GAMES AND ECONOMIC BEHAVIOR 3, 82–100 (1991)

Adaptive and Sophisticated Learning in Normal Form Games*

PAUL MILGROM

Department of Economics, Stanford University, Stanford, California 94305

AND

JOHN ROBERTS

Graduate School of Business, Stanford University, Stanford, California 94305

Received December 13, 1989

In a class of games including some Cournot and Bertrand games, a sequence of plays converges to the unique Nash equilibrium if and only if the sequence is "consistent with adaptive learning" according to the new definition we propose. In the Arrow–Debreu model with gross substitutes, a sequence of prices converges to the competitive equilibrium if and only if the sequence is consistent with adaptive learning by price-setting market makers for the individual goods. Similar results are obtained for "sophisticated" learning. All the familiar learning algorithms generate play that is consistent with adaptive learning. *Journal of Economic Literature* Classification Numbers: 026, 021. © 1991 Academic Press, Inc.

Equilibrium analysis dominates the study of games of strategy, but even many of its foremost exponents are troubled by its assumption that players immediately and unerringly identify and play a particular vector of equilibrium strategies, that is, by the assumption that the equilibrium is common knowledge.[1] An alternative (and to some extent complementary) approach to analyzing behavior in games focuses on learning. The typical

* We thank Xinghai Fang for his able research assistance, Andreu Mas Colell and Tom Sargent for encouraging us to pursue this subject, and Frank Hahn for his correct guess that our theory of stability in games with strategic complements could be applied to Walrasian equilibrium when demand satisfies the "gross substitutability" condition.

[1] For example, see Kreps (1990).

0899-8256/91 $3.00

analysis of this sort considers the game being played repeatedly and posits some specific rules according to which players form expectations regarding what others' current choices will be as a function of past plays. Assuming that the players attempt to maximize their current payoffs given these expectations defines a dynamic process generating a sequence of plays, and concern then centers on the behavior of the sequence. Does play converge over time? And, if so, does it approach the behavior predicted by equilibrium analysis?

This approach is as venerable as equilibrium analysis itself: Cournot's study of duopoly (1838) introduced both the Nash equilibrium and a particular learning process. Yet this sort of analysis is subject to criticisms that are perhaps as bothersome as those leveled at equilibrium theorizing.

Cournot supposed that at each round each firm selects the quantity that would maximize its payoff if its competitors continued to produce the same quantities as at the preceding round. Now called "best-reply" dynamics, this dynamic process still receives attention as a model of learning in games (e.g., Bernheim, 1984; Moulin, 1986). Yet it often seems unreasonable to suppose that real firms would behave in the particular way Cournot described. This is especially true when best-reply dynamics lead to nonconvergent, cycling behavior (as happens for some specifications of costs and demand). When there is cycling, an outsider with no information about payoffs could eventually predict the behavior of Cournot competitors more accurately than the firms in the model do, simply by predicting continuation of the historical frequency of choices.

Brown (1951) suggested a model in which the players themselves follow a similar procedure, that is, they choose the strategies that maximize their individual payoffs given the prediction that the probability distribution of competitor's play at the next round is the same as the empirical frequency distribution of past play. This dynamic model, known as "fictitious play," initially led to encouraging results: Robinson (1951) showed that the empirical distribution of strategy choices under fictitious play converges to an equilibrium distribution for any two-player, finite-strategy, zero-sum game. However, Shapley (1964) established that, without the zero-sum restriction, fictitious play can lead to cycles of exponentially increasing length, so that the empirical frequency distribution does not converge at all. Moreover, the empirical probabilities in the cycles are bounded away from the equilibrium distribution in Shapley's example. Ironically, an outsider who wants to predict the behavior of players under fictitious play

[2] In this example, the outsider could be made into a player whose actions (predictions) do not affect the other players' payoffs and whose own payoff is 1 for a correct prediction and 0 for an incorrect one. Then, we would have an example of a game in which a player can do better, along the paths actually generated, by making Cournot forecasts than by forecasting based on past empirical frequencies.

in Shapley's example can hardly do better than to employ Cournot's suggestion of supposing that each round's choices will be the same as the preceding round: In the long run, the fraction of errors made by such an outsider would converge to zero.[2]

Fictitious play is a variant of what one may call "stationary Bayesian learning," according to which the players analyze past observations as if the behavior of their competitors were stationary, assigning as much weight to observations from the distant past as to more recent observations. Of course, the actual behavior of players during learning is nonstationary, so these stationary Bayesian models are misspecified and may often place too much weight on distant past behavior. Cournot's rule, which bases its forecasts only on the competitors' most recent past play, lies at the opposite extreme. A sophisticated player would be unlikely to adopt either kind of rule.

Cycling is not the only problem to arise from learning models. Fudenberg and Kreps (1988) have shown that models like stationary Bayesian learning applied in the extensive form of the game generate a sequence of choices that may converge, but *to a strategy combination that is different from any perfect equilibrium*. As they show, the players in general extensive form games cannot always learn what strategies their competitors are playing, because a strategy is in general a function from information sets to actions and it may be difficult to gather reliable information about how a competitor would behave at information sets that have occurred only rarely.

