = \times_{i=1}^{I} S_i$ denote a strategy profile for all players, and $s_{-i} \in S_{-i} = \times_{j \neq i} S_j$. Each player i receives a payoff that depends on the terminal node. Player i's payoff function is denoted u_i: $Z \to \mathbb{R}$; each player knows (at least) his own payoff function, the extensive form of the game, and the probability distribution over Nature's moves. Let $\Delta(\cdot)$ denote the space of probability distributions over a set. Then a mixed strategy profile is $\sigma \in \times_{i=1}^{I} \Delta(S_i)$.

Let $Z(s_i)$ be the subset of terminal nodes that are reachable when s_i is played. Let $H(s_i)$ be the set of all information sets that can be reached if s_i is played; for a mixed strategy σ_i, set $H(\sigma_i) = \bigcup_{s_i \in \text{support}(\sigma_i)} H(s_i)$.

We will also need to refer to the information sets that are reached with positive probability under σ, denoted $\bar{H}(\sigma)$. Notice that if σ_{-i} is completely mixed, then $\bar{H}(s_i, \sigma_{-i}) = H(s_i)$, as every information set that is potentially reachable given s_i has positive probability.

In addition to mixed strategies, we define behavior strategies. A behavior strategy for player i, π_i, is a map from information sets in H_i to probability distributions over moves: $\pi_i(h_i) \in \Delta(A(h_i))$, and Π_i is the set of all such strategies. As with pure strategies, $\pi \in \Pi = \times_{i=1}^{I} \Pi_i$, and $\pi_{-i} \in \Pi_{-i} = \times_{j \neq i} \Pi_j$. Let $p(x|\pi)$ be the probability that node $x \in X$ is reached under profile π, and let $p(h_i|\pi) = \Sigma_{x \in h_i} p(x|\pi)$. (Note that the probability p will reflect the probability distribution on nature's moves.)

Since the game has perfect recall, by Kuhn's Theorem each mixed strategy σ_i induces an equivalent behavior strategy denoted $\hat{\pi}_i(\cdot|\sigma_i)$.[3] In other words, $\hat{\pi}_i(h_i|\sigma_i)$ is the probability distribution over actions at h_i induced by σ_i.

Since we have assumed that all players know the structure of the extensive form, their own payoff function, and the probability distribution on nature's moves, the only uncertainty each player faces concerns the strategies his opponents will play. To model this "strategic uncertainty," we let μ_i be a probability measure over Π_{-i}, the set of other players' behavior strategies. Fix s_i. Then the marginal probability of a terminal node z is

$$(2.1) \qquad p_i(z|s_i,\mu_i) = \int p_i(z|s_i,\pi_{-i})\mu_i(d\pi_{-i});$$

the marginal probability $p(h_j|s_i,\mu_i)$ of an information set is the sum of the marginal probabilities of its terminal successors.

This in turn gives rise to preferences

$$(2.2) \qquad u_i(s_i,\mu_i) = \sum_{z \in Z(s_i)} p_i(z|s_i,\mu_i)u_i(z).$$

It is important to note that even though the beliefs μ_i are over opponents' behavior strategies, and thus reflect player i's knowledge that his opponents choose their randomizations independently, the marginal distribution $p_i(\cdot|s_i,\mu_i)$ over terminal nodes can involve correlation between the opponents' play. For example, if players 2 and 3 simultaneously choose between U and D, player 1 might assign probability $1/4$ to $\pi_2(U) = \pi_3(U) = 1$, and probability $3/4$ to $\pi_2(U) = \pi_3(U) = 1/2$. Even though both profiles in the support of μ_i suppose independent randomization by players 2 and 3, the marginal distribution on their joint actions is $p(U,U) = 7/16$ and $p(U,D) = p(D,U) = p(P,D) = 3/16$, which is a correlated distribution. This correlation reflects a situation where

[3] Two strategies s_i and s_i' for player i are said to be equivalent if they lead to the same distribution over terminal nodes for all s_{-i}. The behavior strategy equivalent to σ_i is uniquely determined at all information sets that are not precluded by σ_i (that is, at all information sets $h_i \in H_i \cap H(\sigma_i)$) by the following formula, where $R(h_i) = \{s_i | h_i \in H(s_i)\}$ are those s_i that do not preclude h_i:

$$\hat{\pi}_i(h_i|\sigma_i)(a_i) = \sum_{\{s_i|s_i \in R_i(h_i),\, s_i(h_i) = a_i\}} \sigma_i(s_i) \Big/ \sum_{\{s_i|s_i \in R_i(h_i)\}} \sigma_i(s_i).$$

At information sets in $H_i \setminus H(\sigma_i)$, we arbitrarily define $\hat{\pi}_i(h_i|\sigma_i)$ by $\hat{\pi}_i(h_i|\sigma_i)(a_i) = \Sigma_{\{s_i|s_i(h_i)=a_i\}}\sigma_i(s_i)$. For a discussion of Kuhn's Theorem, see for example Fudenberg and Tirole (1991, Chapter 3), or Kreps (1990, Chapter 11).

player 1 believes some unobserved common factor has helped determine the play of both of his opponents.[4]

3. SELF-CONFIRMING EQUILIBRIUM AND CONSISTENT SELF-CONFIRMING EQUILIBRIUM

One way to define a *Nash equilibrium* is as a mixed profile σ such that for each $s_i \in \text{support}(\sigma_i)$ there exists beliefs μ_i such that

$$s_i \text{ maximizes } u_i(\cdot, \mu_i), \quad \text{and}$$

$$\mu_i\left[\left\{\pi_{-i} | \pi_j(h_j) = \hat{\pi}_j(h_j | \sigma_j)\right\}\right] = 1 \quad \text{for all } h_j \in H_{-i}.$$

In other words, each player optimizes given his beliefs, and his beliefs are a point mass on the true distribution.

One of the goals of this paper is to introduce the notion of a self-confirming equilibrium, which weakens Nash equilibrium by relaxing the second requirement above. Instead of requiring that beliefs are correct at each information set, self-confirming equilibrium requires only that, for each s_i that is played with positive probability, beliefs are confirmed by the information revealed when s_i and σ_{-i} are played, which we take to be the corresponding distribution on terminal nodes. This corresponds to the idea that the terminal node reached is observed at the end of each play of the game: Learning should not be expected to lead to correct beliefs about play at information sets that are never reached.[5]

DEFINITION 1: Profile σ is a *self-confirming equilibrium* if for each player i, $s_i \in \text{support}(\sigma_i)$ there exists beliefs μ_i such that

(i) s_i maximizes $u_i(\cdot, \mu_i)$, and

(ii) $\mu_i\left[\left\{\pi_{-i} | \pi_j(h_j) = \hat{\pi}_j(h_j | \sigma_j)\right\}\right] = 1$ for all $j \neq i$ and $h_j \in \bar{H}(s_i, \sigma_{-i})$.

Condition (ii) requires that player i's beliefs be concentrated on strategy profiles that coincide with the true distribution at information sets that are reached with positive probability when player i plays s_i. His beliefs about play at other information sets need not be concentrated on a single behavior strategy, and at these information sets his beliefs can incorporate correlation of the kind discussed in the last section. We emphasize that each $s_i \in \text{support}(\sigma_i)$ may be confirmed by a different belief μ_i. In the definition of Nash equilibrium, this flexibility is vacuous, as each μ_i must be exactly correct; the flexibility

[4] We thank Robert Aumann for convincing us of the importance of this kind of subjective correlation.