Taken together, these results raise serious doubts about the validity of Nash equilibrium and its refinements as a general model of the likely outcomes of adaptive learning. More fundamentally, they indicate that the "rationality" of any particular learning algorithm is situation dependent: An algorithm that performs well in some situations may work poorly in others. Apparently, real biological players tailor rules-of-thumb to their environments and experience: They learn how to learn. Thus, any single, simple specification of a learning algorithm is unlikely to represent well the behavior that actual players would adopt.

A further troubling aspect of existing models of learning in games is that they all force the players to be "unsophisticated," that is, the players can use *only* information about past play, without giving any weight to information about their competitors' information, payoffs, and rationality. The competing approaches of rationalizability (Bernheim, 1984; Pearce, 1984) and Nash or correlated (Aumann, 1987) equilibrium lie at the opposite extreme: they allow weight to be placed only on payoff information. Real players often make use of both kinds of information.

We provide a general formulation that allows players to *combine* whatever they may know about the past history of play with whatever they may know about their competitors' information, alternatives, payoffs,

and rationality to form a forecast of future play. Our approach is flexible enough to encompass bounded rationality of the kind implicit in best-reply dynamics and stationary Bayesian learning, but it also applies when players are more sophisticated. It encompasses processes in which the players can identify cyclical patterns in past play and others in which players learn about which of several forecasting models work best. In the latter case, the players might track and compare the performance of numerous alternative forecasting models over time, using past performance to select among them or even to form some weighted-average consensus forecast to guide their strategy choices at the current stage. Intelligent people employ a variety of learning strategies, and in constructing our general theory we have strived to encompass them all.

To achieve the desired degree of generality, we avoid any detailed description of how the players actually reach their decisions. Instead, we focus on the sequence of plays over time. For an individual player n, such a sequence is denoted by $\{x_n(t)\}$, where for each t in the (discrete or continuous) index set, $x_n(t)$ is a pure strategy. We then identify two properties which these sequences might satisfy—one property each for "adaptive" and "sophisticated" learning. One or the other of these properties is satisfied under all of the various more specific models of learning. We take these as defining when the observed sequence is consistent with learning.

Roughly, $\{x_n(t)\}$ is "consistent with adaptive learning" if player n *eventually* chooses only strategies that are *nearly* best-replies to some probability distribution over his competitors' joint strategies, where *near zero* probability is assigned to strategies that have not been played for a *sufficiently long* time. Similarly, $\{x_n(t)\}$ is "consistent with sophisticated learning" if the player *eventually* chooses only *nearly* best-replies to his probabilistic forecast of competitors' choices, where the support of the probability distribution may include not only past plays but also strategies that the competitors might choose if they themselves are adaptive or sophisticated learners. Thus, any process $\{x_n(t)\}$ that is consistent with adaptive learning is also consistent with sophisticated learning. Sophisticated learning models allow the player to make full use of any information gleaned from past play, but they also allow the player to assimilate fully the same kind of payoff information that is used in equilibrium analyses. Sophisticated learning is differentiated from equilibrium analysis because no fulfilled expectations assumption is imposed.

The analysis is set in a class of finite-player games with compact strategy sets and continuous payoffs. The results we report are of three types. First are results about what is included in the class of processes consistent with adaptive learning. For example, we prove that if a sequence of strategy profiles $\{x_n(t)\}$ converges to a (Nash or correlated) equilibrium, then each player's play is consistent with adaptive learning. Thus, one

cannot obtain positive equilibrium convergence results with a broader class of learning models than the ones that we analyze. We also present a theorem about a class of algorithms that includes all the specific ones described above, showing that these always generate play that is consistent with adaptive learning. Second, we report results about the implications of adaptive and sophisticated learning: Given any process consistent with adaptive learning, play tends toward the *serially undominated set*, namely, the set of strategies that remain after iterated elimination of each player's strongly dominated strategies.[3] In particular, for games that have but one serially undominated strategy profile, every process consistent with adaptive learning converges to the unique serially undominated profile. Third, we report results that we and others have obtained elsewhere about the many examples in economics in which the serially undominated set is a singleton and for which our results on the convergence of learning processes are especially germane.

In Section 1 below, we state the definitions and prove the theorems described above, which were developed with deterministic learning models in mind. In Section 2, we study a stochastic learning model in which the players experiment and show how to include it in our general framework, so that play eventually converges in an appropriate probabilistic sense toward the serially undominated set. Some important economic applications are developed in Section 3.

1. Formulation and Main Theorems

We begin with a noncooperative game $\Gamma = (N, (S_n; n \in N), \pi)$, where N is the finite player set, S_n is player n's strategy set with typical element x_n, and π is the payoff function. We assume that each S_n is a compact subset of some normed space. Let $S = \times_{n \in N} S_n$. A typical element of $x \in S$ is often usefully written as $x = (x_n, x_{-n})$, where x_{-n} designates the strategy choices of everyone besides player n. The payoff function π: $S \to \mathbf{R}^N$ specifies a payoff $\pi_n(x_n, x_{-n})$ for each player n; we assume it to be continuous. Given a set T, let $\Delta(T)$ denote the set of probability distributions over T; for example, $\Delta(S_n)$ denotes the set of probability distributions on S_n (mixed strategies). Also, let $\Delta_{-n}(T_{-n}) = \times_{j \neq n} \Delta(T_j)$ denote the mixed strategies of n's competitors. In the usual way, we may identify any pure strategy with the mixed strategy that assigns it probability one, and we may correspondingly extend the domain of π to include the mixed strategies.

[3] For two-player finite games, the serially undominated strategies are the same as the rationalizable strategies of Bernheim (1984) and Pearce (1984). When there are more than two players, the serially undominated set is, in general, larger than the rationalizable set. Gul (1990) has independently developed a theory similar to ours that emphasizes the rationalizable strategies.

A strategy $x_n \in S_n$ is *ε-dominated* by another strategy $\bar{x}_n \in \Delta(S_n)$ if for all $z_{-n} \in S_{-n}$, $\pi_n(x_n, z_{-n}) + \varepsilon < \pi_n(\bar{x}_n, z_{-n})$. Being ε-dominated is more than being strongly dominated; it requires that x_n would still be strongly dominated even if ε were added to all the payoffs that could result from its being played.

Given any set $T \subseteq S$, let T_n denote the projection of T onto S_n. Let $T_{-n} = \times_{j \neq n} T_j$. Much of our analysis centers on the following operator:

DEFINITION. Given $T \subseteq S$, let

$$U_n^\varepsilon(T) = \{x_n \in S_n \mid (\forall y_n \in \Delta(S_n))(\exists z_{-n} \in T_{-n})\ \pi_n(x_n, z_{-n}) + \varepsilon \geq \pi_n(y_n, z_{-n})\}$$

$$U^\varepsilon(T) = \times_{n \in N} U_n^\varepsilon(T).$$

The letter U is mnemonic for undominated. The set $U_n^\varepsilon(T)$ is the set of pure strategies in S_n that are not ε-dominated when each other player m is limited to strategies in T_m. If a player n believes that each other player m will choose strategies in T_m, then it would be "unjustified" or "irrational" for n to play any strategy not in $U_n^\varepsilon(T)$, for every such strategy is strongly dominated (and more!) by some other strategy.

LEMMA 1. *The operator U^ε is monotone: If R and T are sets of strategy profiles with $R \subseteq T$, then $U^\varepsilon(R) \subseteq U^\varepsilon(T)$.*

The proof of Lemma 1 follows directly from the definition of U^ε. Except in the applications, monotonicity of U^ε is the only property we shall use, so we cast the analysis entirely in terms of monotone operators.

Note that the definition of U^ε allows the possibility that for some set of profiles T, $U^\varepsilon(T) \not\subseteq T$. In particular, if we begin from some arbitrary set $T \subseteq \times_{n \in N} S_n$ and proceed to apply U^ε in an iterative fashion, it is possible that some strategies not in T will be introduced somewhere in the process. Nevertheless, as the following lemma verifies, if $T = S$, the entire strategy space, this cannot occur: $U^{\varepsilon k}(S)$, the kth iterate of $U^\varepsilon(S)$, is the outcome of k rounds of crossing out ε-dominated strategies starting from S, because when the starting set is S, no crossed-out strategy is ever reintroduced.

LEMMA 2. *For any monotone operator J and for all $k \geq 0$, $J^{k+1}(S) \subseteq J^k(S)$.*

Proof. The conclusion is obvious for $k = 0$. Suppose it holds for $k = j$, so that $J^{j+1}(S) \subseteq J^j(S)$. Then by monotonicity of J, $J^{j+2}(S) \subseteq J^{j+1}(S)$. ∎

In view of Lemma 2, it is natural to define $J^\infty(S)$ as follows:

DEFINITION. For any monotone operator J, $J^\infty(S) \equiv \bigcap_{k=1}^{\infty} J^k(S)$.

It is in terms of $U^{0\infty}$ that we define the serially undominated strategy sets.

DEFINITION. The strategy profile x is said to be *serially undominated* if $x \in U^{0\infty}(S)$.

We may use the operators U^{ε} to define what we mean by adaptive learning. The first definition holds that the process $\{x(t)\}$ is consistent with adaptive learning if each player can eventually find a way to justify its choices in terms of the competitors' past play.

DEFINITIONS. A sequence of *strategies* $\{x_n(t)\}$ is *consistent with adaptive learning* by player n if $(\forall \varepsilon > 0)(\forall \hat{t})(\exists \bar{t})(\forall t \geq \bar{t})\, x_n(t) \in U_n^{\varepsilon}(\{x(s) \mid \hat{t} \leq s < t\})$. A sequence of *strategy profiles* $\{x(t)\}$ is *consistent with adaptive learning* if each $\{x_n(t)\}$ has the property.

For games with finite strategy spaces, $\{x_n(t)\}$ is consistent with adaptive learning if and only if the condition in the definition holds with $\varepsilon = 0$, leading to a simpler theory. For infinite games, however, adaptive learning is somewhat more inclusive.

Let G_n^t be the empirical distribution of player n's choices up to and including date t, G_{-n}^t the empirical joint distribution of the other player's choices, and G^t the empirical joint distribution of all the players' choices.

DEFINITION. A sequence of profiles $\{x(t)\}$ *converges omitting correlation to a correlated strategy profile* $G \in \Delta(S)$ if (1) and (2) hold, where: (1) G_n^t converges weakly to the marginal distribution G_n for all n and (2) $(\forall \varepsilon > 0)(\exists \bar{t})(\forall t \geq \bar{t})(\forall n)\, d[x_n(t), \operatorname{supp}(G_n)] < \varepsilon$, where $d(x, T) = \inf_{y \in T} \|x - y\|$. The sequence *converges to the correlated strategy* $G \in \Delta(S)$ if, in addition, G^t converges weakly to G.