[5] One way to obtain Hahn's conjectural equilibrium in our setting is to suppose that the players do not observe the terminal node, but rather observe only the market price, that this price depends both on the outputs chosen and on a demand curve by Nature, and that (in contrast to our assumption of the last section) the players do not know the probability distribution over Nature's moves. (If players did know the distribution, then repeated observations of the market price would identify the output of their opponents.)

matters once beliefs are allowed to be wrong. This diversity of beliefs is natural in a learning model with populations of each type of player: different player i's may have different beliefs, either due to different priors or to different observations. Section 6 considers the restriction of "unitary beliefs," which requires that the same beliefs μ_i be used to rationalize each $s_i \in \text{support}(\sigma_i)$.

DEFINITION 2: Profile σ is a *consistent self-confirming equilibrium* if for all players i and each $s_i \in \text{support}(\sigma_i)$ there are beliefs μ_i such that

(i) $\quad s_i$ maximizes $u_i(\cdot, \mu_i)$, and

(ii′) $\quad \mu_i\left[\left\{\pi_{-i} | \pi_j(h_j) = \hat{\pi}_j(h_j|\sigma_j)\right\}\right] = 1$

for all $j \neq i$ and h_j such that $h_j \in H(s_i)$.

In words, self-confirming equilibrium requires that, for each s_i that player i gives positive probability, player i correctly forecasts play at all information sets that will be reached when player i plays s_i and the opponents play σ_{-i}. Consistent self-confirming equilibrium requires further that player i's beliefs be correct at all information sets that could be reached if player i plays s_i. Notice that if there are two players, i and j, either of whom can deviate and cause information set h to be reached, then both players' beliefs about play at h are correct, and in particular are equal to each other. This is why we call the equilibrium consistent.

Note that consistency as we have defined it is a property of the equilibrium strategies, and not of the particular beliefs used to support them: a strategy profile σ is consistent if for each $s_i \in \text{support}(\sigma_i)$ there are *some* beliefs that satisfy the required conditions. Moreover, consistency is less stringent than it might first appear, because Definition 2 is equivalent to the requirement that for each $s_i \in \text{support}(\sigma_i)$, there is a μ_i that satisfies (i) and the apparently weaker condition (ii″):

(ii″) $\quad \mu_i\left[\left\{\pi_{-i} | \pi_j(h_j) = \hat{\pi}_j(h_j|\sigma_j)\right\}\right] = 1$ for all $j \neq i$ and all h_j such that

(a) $\quad h_j \in \overline{H}(s_i', \sigma_{-i})$ for some s_i' and

(b) $\quad h_j \in H(s_i)$.

Condition (ii″) differs from (ii′) in that it concerns only those information sets that can be reached both by player i playing s_i, and by player i playing some s_i' and his opponents sticking with the equilibrium strategies σ_{-i}. Intuitively, the reason these two definitions are equivalent is that no player cares about opponents' play at information sets that he does not expect to be reached regardless of how he plays. To verify this intuition we will define what it means for an information set to be relevant under given beliefs, show that the expected payoffs depend only on beliefs at relevant information sets, and then verify that conditions (ii′) and (ii″) differ only at information sets that are not relevant.

DEFINITION 3: Information set h_j is *relevant to player i given beliefs* μ_i if there exists an $s_i \in S_i$ such that $p_i(h_j|s_i, \mu_i) > 0$.

Let $R_i(\mu_i)$ denote the set of all information sets that are relevant to player i given μ_i. For any subset Q of H_{-i}, $\Pi^Q_{-i} = \times_{h_j \in Q} \Delta(A(h_j))$ may be viewed as a subspace of Π_{-i} corresponding to the information sets in Q, and $\Pi^{\sim Q} = \times_{h_j \in H_{-i}/Q} \Delta(A(h_j))$ is the complementary subspace. Let μ_i^Q denote the marginal distribution that μ_i induces on Q, so that

$$\mu_i^Q(B^Q_{-i}) = \int_{\Pi^{\sim Q}_{-i}} \mu_i(B^Q_{-i}, \pi^{\sim Q}_{-i}) d\pi^{\sim Q}_{-i},$$

where B^Q_{-i} is a (Borel) subset of Π^Q_{-i}.

LEMMA 1: *If* μ_i *and* $\hat{\mu}_i$ *are two distributions on* Π_{-i} *such that* $\mu_i^{R(\mu_i)} = \hat{\mu}_i^{R(\mu_i)}$, *then*

(a) $R(\mu_i) = R(\hat{\mu}_i)$, *and*

(b) $u_i(s_i, \mu_i) = u_i(s_i, \hat{\mu}_i)$ *for all* s_i.

PROOF: (a) If this is false, then either $\exists h_j \in R(\hat{\mu}_i)/R(\mu_i)$ or $\exists h_j \in R(\mu_i)/R(\hat{\mu}_i)$. In the first case, choose a path of actions $\acute{a}$ leading to $x \in h_j$, and such that $p_i(x|s_i, \hat{\mu}_i) > 0$ for some s_i. (Such a path exists because $h_j \in R(\hat{\mu}_i)$.) This path of actions is ordered by the precedence relation in the tree; index the path by t, and let $\iota(t)$ denote the player who moves at the tth step of the path. Let t^* be the lowest index t such that

$$\mu_i\left[\left\{\pi_{-i}|\left(\acute{a}(t) \in \text{support}\left(\pi_{\iota(t)}\left(h_{\iota(t)}\right)\right)\right\}\right] = 0.$$

(Such a t exists because $h_j \notin R(\mu_i)$.)

Then the information set corresponding to t^* is relevant under μ. Since $\mu_i^{R(\mu_i)} = \hat{\mu}_i^{R(\mu_i)}$, and the information set corresponding to $t^* + 1$ is relevant under $\hat{\mu}_i$, this information set must be relevant under μ_i, a contradiction. The proof for the case where $h_j \in R(\mu_i)/R(\hat{\mu}_i)$ is similar.

(b) This follows immediately from (a) and the fact that the expected payoff to any strategy s_i depends only on play at nodes which have positive probability under (s_i, μ_i). *Q.E.D.*

THEOREM 1: *Conditions* (ii′) *and* (ii″) *give the same set of consistent self-confirming equilibria.*

PROOF: Suppose that for each player i and each $s_i \in \text{support}(\sigma_i)$ there are beliefs μ_i that satisfy conditions (i) and (ii″). For each such s_i, define new beliefs $\hat{\mu}_i$ by the conditions

$$\hat{\mu}_i\left[\left\{\pi_{-i}|\pi_j(h_j) = \hat{\pi}_j(h_j|\sigma_j)\right\}\right] = 1 \quad \text{for all } j \neq i \text{ and } h_j \in H(s_i), \quad \text{and}$$

$$\hat{\mu}_i^{\sim H(s_i)} = \mu_i^{\sim H(s_i)}.$$

Since $\hat{\mu}_i$ and μ_i only differ at nodes that are irrelevant under μ_i, the conclusion follows. *Q.E.D.*

We have three reasons for being interested in the additional restrictions imposed by the consistency conditions (ii′) and (ii″). The first is as a tool for understanding the reasons that self-confirming equilibria can fail to be Nash. Inconsistent beliefs, as in Example 1 below, are one such reason, but as we will see, even consistent self-confirming equilibria can have non-Nash outcomes. Second, and relatedly, all self-confirming equilibria are consistent in games with observed deviators.