The definition implies in particular that a sequence $\{x(t)\}$ converges omitting correlation to a mixed strategy Nash equilibrium if it replicates the empirical frequency of the *separate* mixed strategies and if it eventually plays only pure strategies that are in or near the support of the equilibrium mixed strategies. Full convergence requires in addition that the correlation among the individual strategies be replicated asymptotically.

THEOREM 3. (i) *If* $\{x(t)\}$ *converges omitting correlation to a correlated equilibrium in the game* Γ, *then* $\{x(t)\}$ *is consistent with adaptive learning.*

(ii) *Suppose that the sequence* $\{x(t)\}$ *is consistent with adaptive learning and that it converges to a point* x^*. *Then* x^* *is a pure strategy Nash equilibrium.*

Theorem 3 is elementary, and we omit its proof. Part (i) of the theorem is about the inclusiveness of our definitions; anything that converges to a mixed or correlated equilibrium is covered. Part (ii) is about the most

important excluded case: adaptive learning excludes any sequence that converges to any pure strategy profile other than a Nash equilibrium.

For players who know the payoff functions, even the relatively weak restrictions of adaptive learning may be too severe to encompass all "rational" behavior. For example, in a three-player game, if player 3 has played a strategy at date t that it had never played before, then player 1 might anticipate the possibility that player 2 will introduce a new strategy next period that performs better against 3's possible new pattern of play. Of course, in the usual fashion of game-theoretic reasoning, 2 might anticipate 1's response in deciding on its own choice, and more rounds of reasoning might be required.

To accommodate these possibilities, we must introduce some notation. Let $F^{\varepsilon 0}(\hat{t}, t) = U^{\varepsilon}(\{x(s) \mid \hat{t} \leq s < t\})$ and, for $k \geq 1$, let $F^{\varepsilon k}(\hat{t}, t) = U^{\varepsilon}(F^{\varepsilon,k-1}(\hat{t}, t) \cup \{x(s) \mid \hat{t} \leq s < t\})$. The $F^{\varepsilon k}$ notation is mnemonic for k-step *forward* looking. Thus, the $F^{\varepsilon 0}$ notation can be used to describe adaptive rules, which are 0-step forward looking: the players make choices that are justified in terms of competitors' past play. The profiles in $F^{\varepsilon 1}(\hat{t}, t)$ are those that can be justified by players who think like player 1 in the discussion in the last paragraph. As we add more possible rounds to the reasoning process, the set of possible justifications for any particular choice at time t expands and so the set of possible justifiable choices expands as well. The following lemma verifies that our mathematical formulation captures that intuition.

LEMMA 4. *For all* $k \geq 0$, $F^{\varepsilon k}(\hat{t}, t) \subseteq F^{\varepsilon,k+1}(\hat{t}, t)$.

Proof. Fix $(\hat{t}, t)$. It is immediate from the definitions that $F^{\varepsilon 0}(\hat{t}, t) \subseteq F^{\varepsilon 1}(\hat{t}, t)$. Applying the monotonicity of U^{ε} inductively, $F^{\varepsilon k}(\hat{t}, t) \subseteq F^{\varepsilon,k+1}(\hat{t}, t)$. ■

Whereas adaptive learning was defined to include only processes that justify choices in terms of past play, we now define sophisticated learning to be more inclusive. It incorporates the possibility that a player *may* forecast its competitors' behavior based jointly on how all the players (including itself) have acted in the past and on what all their payoffs are. (The player's own past actions and current payoffs may influence its competitors' current choices, and the player may recognize that.) Moreover, it imposes no a priori restrictions on the number of iterations of the "he may think that I may think that . . ." style of reasoning that is so central in rationalizability and traditional equilibrium analyses.

DEFINITION. A sequence of strategies $\{x_n(t)\}$ is *consistent with sophisticated learning* if $(\forall \varepsilon > 0)(\forall \hat{t})(\exists \bar{t})(\forall t > \bar{t})\ x_n(t) \in U_n^{\varepsilon}(F^{\varepsilon\infty}(\hat{t}, t))$, where

$$F^{\varepsilon\infty}(\hat{t}, t) = \bigcup_{k=1}^{\infty} F^{\varepsilon k}(\hat{t}, t),$$

and the sequence of profiles $\{x(t)\}$ is *consistent with sophisticated learning* if each $\{x_n(t)\}$ had the property or, equivalently, if $(\forall \varepsilon > 0)(\forall \hat{t})(\exists \bar{t})(\forall t > \bar{t})\ x(t) \in F^{\varepsilon\infty}(\hat{t}, t)$.

Sophisticated learners are not necessarily more successful than adaptive learners. For example, in the Battle of the Sexes game shown below,

	L	R
U	2,1	0,0
D	0,0	1,2

the process $\{x(t)\}$ in which $x(t) = (D, L)$ for all t is consistent with "sophisticated learning," because each player is permitted to forecast that the other will recognize its true interest and switch at the next round. (Of course, as Theorem 5 showed, no such outcome is possible when the process is consistent with adaptive learning.) Although sophisticated learning does not ensure that only Nash equilibria can be limit points of the process, it does impose some restrictions.