Finally, the consistency condition (ii″) is of interest because player i will accumulate at least this much information in a learning model if his opponents play each of their strategies sufficiently often, but player i repeatedly plays s_i. That is, consistent self-confirming equilibrium describes a situation in which player i does not experiment himself, but does observe experiments by his opponents. The asymmetry involved in this state of affairs may be puzzling at first sight, but it corresponds to a situation that we think is of interest. Specifically, suppose that there are a large number of player 1's, player 2's, etc., and consider the sort of "independent types" perturbations studied in Fudenberg, Kreps, and Levine (1988), where with probability near 1 a player's payoff function is the u_i originally specified, while with some small probability the payoff functions are different, and in particular every pure strategy of each player is the unique optimal response for some "type" that has positive probability.

In a large population of players, most player i's will be "normal" and have payoff function u_i, while a small proportion will be "crazy." An individual player i who plays many times will then usually meet "normal" opponents, so that his expected payoff will be (approximately) determined by how the "normal" types play. However, a "normal" type of player i who plays a fixed strategy s_i in every period will eventually encounter enough "crazy" player j's, $j \neq i$, to learn the overall distribution of play in the population at every $h_j \in H(s_i)$, conditional on h_j being reached. To show that player i's beliefs will eventually satisfy condition (ii″), it thus suffices to explain why the observed distribution of play corresponds to the play of the normal types at all information sets that satisfy clauses (a) and (b) of the condition.

The observed distribution at an information set h_j will correspond to the play of the normal player j's if, conditional on h_j being reached, player j is (much) more likely to be normal than to be crazy. In particular, this will be the case if h_j is on the equilibrium path, or if it can only be reached if player j doesn't deviate. However, as pointed out by a referee, observed play at h_j need not correspond to the play of the normal type of player j if h_j can be reached by j's deviations. But since the game has perfect recall, if h_j is not on the equilibrium path $\bar{H}(\sigma)$, and if for some deviation $s_j, h_j \in \bar{H}(s_j, \sigma_{-j})$, then for $i \neq j$ there is no s_i' such that $h_j \in \bar{H}(s_i', \sigma_{-i})$. (Otherwise, player j wouldn't be able to distinguish player i's deviations from his own.) Hence, although player i's beliefs need not correspond to the play of the normal player j at every h_j, player i's beliefs do correspond to the play of the normal player j at every h_j that satisfies clauses (a) and (b) of condition (ii″).[6]

Note, though, that if a normal player i always plays s_i, he will not learn the play at information sets not in $H(s_i)$, even though the "crazy" types of player i play strategies other than s_i. This explains why consistent self-confirming equilibrium treats a player's own experiments differently than those of his opponents.[7]

Note also that in a one-shot simultaneous-move game, all information sets are on the path of every profile, so the sets $\overline{H}(s_i, \sigma_{-i})$ are all of H, and condition (ii) requires that beliefs be exactly correct. Hence in these games, all self-confirming equilibria are Nash. In more general games, the self-confirming equilibria can be a larger set, as shown by the examples of the next section.

4. WHEN ARE SELF-CONFIRMING EQUILIBRIA CONSISTENT?

This section begins with an example of a self-confirming equilibrium that is not consistent self-confirming. The example has the property that one player cannot distinguish between deviations by two of his opponents; we show that in the opposite case of "unobserved deviators" any self-confirming equilibrium is a consistent self-confirming equilibrium.

EXAMPLE 1 (Fudenberg-Kreps): In the three payer game illustrated in Figure 1, player 1 moves first. If he plays A_1, player 2 moves next; if he plays D_1, player 3 gets the move. If player 2 gets the move, he can either play A_2, which ends the game, or play D_2, which gives the move to player 3. The key feature of the game is that if player 3 gets the move, he cannot tell whether player 1 played D_1, or player 1 played A_1 and player 2 played D_2.

Fudenberg and Kreps (1988) use this game to show that learning need not lead to Nash equilibrium even if players are long-lived. Suppose that player 1 expects player 3 to play R and player 2 expects player 3 to play L. Given these beliefs, it is optimal for players 1 and 2 to play A_1 and A_2. Moreover, (A_1, A_2)

[6] Our discussion here glosses over the fact that for any small but noninfinitesmal ε' probability of the crazy types, the observed distribution will approximate the play of the normal types but will not exactly equal it, so that the beliefs would be within some ε of the play of the normal types at the information sets covered by (a) and (b). We believe that the standard sort of upper hemicontinuity argument would show that the resulting "ε-consistent self-confirming equilibria" converge to consistent self-confirming equilibria as ε goes to 0, but we have not written out the details. Kalai and Lehrer (1991a) provide a formal limit argument for a closely related sort of approximately correct beliefs.

[7] A second motivation yields a concept related to, but weaker than, the consistency condition. This motivation supposes that *all* player i's have the stage game payoffs u_i, that most of them have low discount factors, but that a small proportion have discount factors near 1. Then these patient players will do every worthwhile experiment infinitely often, and in the process the nonpatient players will get a chance to learn about some information sets off of the equilibrium path. However, we show by example in Fudenberg and Levine (1991) that even a completely patient player may optimally choose never to play some of his strategies, so this model will not lead a myopic player i to have to correct beliefs at all information sets in $H(s_i)$.

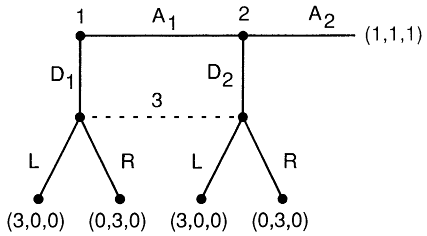

FIGURE 1

is a self-confirming equilibrium. However, it is not a Nash equilibrium outcome: Nash equilibrium requires players 1 and 2 to have the same (correct) beliefs about player 3's play, and if both have the same beliefs, at least one of the players must choose D. (If the beliefs assign probability more than $1/3$ to L and 2 plays A_2, then 1 plays D_1, while if the beliefs assign probability more than $1/3$ to R and 1 plays A_1 then 2 plays D_2.)

When this example has been presented in seminars, the following question has frequently been raised: Shouldn't player 2 revise his beliefs about player 3 in the direction of 3 playing R when he sees player 1 play A_1? And, in the spirit of the literature on the impossibility of players "agreeing to disagree" (Aumann (1976), Geanakoplos and Polemarchakis (1982), and so forth) shouldn't players 1 and 2 end up with the same beliefs about player 3's strategy?

Our response is to note that, while this sort of indirect learning could occur in our model, it need not do so. First, the indirect learning supposes that players know (or have strong beliefs about) one another's payoffs, which is consistent with our model but is not necessarily the case. Second, even if player 2 knows player 1's payoffs, and hence is able to infer that player 1 believes player 3 will play R, it is not clear that this will lead player 2 to revise his own beliefs. It is true that player 2 will revise his beliefs if he views the discrepancy between his own beliefs and player 1's as due to information that player 1 has received but player 2 has not, but player 2 might also believe that player 1 has no objective reason for his beliefs, but has simply made a mistake. The "agreeing to disagree" literature ensures all differences in beliefs are attributable to objective information by supposing that the players' beliefs are consistent with Bayesian updating from a common prior distribution. But when equilibrium is interpreted as the result of learning, the assumption of a common prior is inappropriate. Indeed, the question of whether learning leads to Nash equilibrium can be rephrased as the question of whether learning leads to common *posterior* beliefs starting from arbitrary priors.

While (A_1, A_2) in Example 1 is a self-confirming equilibrium (with unitary beliefs) it is not a consistent self-confirming equilibrium. To see this, note that player 3's information set can be reached if player 1 sticks to his equilibrium action A_1 and player 2 deviates, and also if player 2 sticks to A_2 and player 1 deviates. Hence condition (ii′) requires that both player 1 and player 2 have correct beliefs at h_3, and in particular both of their beliefs must be the same. The reason that non-Nash outcomes can be obtained by allowing players 1 and 2 to disagree about play at h_3 is that either of them can deviate and cause h_3 to be reached. This particular source of non-Nash behavior is absent in games that satisfy the following definition.

DEFINITION 4: A game has *observed deviators* if for all players i, all strategy profiles s and all deviations $s'_i \neq s_i$, $h \in \overline{H}(s'_i, s_{-i}) \setminus \overline{H}(s)$ implies that there is no s'_{-i} with $h \in \overline{H}(s_i, s'_{-i})$.

In words, this definition says that if some deviation from σ by player i leads to a new information set h, that is not on the equilibrium path, then the information set cannot be reached if player i plays s_i. Games of perfect information satisfy this condition, as do repeated games with observed actions. More generally, the condition is satisfied by all "multi-stage games with observed actions," meaning that the extensive form can be parsed into "stages" with the properties that the beginning of each stage corresponds to a proper subgame (Selten (1975)), and that within each stage all players move simultaneously.[8] The following result shows that the condition is also satisfied in all two-player games of interest:

LEMMA 2: *Every two-player game of perfect recall has observed deviators.*

PROOF: Suppose to the contrary that there exists a profile $s = (s_1, s_2)$, and information set h such that $h \notin \overline{H}(s_1, s_2)$, but $h \in \overline{H}(s_1, s'_2)$ for some s'_2 and $h \in \overline{H}(s'_1, s_2)$ for some s'_1. If $h \in H_1$, then player 1 cannot distinguish between s_1 and s'_1, while $h \in H_2$ implies that player 2 cannot distinguish between s_2 and s'_2.
Q.E.D.

THEOREM 2: *In games with observed deviators, self-confirming equilibria are consistent self-confirming.*

PROOF: The idea is simply that in games with observed deviators, the only information sets covered by condition (ii″) are those that are on the equilibrium path. Suppose that σ is a self-confirming equilibrium, and for each player i and $s_i \in \text{support}(\sigma_i)$ let $\mu_i(\cdot\,; s_i)$ be beliefs satisfying conditions (i) and (ii) of Definition 1. We claim that these same beliefs satisfy condition (ii″) of Definition 2. If

[8] See Fudenberg and Tirole (1991) for a more detailed explanation of multi-stage games; we introduced the definition in Fudenberg and Levine (1983). Note that the extensive form in Example 2 below is not a multi-stage game with observed actions, but is a game with observed deviators. Moreover, splitting player 1's information into two consecutive choices, the first one being A or $\sim A$, yields a multi-stage game with observed actions that has the same reduced strategic form and the same set of self-confirming equilibria. This emphasizes that from the viewpoint of self confirming equilibria, observed deviators is the more fundamental property. In a private communication, Battigalli has shown that observed deviators is equivalent to the condition that the set of strategy profiles that reaches any given information set be a cross-product; this condition is called "strategic decomposability" in Battigalli (1991).

not, there must be a player i, $s_i \in \text{support}(\sigma_i)$, and h_j where (ii′) is violated for beliefs $\mu_i(\cdot\,; s_i)$; since σ is self-confirming, this h_j is not in $\bar{H}(s_i, \sigma_{-i})$. Hence we have that (a) $h_j \in \bar{H}(s_i', \sigma_{-i}) \setminus \bar{H}(s_i, \sigma_{-i})$ for some s_i' and (b) $h_j \in H(s_i)$. Disregard (b) for the moment and consider the implications of (a). Since $h_j \in \bar{H}(s_i', \sigma_{-i})$ implies that $h_j \in \bar{H}(s_i', s_{-i})$ for some $s_{-i} \in \text{support}(\sigma_{-i})$, and $h_j \notin \bar{H}(s_i, \sigma_{-i})$ implies that $h_j \notin \bar{H}(s_i, s_{-i})$ for all $s_{-i} \in \text{support}(\sigma_{-i})$, there is an $s_{-i} \in \text{support}(\sigma_{-i})$ such that $h_j \in \bar{H}(s_i', s_{-i}) \setminus \bar{H}(s_i, s_{-i})$. Hence observed deviators implies that there is no s_{-i}' with $h_j \in \bar{H}(s_i, s_{-i}')$. This contradicts the hypothesis (b) that $h_j \in H(s_i)$. *Q.E.D.*

COROLLARY: *Self-confirming equilibria are consistent self-confirming in two-player games.*

5. CONSISTENT SELF-CONFIRMING EQUILIBRIUM AND EXTENSIVE-FORM CORRELATED EQUILIBRIUM

Even consistent self-confirming equilibria need not be Nash. There are two reasons for this difference. First, consistent self-confirming equilibrium allows a player's uncertainty about his opponents' strategies to be correlated, while Nash equilibrium requires that the beliefs be a point mass on a behavior strategy profile. Second, consistent self-confirming equilibrium allows each strategy that a player assigns positive probability to be a best response to different beliefs. Our first example displays the consequences of correlated uncertainty.

EXAMPLE 2 [Untested Correlation]: In the game in Figure 2, player 1 can play A, which ends the game, or play L_1, M_1, or R_1, all of which lead to a simultaneous-move game between players 2 and 3, neither of whom observes player 1's action. In this game, A is a best response to the correlated distribution $p(L_2, L_3) = p(R_2, R_3) = 1/2$. Thus if player 1's prior beliefs are either that 2 and 3 always play L, or that they always play R, then player 1's best response is to play A, and so A is the outcome of a self-confirming equilibrium.

However, we claim that A is not a best response to any strategy profile for players 2 and 3. Verifying this is straightforward but tedious: Let p_2 and p_3 be the probabilities that players 2 and 3, respectively, assign to L_2 and L_3. In order

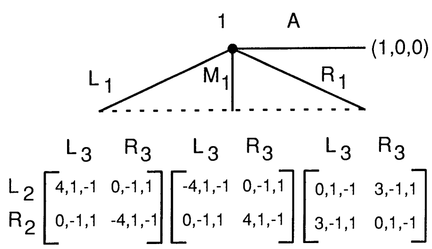

FIGURE 2

for A to be a best response, the following 3 inequalities must be satisfied:

$$4[p_2p_3-(1-p_2)(1-p_3)] \leqslant 1, \quad \text{or} \quad p_2+p_3 \leqslant 5/4, \tag{5.1}$$

$$4[-p_2p_3+(1-p_2)(1-p_3)] \leqslant 1, \quad \text{or} \quad p_2+p_3 \geqslant 3/4, \quad \text{and} \tag{5.2}$$

$$p_2(1-p_3)+(1-p_2)p_3 \leqslant 1/3. \tag{5.3}$$

We will show that when constraints (5.1) and (5.2) are satisfied, (5.3) cannot be. For any $p_2 \leqslant 1/2$, the left-hand side of (5.3) is minimized when p_3 is as small as possible, that is, for $p_3 = 3/4 - p_2$. The minimized value is $2p_2^2 - 3/2p_2 + 3/4$, and this expression is minimized over p_2 at $p_2 = p_3 = 3/8$; at this point the left-hand side of (5.3) equals $30/64 > 1/3$. The case $p_2 > 1/2$ is symmetric.