THEOREM 5. *Let $\{x(t)\}$ be consistent with sophisticated learning. Then, for each $\varepsilon > 0$ and k there exists a time $t_{\varepsilon k}$ after which (i.e., for $t \geq t_{\varepsilon k}$) $x(t) \in U^{\varepsilon k}(S)$.*

Proof. Fix any $\varepsilon > 0$ and write t_k instead of $t_{\varepsilon k}$. For $k = 0$, the conclusion holds trivially (choosing $t_0 = 0$). Suppose the conclusion holds for $k = j$. Then, there is a t_j such that for all $t \geq t_j$, $\{x(s) \mid t_j \leq s \leq t\} \subseteq U^{\varepsilon j}(S)$.

By hypothesis, the process is consistent with sophisticated learning. So, in the definition of such processes, we may let $\hat{t} = t_j$ and we may take $t_{j+1} = \max(\hat{t}, \bar{t})$. Let $t \geq t_{j+1}$. Then, $x(t) \in F^{\varepsilon\infty}(t_j, t)$ and we want to show that $F^{\varepsilon\infty}(t_j, t) \subseteq (U^{\varepsilon})^{j+1}(S)$. We show, equivalently, that $F^{\varepsilon i}(t_j, t) \subseteq (U^{\varepsilon})^{j+1}(S)$ for all i.

For $i = 0$, $F^{\varepsilon 0}(t_j, t) = U^{\varepsilon}(\{x(s) \mid t_j \leq s \leq t\}) \subseteq U^{\varepsilon}[U^{\varepsilon j}(S)] = (U^{\varepsilon})^{j+1}(S)$, by monotonicity of U^{ε}. Suppose the conclusion is true for i. In particular, by Lemma 2, $(U^{\varepsilon})^{j+1}(S) \subseteq (U^{\varepsilon})^{j}(S)$. Then, by monotonicity of U^{ε},

$$\begin{aligned} F^{\varepsilon,i+1}(t_j, t) &= U^{\varepsilon}(F^{\varepsilon i}(t_j, t) \cup \{x(s) \mid t_j \leq s \leq t\}) \subseteq U^{\varepsilon}[U^{\varepsilon j}(S)] \\ &= (U^{\varepsilon})^{j+1}(S). \quad \blacksquare \end{aligned}$$

COROLLARY 6. *Let $\{x(t)\}$ be consistent with sophisticated learning and let S_n^{∞} be the set of strategies that are played infinitely many times in $\{x_n(t)\}$. Then, $\times_{n \in N} S_n^{\infty} \subseteq U^{0\infty}(S)$. In particular, for any finite game Γ, all play lies eventually in the set of serially undominated strategies $U^{0\infty}(S)$.*

Proof. Applying Theorem 5, $S^\infty \subseteq \cap_k \cap_{\varepsilon>0} U^{\varepsilon k}(S) = \cap_k U^{0k}(S) = U^{0\infty}(S)$. ■

THEOREM 7. *Suppose $U^{0\infty}(S) = \{\bar{x}\}$. Then $\|x(t) - \bar{x}\| \to 0$ if and only if $\{x(t)\}$ is consistent with adaptive learning. Similarly, $\|x(t) - \bar{x}\| \to 0$ if and only if $\{x(t)\}$ is consistent with sophisticated learning.*

Proof. Suppose $\|x(t) - \bar{x}\| \to 0$. Since π is continuous, $(\forall \varepsilon > 0)(\exists \bar{t})(\forall t > \bar{t})(\forall n \in N)$,

$$\pi_n(x_n(t), x_{-n}(t)) - \max\{\pi_n(y_n, x_{-n}(\bar{t})) \mid y_n \in S_n\} < [\pi_n(\bar{x}) + \varepsilon/2] - [\max\{\pi_n(y_n, \bar{x}_{-n}) \mid y_n \in S_n\} - \varepsilon/2] = \varepsilon.$$

Hence, $x_n(t) \in U_n^\varepsilon(\{x(\bar{t})\}) \subseteq U_n^\varepsilon(\{x(s) \mid \bar{t} \le s < t\})$. Then, $x(t) \in F^{0\varepsilon}(\bar{t}, t)$, which establishes that convergence can occur only if $\{x(t)\}$ is consistent with adaptive learning and hence only if it is consistent with sophisticated learning.

For the "if" part, let x^* be an accumulation point of $\{x(t)\}$ and assume that $\{x(t)\}$ is consistent with (adaptive or) sophisticated learning. By Theorem 5, $(\forall k)(\exists \bar{t})(\forall t > \bar{t})\, x(t) \in U^{\varepsilon k}(S)$. By Lemma 2 and the compactness of $U^{\varepsilon k}(S)$,

$$x^* \in \cap_{\varepsilon>0} \cap_{k=1}^\infty U^{\varepsilon k}(S) = \cap_{k=1}^\infty \cap_{\varepsilon>0} U^{\varepsilon k}(S) = U^{0\infty}(S) = \{\bar{x}\},$$

where the reversal of intersections is justified because $U^{\varepsilon k}(S)$ is doubly monotone (decreasing in k, increasing in ε). ■

Although our formulation of learning allows the possibility that players may be learning how to optimize in addition to learning what to expect from competitors,[4] all the studies and algorithms cited in the introduction assume that players are always able to optimize given their possibly inaccurate forecasts. To establish that our theory subsumes these earlier ones, let p_n denote a forecasting algorithm for player n, so that $p_n(\cdot \mid x(s); s < t)$ is a probability distribution over S_{-n} representing what player n expects to be played at date t given the history of play (including his own play) up to that date. Let A_n^p be a learning algorithm that makes the optimizing choices associated with p, that is, $A_n^p[x(s); s < t] \in \operatorname{argmax} E\{\pi_n(x_n, x_{-n}) \mid p_n[\cdot \mid x(s); s < t]\}$. Best-reply dynamics, fictitious play, and many others are algorithms of this sort.