We stress that the correlation in this example need not describe a situation in which player 1 believes that players 2 and 3 actually correlate their play. To the contrary, player 1 might be certain that they do not do so, and that she could learn which (uncorrelated) strategy profile they are using by giving them the move a single time. These competing explanations for the correlation—call them "objective" correlation and "subjective" correlation—cannot be distinguished in a static, reduced-form model of the kind considered in this paper. However, our (1991) paper on steady-state learning shows that the non-Nash outcome of Example 2 can be the steady state of a learning process where players are certain that their opponents' actual play is an uncorrelated behavior profile.

As we indicated above, there is another way that consistent self-confirming equilibria can fail to be Nash: The self-confirming concept allows each strategy that a player assigns positive probability to be a best response to different beliefs. This possibility allows for non-Nash play even in two-player games. The most immediate consequence of these differing beliefs is a form of convexification, as in the following example.

EXAMPLE 3 [Public Randomization]: In the game in Figure 3, player 1 can end the game by moving L or he can give player 2 the move by choosing R. Player 1 should play L if he believes 2 will play D, and should play R if he believes 2 will play U. If player 1 plays R with positive probability, player 2's unique best response is to play U, so there are two Nash equilibrium outcomes, (L) and (R,U). The mixed profile $((1/2L, 1/2R), U)$ is a self-confirming equilibrium whose outcome is a convex combination of the Nash outcomes:

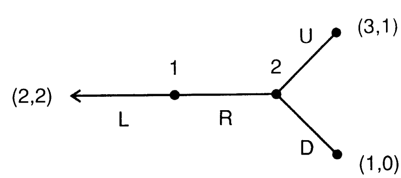

FIGURE 3

Player 1 plays L when he expects player 2 to play D, and R when he expects 2 to play U, and when he plays L his forecast of D is not disconfirmed. (Moreover, this equilibrium is clearly independent.)

Besides convexifying the set of Nash equilibria, the additional randomization by one player may lead to entirely different play by another. Example 3 is a two-stage game, with only one player moving in each stage. Our next example presents a more complicated situation, in which players 1 and 2 move simultaneously in the first stage, followed by player 2 moving alone in the second stage. As in Example 3, player 1 has two possible beliefs about player 2's second-stage play, which leads player 1 to randomize in the first stage, even though he cannot play a mixed strategy in Nash equilibrium. Moreover, because both players move in the first stage, this randomization leads player 2 to play a first-period action that is not a best response to any pure strategy of player 1, and hence must have probability zero in any Nash equilibrium.

EXAMPLE 4: The extensive form shown in Figure 4 corresponds to a two-stage game: In the first stage, players 1 and 2 play simultaneously, with player 1 choosing U or D and player 2 choosing L, M, or R. Before the second stage, these choices are revealed. In the second stage, only player 2 has a move, choosing between R ("Reward") costing both players 0, and P ("Punish") costing both players 10. The payoffs are additively separable between periods.

We claim first that in any Nash equilibrium of this game, player 1 must play a pure strategy and player 2 must play M in stage 1 with probability zero. If both U and D have positive probability, then following every outcome of the first stage that occurs with positive probability, player 2 must choose R. But then player 1 would play U with probability 1. We conclude that player 1 must play a pure strategy, and consequently player 2 cannot play M.

However, player 2 can play M with probability 1 in a self-confirming equilibrium: Let player 1's strategy be $\sigma_1 = (\frac{1}{2}U, \frac{1}{2}D)$, and let player 2's strategy σ_2 be "play M in the first stage and play R in the second stage regardless of the first-period outcome." Player 2's strategy is a best response to the strategy σ_1 that player 1 is actually playing, and $U \in \text{support}(\sigma_1)$ is a best response to σ_2. The strategy $D \in \text{support}(\sigma_1)$ is not a best response to σ_2, but it is a best response to the belief that player 2 will play R if player 1 plays D and P if

Stage 1

	L	M	R
U	13,15	13,14	13,11
D	12,11	12,14	12,15

Stage 2

R	P
0,0	-10,-10

FIGURE 4

player 1 plays U; and when player 1 plays D his forecast of what would have happened if he had played U is not disconfirmed.

Although consistent self-confirming equilibria need not be Nash equilibria, they are a special case of another equilibrium concept from the literature, namely the extensive-form correlated equilibria defined by Forges (1985). These equilibria, which are only defined for games whose information sets are ordered by precedence (the usual case), are the Nash equilibria of an expanded game where an "autonomous signalling device" is added at every information set, with the joint distribution over signals independent of the actual play of the game and common knowledge to the players, and the player on move at each information set h is told the outcome of the corresponding device before he chooses his move.[9] Extensive-form correlated equilibrium includes Aumann's (1974) correlated equilibrium as the special case where the signals at information sets after stage 1 have one-point distributions and so contain no new information. The possibility of signals at later dates allows the construction of extensive-form correlated equilibria that are not correlated equilibria, as in Myerson (1986). Another example is based on the self-confirming equilibrium we constructed in Example 4.

EXAMPLE 4 REVISITED: We construct an extensive-form correlated equilibrium with the same distribution over outcomes as the self-confirming equilibrium in Example 4. The first-stage private signals describe play in that stage: There is a probability $1/2$ of the signals (U, M) and (D, M) in stage 1. The strategies in stage 1 are to play the recommended action. The second-stage pubic signal takes on two values, U and D. The strategy for player 2 in stage 2 is to play P if player 1 played U and the second signal is D, and R otherwise. The second-stage public signal is perfectly correlated with player 1's first-stage private signal. Let us check that it is a Nash equilibrium for the players to use the strategies their signals recommend: Since player 1's signal reveals whether or not he will be punished for playing U, player 1 finds it optimal to obey his signal. Player 2's first signal is uninformative about player 1's stage 1 play, and so player 2 expects player 1 to randomize $(1/2)-(1/2)$ in the first stage and thus plays M. Player 2 cannot improve on the recommended strategies in the second stage because he is only told to punish U when player 1's first signal was to play D, and if player 1 obeys his signal this will not occur. The role of the second signal is to tell player 2 when to punish player 1 without revealing player 1's play at the beginning of the first stage; if player 1's play was revealed at this point this would remove player 2's incentive to play M. Note that while the extensive-form correlated equilibrium and the self-confirming equilibrium have the same distribution over outcomes, they involve different distributions over

[9] Forges shows, in the spirit of the revelation principle, that it suffices to work with a smaller set of signalling devices. She also defines "communications equilibria," which allow the players to send messages in the course of play that influence subsequent signals.

strategies: In a self-confirming equilibrium, if player 1 mixes between U and D, then player 2 must respond to both U and D with R; player 1 sometimes plays D because he incorrectly believes 2 will respond to U with P. In an extensive-form correlated equilibrium, each player's predictions about his opponents' strategies are on average correct, so if player 1 sometimes believes that player 2 responds to U with P, then player 2 must assign positive probability to a strategy that does so.

We should point out that the extensive-form nature of the correlations is required, as player 2 cannot play M with probability 1 in a correlated equilibrium of the usual kind, which allows signals only at the initial stage. To see this, suppose that probability distribution $p \in \Delta(S)$ is a correlated equilibrium. If 1 plays U with probability 1, 2 must play L, while if he plays D with probability 1, 2 must play R. So in this case the probability of M is zero. So now suppose that both U and D have positive probability, and that player 2 plays M with probability 1. Since player 1 is willing to play D, which has a lower first-period payoff, he must expect that P has positive probability conditional on the first-period outcome (U, M). But since P gives player 2 a lower payoff than R does, and (U, M) has positive probability by assumption, player 2 cannot follow (U, M) with a positive probability of P, and hence there is no correlated equilibrium in which the probability of M is 1. Moreover, the probability of M is bounded away from 1 in all correlated equilibria, because the set of correlated equilibria is closed.

THEOREM 3: *For each consistent self-confirming equilibrium of a game whose information sets are ordered by precedence, there is an equivalent extensive-form correlated equilibrium, that is, one with the same distribution over terminal nodes.*

PROOF: Let σ be consistent self-confirming, and for each $s_i \in \operatorname{support} \sigma_i$, let $\mu_i(s_i)$ be beliefs satisfying conditions (i) and (ii′) of Definition 2. We now expand the game by adding an initial randomizing device whose realization is partially revealed as private information at various information sets. A realization of this device is an I-vector with the ith component a pair (s_i, π^i_{-i}) with $s_i \in S_i$ and $\pi^i_{-i} = (\pi_j)_{j \neq i} \in \Pi_{-i}$. The s_i follow the probability distribution σ (and in particular s_i and s_j are independent for $i \neq j$). The distribution of π^i_{-i} conditional on s is $\mu_i(s_i)$. Intuitively, profile π^i_{-i} is the way player i expects to be "punished" if he deviates from strategy s_i.

Initially each player i is told s_i. Subsequent revelations also depend upon s. At information sets on the path of s, $h \in \overline{H}(s)$, no additional information is revealed. At information sets that can be reached only if two or more players deviate from s no information is revealed. If $h_j \in \overline{H}(s'_i, s_{-i}) \setminus \overline{H}(s)$, so h_j is reached by player i's deviation, and $j \neq i$, then player j is told $\pi^i_j(h)$.

If $h_j \in \overline{H}(s'_i, s_{-i}) \setminus \overline{H}(s)$ and $h_j \in \overline{H}(s'_k, s_{-k}) \setminus \overline{H}(s)$, then $h_j \in \overline{H}(s_i)$ and $h_j \in \overline{H}(s'_i, s_{-i})$, so that in a consistent self-confirming equilibrium player i's beliefs about play at h_j are correct. A similar argument shows that player k's beliefs are correct as well, so that the two beliefs are equal. It follows that $\pi_j^i(h_j) = \pi_j^k(h_j)$ for $j \neq i, k$, so only one distinct signal is received by j.

Now consider the strategy profile $\hat{s}$ for the expanded game in which each player j plays s_j except at information sets (in the expanded game), where the signal $\pi_j^i(h_j)$ is received. At such information sets j plays according to $\pi_j^i(h_j)$.

By construction, $\hat{s}$ induces the same distribution over terminal nodes as σ does. If player i's opponents follow $\hat{s}$, player i will never receive an additional message, so player i is willing to play $\pi_i^j(h_i)$ at the probability-zero information sets where player j deviates and i is told $\pi_i^j(h_i)$. Moreover, given the initial message s_i, opponents' play is drawn from $\mu_i(s_i)$, and s_i is a best response to $\mu_i(s_i)$ from condition (i) in the definition of self-confirming equilibrium. Hence $\hat{s}$ is a Nash equilibrium of the expanded game. *Q.E.D.*

COROLLARY: *In games with observed deviators, every self-confirming equilibrium outcome is the outcome of an extensive form correlated equilibrium.*

Theorem 3 fails if the hypothesis of consistency is dropped. To see this, let us return to the game of Example 1. Since each player has only one information set, extensive-form correlated equilibrium is equivalent to correlated equilibrium; since there is no Nash equilibrium in which players 1 and 2 play (A_1, A_2), the probability of (A_1, A_2) is bounded away from 1 in every correlated equilibrium. The reason is that the common prior supposed by correlated equilibrium rules out situations in which player 1 always believes player 3 is likely to play R and player 2 always believes player 3 is likely to play L.[10] As shown by Theorem 3, consistent self-confirming equilibrium ensures that any two players' beliefs are compatible with a common prior, at least at the information sets that matter to both of them.

Note also that not all outcomes of extensive-form correlated equilibria are the outcomes of consistent self-confirming equilibria: Because self-confirming equilibrium supposes that players choose their actions independently, the equilibrium path of play must be uncorrelated, so not even every correlated equilibrium outcome can be attained. In particular, since publicly observed signals can lead to correlated play, consistent self-confirming equilibria cannot generate every outcome that can be supported using public randomizing devices. This suggests that it might be possible to find an interesting and tighter characterization of consistent self-confirming equilibrium; we have not been able to do so.

6. UNITARY BELIEFS AND NASH EQUILIBRIUM

So far we have seen three ways in which self-confirming equilibria can fail to be Nash. First, two players can have inconsistent beliefs about the play of a third, as in Example 1. Second, a player's subjective uncertainty about his opponents' play may induce a correlated distribution on their actions, even

[10] However, there is a correlated equilibrium in which (A_1, A_2) has probability $1/4$.

though he knows that their actual play is uncorrelated; this was the case in Example 2. Finally, the fact that each player can have heterogeneous beliefs—that is, different beliefs may rationalize each $s_i \in \text{support}(\sigma_i)$—introduces a kind of extensive-form correlation. Theorem 2 showed that in games with observed deviators, self-confirming equilibria are consistent, thus precluding the kind of non-Nash situation in Example 1. Theorem 4 shows that the combination of off-path correlation and heterogeneous beliefs encompass all other ways that self-confirming equilibria can fail to be Nash.

DEFINITION 5: A self-confirming equilibrium σ has *independent beliefs* if for all players i and all $s_i \in \text{support}(\sigma_i)$, the associated beliefs μ_i satisfy

$$\mu_i\Big(\bigtimes_{j \neq i} \overline{\Pi}_j\Big) = \bigtimes_{j \neq i} \mu_i\big(\overline{\Pi}_j\big) \quad \text{for all (measurable) } \overline{\Pi}_j \subseteq \Pi_j.$$

A self-confirming equilibrium has *unitary beliefs* if for each player i the same beliefs μ_i can be used to rationalize every $s_i \in \sigma_i$. That is, in Definition 1 (or 2), we replace "[for all] $s_i \in \text{support}(\sigma_i)$ $\exists \mu_i$" with "$\exists \mu_i$ for all $s_i \in \text{support}(\sigma_i)$."

THEOREM 4: *Every consistent self-confirming equilibrium with independent, unitary beliefs is equivalent to a Nash equilibrium.*

OUTLINE OF PROOF: Fix a consistent self-confirming equilibrium σ with independent, unitary beliefs. Thus for each player i, there is a μ_i such that conditions (i) and (ii′) of Definition 2 are satisfied for all $s_i \in \text{support}(\sigma_i)$, and μ_i is a product measure on Π_{-i}.