For simplicity, let us restrict attention here to games with finite strategy sets and to learning that occurs in discrete time. Then, in a small stretch of

[4] In some environments, this extension allows us to encompass "genetic algorithms" of the kind studied by Marimon *et al.* (1989).

terminology, we may say that the forecasting algorithm p_n is *adaptive* if for any strategy x_m of any player $m \neq n$ that is played only finitely many times in the sequence $\{x(t)\}$, $p_n(x_m \mid x(s);\ s < t)$ converges to zero as $t \to \infty$.[5]

THEOREM 8. *Let p_n be an adaptive forecasting algorithm. Then for any sequence of opposing strategy profiles $\{x_{-n}(t)\}$, the induced sequence $\{x_n(t)\} = \{A_n^p[x(s);\ s < t]\}$ is consistent with adaptive learning.*

The proof of Theorem 8 for this finite strategy case is obvious: Player n's choices under A_n^p are eventually best-responses to some probability distribution over those opposing strategies that will be played infinitely often. Thus, for all $\bar{t}$, the choices $\{x_n(t)\}$ lie eventually in $U_n^\varepsilon(\{x(s) \mid \bar{t} \leq s \leq t\})$, which is the definition of the phrase "$\{x_n(t)\}$ is consistent with adaptive learning."

Since the forecasts implicit in Cournot's best-reply dynamics, Brown's fictitious play, and Bayesian learning are all adaptive, our theory contains those examples as special cases. Similarly, if a player were to use Bayesian statistical methods, for example, to estimate a matrix whose elements are the conditional probabilities that a particular profile x_{-n} is played given a specification of the k previous plays $\{x(t-k), \ldots, x(t-1)\}$, the forecasts would be adaptive. This example illustrates how our theory encompasses learning strategies that can recognize cycles. Furthermore, if $\{p_n^1, \ldots, p_n^k\}$ is a set of adaptive forecasting algorithms, then one can verify that any algorithm q_n whose forecasts always lie in the convex hull of those generated by the set $\{p_n^1, \ldots, p_n^k\}$ is another adaptive forecasting algorithm. This means that any procedure for selecting among or weighting the forecasts of a finite set of adaptive forecasting algorithms, for example, on the basis of past performance, is another adaptive forecasting algorithm: Learning how to forecast is consistent with adaptive learning.

2. STOCHASTIC LEARNING PROCESSES INVOLVING EXPERIMENTATION

So far, we have formulated our concepts so that if occasional mistakes are made over an infinite horizon, then play is inconsistent with adaptive learning even if the mistakes eventually become very rare. This formulation is at odds with the idea that learning might be based on experimentation—an idea that has recently been incorporated in a formal model by Fudenberg and Kreps (1988).

[5] For the general case of compact strategy sets and time which may be modeled as continuous, we may define p to be an adaptive forecasting algorithm if the probability assigned to any compact set of strategies from which no plays are chosen after some date $\bar{t}$ converges to zero as t goes to infinity. With this definition, Theorem 8 can be proved for the general case.

Let Γ be a noncooperative game with a finite player set and a finite strategy set S. We follow Fudenberg and Kreps (1988) in supposing that at each date t, player n conducts an experiment with probability ε_{nt} in an attempt to learn its best play. Player n's decision to experiment at any date is assumed to be independent of any contemporaneous decisions of the other players, and n's experiments eventually become rare ($\varepsilon_{nt} \to 0$ as $t \leftarrow \infty$). However, the number of experiments that player n conducts eventually is infinite ($\sum_t \varepsilon_{nt} = \infty$). This latter assumption ensures that the player experiments often enough that if its competitors' behavior ever settles down to a stationary distribution, the player would learn to play a best reply. We assume that when the player experiments, it selects each of its finite number of strategies with equal probability.

Suppose that when player n does not experiment at a given date, it picks a strategy that is among those with the highest average payoff on past dates of experimentation. Adaptive behavior of this general kind might be sensible if the player has no idea what the environment is like, how many players there may be, what strategies they have played at any round, or even whether a game against maximizing players is being played, but knows only what strategies it has played and what payoffs it has earned on the dates when it has experimented. Given any realization ω of the player's randomized choices, let $\{t(k, \omega)\}$ be the subsequence of dates at which player n conducts no experiment.

THEOREM 9. *For any finite strategy game* Γ, *the sequence* $\{x_n(t(k, \omega))\}$ *constructed as described above is consistent with adaptive learning (almost surely).*

The proof is given in the Appendix.

In formulating Theorem 9, we have assumed that at each date player n uses only information about the payoffs earned *at dates of experimentation*. The reader is warned that the extension to the case where the players use information about the strategies they have chosen and the payoffs they have earned at *all* dates is not straightforward. Because the player randomizes strategy choices at dates of experimentation but not at other dates, naive estimates based on the payoffs at the experimentation dates are always unbiased. A player who wishes to take advantage of experience gained at other dates may need to use sophisticated statistical techniques to avoid having the estimates contaminated by selection bias.

3. APPLICATIONS

The key to the applications of this theory lies in an investigation of the sets $U^\infty(S)$. For a game with N players each having k strategies, if the payoffs are picked at random from [0,1] using some continuous distribu-

tion, then the expected number of strongly dominated strategies can be shown to be equal to $N(k-1)k^{1-\Lambda}$, which tends to zero as k and N grow large. So, for "generic large games," there are usually no strongly dominated strategies, and then no general restrictions are implied by our theory about which strategies can be played infinitely often. Nevertheless, for many game models that have attracted the interest of applied researchers, our theory does imply surprisingly strong restrictions.

An important class of games for which our theory is useful is the class of "games with strategic complementarities," as variously defined by several authors. This class includes the supermodular games of Topkis (1979) and Vives (1990), in which the strategy sets are complete lattices[6] and the incremental return to any player from increasing his strategy in the lattice order is a nondecreasing function of the strategy choices of the other players. Membership in this class of games is often easy to check. Indeed, if the payoff functions are smooth and (as in many applications) the strategy spaces are compact intervals in $\mathbf{R}$, then Γ is a supermodular game if and only if $\partial^2\pi_n/\partial x_n\partial x_m \geq 0$ for all $m \neq n$.

Milgrom and Roberts (1990) (hereafter denoted as MR) use a more inclusive definition, defining a game as having strategic complementarities if there exist any strictly increasing functions f_n such that the game with the transformed payoffs $f_n(\pi_n(x_n, x_{-n}))$ is a supermodular game. Milgrom and Shannon (1991) (hereafter MS) provide a still more inclusive (but harder to check) definition. For the case where the strategy sets are totally ordered (e.g., subsets of the real line), they define the class of games with strategic complementarities as those for which the following ("single-crossing") conditions hold:

For all $x, y \in S$ with $x \geq y$,[7]

$$[\pi_n(x_n, y_{-n}) \geq \pi_n(y_n, y_{-n})] \Rightarrow [\pi_n(x_n, x_{-n}) \geq \pi_n(y_n, x_{-n})] \text{ and}$$

$$[\pi_n(x_n, y_{-n}) > \pi_n(y_n, y_{-n})] \Rightarrow [\pi_n(x_n, x_{-n}) > \pi_n(y_n, x_{-n})].$$

For the analysis of these games, we may define monotone operators UP (mnemonic for *undominated* in *pure* strategies) and I (mnemonic for *interval*), as follows.

DEFINITIONS. Given $T \subseteq S$, let

(1) $[T] = \{x \in S \mid \inf(T) \leq x \leq \sup(T)\}$

(2) $UP_n(T) = \{x_n \in S_n \mid (\forall y_n \in S_n)(\exists z_{-n} \in T_{-n})\ \pi_n(x_n, z_{-n}) \geq \pi_n(y_n, z_{-n})\}$

[6] A complete lattice is a partially ordered set S_n with the property that every subset T_n has an infimum and a supremum in S_n.

[7] For our applications here, this may be read as a simple vector inequality, where each player's strategy space is a compact subset of the real line $\mathbf{R}$.

(3) $UP(T) = \times_{n \in N} UP_n(T)$
(4) $I(T) = UP([T])$.

The operator I will be especially useful in studying learning rules in which the player is allowed to reason in the following kind of way: "My competitor has set prices of 5 and 6 in the past two periods, so it seems reasonable to entertain the possibility that it might set some price between 5 and 6 in the current period." Similarly, "Since I have set prices of 3 and 4 in the past two periods, my competitors might expect me to set some price between 3 and 4 next period, and might respond accordingly."

Since I is the composition of the two monotone operators $UP(\cdot)$ and $[\cdot]$, I is itself a monotone operator. So, by the very same arguments that we have used in Theorem 5 for the operator U, if the players eventually choose the strategy profile $x(t)$ from the set $I(\{x(s);\ t \geq s \geq \hat{t}\})$, play eventually lies in the set $I^k(S)$. How large is this set? The next two theorems provide the answer.

THEOREM 10. *Let Γ be a game with strategic complementarities. Then, both $\underline{x} = \inf(I^\infty(S))$ and $\bar{x} = \sup(I^\infty(S))$ are pure strategy Nash equilibrium profiles (and therefore elements of $U^{0\infty}(S)$).*

Proofs of Theorem 10 are given in MR and MS, using their respective definitions of the phrase "strategic compementarities."