We will construct a new strategy profile σ' that is a Nash equilibrium and has the same distribution on terminal nodes as σ does. The idea is simply to change the play of all players $j \neq i$ to that given by player i's beliefs at all the information sets that can be reached if i unilaterally deviates from σ. The unitary beliefs condition implies that "player i's beliefs" are a single object. The requirement that the equilibrium is consistent ensures that this process is well-defined, as if deviations by two distinct players can lead to the same information set, then their beliefs at that information set are identical. Finally, the condition of independence says that player i's beliefs μ_i correspond to the behavior strategy profile $\pi'_{-i}(\cdot\,|\sigma'_{-i})$ corresponding to σ'. The details are given in the Appendix.

COROLLARY: *In two-player games, every self-confirming equilibrium with unitary beliefs is Nash.*[11]

[11] This is proved directly in Fudenberg and Kreps (1993), and in Battigalli (1987) for the case of point-valued beliefs.

7. GENERALIZATIONS AND EXTENSIONS

Self-confirming equilibrium describes a situation in which players know their own payoff functions, the distribution over Nature's moves, and the strategy spaces of their opponents; the only uncertainty players have is about which strategies their opponents will play. Moreover, as we explained in the introduction, the assumption that players' beliefs are correct along the path of play implicitly supposes that players observe the terminal node of the game at the end of each play. Thus it is of some interest to consider how the assumptions might be relaxed.

It would be interesting to see a characterization of the analog of self-confirming equilibrium for the case in which each player's end-of-stage information is precisely his own payoff; the key would be finding a tractable description of how much information the payoffs convey. Another interesting case is that of games of incomplete information, with the assumption that each player observes the entire sequence of play and his own type, but not the types of his opponents. We conjecture that if each player's payoff depends on the sequence of actions played and his own type, but not on the types of his opponents, and if all types of each player i have the same "physical" extensive form, so that the incomplete-information game is an "elaboration" of an underlying complete-information game in the sense of Fudenberg, Kreps, and Levine (1988), then the set of self-confirming equilibria is the same whether or not the opponents' types are observed at the end of each round.

The other informational assumptions of self-confirming equilibrium can be relaxed as well. It is easy to generalize self-confirming equilibrium to allow for players who do not know the distribution of Nature's moves. In this context our convention that all of Nature's moves are at the beginning of the tree is not innocuous: When combined with the maintained assumption that players observe the terminal node of the tree, this convention implies that players will learn the entire distribution of Nature's moves regardless of the actions chosen by players other than Nature. This assumption is natural if the uncertainty modeled by Nature's moves is not directly related to the players' actions, as for example if Nature's move is the weather, for then we would expect both that the weather is observed regardless of how the game is played, and that the distribution over weather is not influenced by the players' actions. However, if Nature's move models uncertainty about the productivity of a new technology, we would not expect the move to be observed unless the technology is used, so that players might persistently maintain incorrect beliefs about Nature's move. To model such situations we would want to embed Nature's move in the tree after the players' decision nodes.

Yet another extension is to games where players do not know the extensive form of the game, and in particular do not know the information that their opponents possess when choosing their actions. At a formal level this uncertainty can be represented as a game of incomplete information, that is by a combination of moves by Nature. If players observe only one another's actions, but do not observe Nature's moves, then players may maintain incorrect beliefs about other players' strategy spaces. For this reason, it appears that the resulting static equilibrium concept may be quite different from the one we have developed here.

This paper has stressed the importance of the players' information about one another's strategies. In applied work, it is also important to consider the econometrician's information about the players' strategies, which may be different. Bresnahan and Reiss (1991) show that this difference can have serious implications when observed play is assumed to correspond to a Nash equilibrium; the implications for estimating self-confirming equilibrium may be even more striking.

Dept. of Economics, M.I.T., Cambridge, MA 02139, U.S.A.
and
Dept. of Economics, U.C.L.A., Los Angeles, CA 90024-1477, U.S.A.

Manuscript received August, 1990; final revision received August, 1992.

APPENDIX

Proof of Theorem 4

We construct σ' in two steps. First, for each player i we construct a profile π^i_{-i} of behavior strategies for player i's opponents, such that (i) for all strategies $s_i, u_i(s_i, \mu_i) = u_i(s_i, \pi^i_{-i})$, where the second argument on the right-hand side of this equality should be read as beliefs corresponding to a point mass on π^i_{-i}, and (ii) π^i_{-i} agrees with σ at all information sets in $\bar{H}(\sigma)$. That such a π^i_{-i} exists will follow from the assumptions that beliefs are independent, and that the game has perfect recall.

The procedure just described constructs $(I-1)$ strategies $\{\pi^i_k\}_{i \neq k}$ for each player k. We now use these strategies to construct a single strategy profile π'. We do this by specifying that $\pi'_k(h_k) = \pi^i_k(h_k)$ if there exists an s_i such that $h_k \in H(s_i, \sigma_{-i})$. The assumption that the equilibrium is consistent will ensure that this construction is well defined: If there exists an s_i such that $h_k \in H(s_i, \sigma_{-i})$, and there exists an s_j such that $h_k \in H(s_j, \sigma_{-j})$, then, as in the proof of Theorem 3, both player i's and player j's beliefs must be correct, and so $\pi^i_k(h_k)$ will equal $\pi^j_k(h_k)$. Having constructed π' and its equivalent mixed strategy representation σ', we then verify that it is a Nash equilibrium.

Step 1: Constructing the π^i_{-i}: For each player k, let $\hat{\sigma}_k$: $\Pi_k: \to \Sigma_k$ be the map that constructs a mixed strategy that is equivalent to the given behavior strategy π_k by the rule

$$\hat{\sigma}_k(\pi_k)(s_k) = \prod_{h_k \in H_k} \pi_k(h_k)(s_k(h_k)).$$

For a profile π_{-i} of behavior strategies of player i's opponents, define an equivalent mixed strategy profile by $\hat{\sigma}_{-i}(\pi_{-i}) = \times_{k \neq i} \hat{\sigma}_k(\pi_k)$, where $\times$ denotes the product of measures. Let $\mu_{i,k}$ be the marginal distribution of player i's beliefs over player k's behavior strategies.

Now for each i, define a probability distribution on S_{-i} by

$$\sigma^i_{-i}(s_{-i}) = \int_{\Pi_{-i}} \hat{\sigma}_{-i}(\pi_{-i})(s_{-i}) \mu_i[d\pi_{-i}]$$

Using the definition of $\hat{\sigma}_{-i}$, we have

$$\sigma^i_{-i}(s_{-i}) = \int_{\Pi_{-i}} \left(\prod_{k \neq i} \hat{\sigma}_k(\pi_k)(s_k) \right) \mu_i[d\pi_{-i}]$$

$$= \prod_{k \neq i} \int_{\Pi_k} \hat{\sigma}_k(\pi_k)(s_k) \mu_{i,k}[d\pi_k].$$

If we now define $\sigma_k^i \in \Sigma_k$ by

$$\sigma_k^i(s_k) = \int_{\Pi_k} \hat{\sigma}_k(\pi_k)(s_k)\mu_{i,k}[d\pi_k],$$

it is clear that $\sigma_{-i}^i(s_{-i}) = \Pi_{k \neq i}\sigma_k^i(s_k)$, so that $\sigma_{-i}^i \in \Sigma_{-i}$. Let $\hat{\pi}_{-i}$: $\Sigma_{-i} \to \Pi_{-i}$ be the map that assigns to each profile σ_{-i} of mixed strategies for player i's opponents the equivalent behavior strategy profile

$$\hat{\pi}_{-i}(\sigma_{-i}) = \bigtimes_{k \neq i} \hat{\pi}_k(\cdot\,|\sigma_k)$$

and let

$$\pi_{-i}^i = \hat{\pi}_{-i}^i(\sigma_{-i}^i) = \bigtimes_{k \neq i} \hat{\pi}_k(\cdot\,|\sigma_k^i)$$

be behavior strategy profile equivalent to σ_{-i}^i.

Now consider replacing player i's possibly diffuse beliefs about his opponents' play by the single strategy profile π_{-i}^i. By construction, $u_i(s_i, \mu_i) = u_i(s_i, \pi_{-i}^i)$ for all s_i: Player i's expected utility to strategy s_i when his beliefs are μ_i is the same as his expected utility to s_i when his beliefs are a point mass on π_{-i}^i. It remains to show that $\pi_j^i(h_j) = \hat{\pi}_j(h_j|\sigma_j)$ for all $j \neq i$ and all $h_j \in \bar{H}(\sigma)$, i.e. that π_{-i}^i coincides with the true distribution along the path of play. Since player i's beliefs μ_i assign probability one to the subset of Π_{-i} for which this is true, it is intuitive that the profile π_{-i}^i obtained by integrating μ_i should have the same property. Verifying this requires working through the formulae for σ_{-i}^i and $\hat{\pi}_k(\sigma_k)$. Fortunately, since

$$\pi_k^i = \hat{\pi}_k(\cdot\,|\sigma_k^i) = \int_{\Pi_k} \hat{\sigma}_k(\pi_k)\mu_{i,k}[d\pi_k]$$

depends only on player i's beliefs about player k, the required verification is exactly the same as that given in Fudenberg and Kreps (1993) for the two-player case, so their calculations show that $\pi_j^i(h_j) = \hat{\pi}_j(h_j|\sigma_j)$ for all $j \neq i$ and all $h_j \in \bar{H}(\sigma)$.

Step 2: We now construct the profile π' by specifying that $\pi_k'(h_k) = \pi_k^i(h_k)$ if there exists an s_i such that $h_k \in \bar{H}(s_i, \sigma_{-i})$. As we remarked earlier, the assumption that the equilibrium is consistent ensures that this construction is well defined: If there exists an s_i such that $h_k \in \bar{H}(s_i, \sigma_{-i})$, and there exists an s_j such that $h_k \in \bar{H}(s_j, \sigma_{-j})$, then $h_k \in H(\sigma_j) \cap H(\sigma_i)$, and so $\pi_k^i(h_k)$ will equal $\pi_k^j(h_k)$.

Note that if for some $k \neq i$ and h_k, $\pi_k'(h_k) \neq \pi_k^i(h_k)$, then h_k cannot be reached unless some player $j \neq i$ deviates from σ. Hence, for all i, s_i, and $k \neq i$,

$$u_i(s_i, \pi_{-i}') = u_i(s_i, \pi_{-i}^i) = u_i(s_i, \mu_i).$$

Moreover, since each $\pi_k^i(h_k)$ agrees with $\hat{\pi}(h_k|\sigma)$ at all information sets in $\bar{H}(\sigma)$, the same is true of π'.

Finally, we check that π' is a Nash equilibrium. Let π denote the behavior strategy profile induced by σ. Because σ is a unitary self-confirming equilibrium, for all $\hat{\pi}_i$

$$u_i(\pi_i, \mu_i) \geqslant u_i(\hat{\pi}_i, \mu_i),$$

and because player i's expected payoff to any strategy is the same under μ_i and π_{-i}', this implies

$$u_i(\pi_i, \pi_{-i}') \geqslant u_i(\hat{\pi}_i\, \pi_{-i}').$$

Since π_i' agrees with π_i at every information set with positive probability under π (and hence at every information set with positive probability under π') we conclude that

$$u_i(\pi_i', \pi_{-i}') \geqslant u_i(\hat{\pi}_i, \pi_{-i}'). \qquad \textit{Q.E.D.}$$

REFERENCES

AUMANN, R. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

——— (1976): "Agreeing to Disagree," *Annals of Statistics*, 4, 1236–1239.

BATTIGALLI, P. (1987): "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali," unpublished dissertation, Università Commerciale "L. Bocconi", Milano.

——— (1988): "Implementable Strategies, Prior Information and the Problem of Credibility in Extensive Games," *Rivista Internazionale di Scienze Economiche e Commerciali*, 35, 705–733.

——— (1991): "Strategic Independence, Generally Reasonable Extended Assessments, and Consistent Assessments," Università Commerciale L. Bocconi, mimeo.

BATTIGALLI, P., AND P. GUATOLI (1988): "Conjectural Equilibrium," mimeo.

BRESNAHAN, T., AND P. REISS (1991): "Empirical Models of Discrete Games," mimeo, forthcoming in the *Journal of Econometrics*.

FORGES, F. (1985): "An Approach to Communications Equilibrium," *Econometrica*, 54, 1375–1386.

FUDENBERG, D., AND D. M. KREPS (1988): "A Theory of Learning, Experimentation and Equilibrium in Games," mimeo, Stanford.

——— (1991): "Learning and Equilibrium in Games, Part I: Strategic-Form Stage Games," mimeo.

——— (1993): "Learning and Equilibrium in Games, Part II: Extensive-Form Games and Nash Equilibrium," in preparation.

FUDENBERG, D., AND D. K. LEVINE (1983): "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games," *Journal of Economic Theory*, 31, 251–268.

——— (1990): "Steady State Learning and Self-Confirming Equilibrium," mimeo.

——— (1993): "Steady State Learning and Nash Equilibrium," *Econometrica*, 61 (1993), 547–573.

FUDENBERG, D., D. M. KREPS, AND D. K. LEVINE (1988): "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44, 354–380.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge, MA: MIT Press.

GEANAKOPLOS, J., AND H. POLEMARCHAKIS (1982): "We Can't Disagree Forever," *Journal of Economic Theory*, 28, 192–200.

HAHN, F. (1977): "Exercises in Conjectural Equilibrium Analysis," *Scandinavian Journal of Economics*, 79, 210–226.

KALAI, E., AND E. LEHRER (1991a): "Private Beliefs Equilibrium," mimeo.

——— (1991b): "Rational Learning Leads to Nash Equilibrium," mimeo.

KOHLBERG, E., AND J. F. MERTENS (1986): "On the Strategic Stability of Equilibrium," *Econometrica*, 54, 1003–1034.

KREPS, D. M. (1990): *A Course in Microeconomic Theory*. Princeton, N.J.: Princeton University Press.

MYERSON, R. B. (1986): "Multistage Games With Communication," *Econometrica*, 54, 323–358.

ROTHSCHILD, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 195–202.

RUBINSTEIN, A., AND A. WOLINSKY (1990): "Rationalizable Conjectural Equilibrium Between Nash and Rationalizability," mimeo.

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