THEOREM 11. *Let Γ be a game with strategic compementarities and let PNE denote the set of pure Nash equilibrium profiles. Then, the bounds on joint behavior predicted by the various "justification" concepts coincide:*

$$[U^{0\infty}(S)] = [UP^\infty(S)] = [I^\infty(S)] = [PNE].$$

Proof. Since for all T, $U^0(T) \subseteq UP(T) \subseteq I(T)$ and since the operators are all monotone nondecreasing, it follows by induction on k that for all T, $U^{0k}(T) \subseteq UP^k(T) \subseteq I^k(T)$ and hence that $U^{0\infty}(S) \subseteq UP^\infty(S) \subseteq I^\infty(S)$. Using the notation and results of Theorem 9, $\{\underline{x}, \bar{x}\} \subseteq U^{0\infty}(S)$ and $I^\infty(S) = [\{\underline{x}, \bar{x}\}] = [PNE]$. Applying the monotone operator $[\cdot]$ to these inclusions yields $[\{\underline{x}, \bar{x}\}] \subseteq [U^{0\infty}(S)] \subseteq [UP^\infty(S)] \subseteq [I^\infty(S)] = [[\{\underline{x}, \bar{x}\}]] = [\{\underline{x}, \bar{x}\}] = [PNE]$, where the next-to-last equality follows from the fact that for all T, $\inf([T]) = \inf(T)$ and $\sup([T]) = \sup(T)$. ■

We will illustrate the power of these theorems with a series of three applications, in each of which the serially undominated set $U^{0\infty}(S)$ is a singleton. It follows that for each application, there exists a unique Nash equilibrium and that $\{x(t)\}$ converges to this equilibrium if and only if it is consistent with adaptive learning (and if and only if it is consistent with sophisticated learning). Additional applications can be found in MR and MS.

EXAMPLE 1: COURNOT'S DUOPOLY MODEL. Let the market demand function be given by $P = f(Q)$, where $f(\cdot)$ is continuous and decreasing and where $f'(Q) + Qf''(Q)$ is nowhere positive. Following Cournot, let the two market participants have zero costs of production and suppose that the strategy spaces are specified as $S_n = [0,\bar{q}]$. Then, essentially as Moulin (1986, Chap. 6) has shown, $U^{0\infty}(S)$ is a singleton (the game is "dominance solvable")[8] and so Theorem 7 applies: all behavior that is consistent with adaptive or sophisticated learning converges to the unique Nash equilibrium. This significantly expands Cournot's conclusion that the quantities chosen using best-reply dynamics converge to the unique equilibrium in the Cournot game.

EXAMPLE 2: BERTRAND OLIGOPOLY WITH DIFFERENTIATED PRODUCTS. Let the demand for product n be given by $q_n = D_n(p)$, where $p = (p_m;\ m \in N)$, and suppose that the demand for good n becomes less elastic as p_{-n} increases ($\partial^2\log(D_n(p))/\partial\log(p_n)\partial p_m > 0$ for $m \neq n$). For substitute products, this condition is satisfied by the linear, logit, and CES demand functions, and by translog demand with certain parameter restrictions. Let the costs of production be $C_n(q_n) = c_n q_n$ and let $S_n = [c_n, \bar{p}]$. Then, as shown in MR, this game has a unique Nash equilibrium and the transformed game with payoffs $\log(\pi_n)$ is a supermodular game.[9] It then follows from Theorem 10 that $U^{0\infty}(S)$ is a singleton, so a process $\{x(t)\}$ converges to the unique equilibrium if and only if $\{x(t)\}$ is consistent with adaptive learning.

EXAMPLE 3: GENERAL EQUILIBRIUM WITH GROSS SUBSTITUTES. Arrow and Hurwicz (1958) and Arrow *et al.* (1959) proved that any general equilibrium system satisfying the gross substitutability condition is stable under a class of continuous-time tatonnement price adjustment processes. Because the definitions and proofs in this paper make no use of the hypothesis that time is discrete, our theorems can be applied to these continuous-time models as well as to the discrete models that are more commonly used to describe learning in games.

For the general equilibrium system, suppose that there are $L + 1$ commodities and that one is specified as the muneraire. Let there be L players and let player n name a price $p_n \in S_n = [0,\infty]$; player n may be called the "market maker" for good n. We assume that the excess demand for commodity n is $q_n = q_n(p)$ and that it has the following properties: q_n is continuous, nonincreasing in p_n and nondecreasing in p_{-n} (gross substi-

[8] Moulin defines dominance solvability using weak dominance, but his argument can be extended to the case of strong dominance with no difficulty. Alternatively, the same conclusion can be reached using the theory of supermodular games. See Milgrom and Roberts (1990).

[9] To verify this, observe that $\partial^2\pi_n/\partial p_n\partial p_m = \partial^2\log[(p_n - c_n)D_n(p)]/\partial p_n\partial p_m = [\partial^2\log(D_n(p))/\partial\log(p_n)\partial p_m] \cdot \partial\log(p_n)/\partial p_n \geq 0$.

tutes). Also, $q_n(\infty, p_{-n}) < 0$ and $q_n(0, p_n) > 0$. Let player n's payoff be given by $\pi_n(p) = -|q_n(p)|$. By inspection, any pure strategy Nash equilibrium $\bar{p}$ of this game is a competitive equilibrium, that is, $q(\bar{p}) \equiv 0$. Arrow and Hurwicz showed that, with these assumptions, there is a unique competitive equilibrium. It is shown in MS that this game is one with strategic complementarities. Therefore, by Theorem 10, $U^{0\infty}(S) = \{\bar{p}\}$.

In conjunction with Theorem 7, this is a powerful conclusion. To illustrate its force, suppose that each market maker bases the price adjustment on some distributed lag estimate using past realized price choices of the other market makers. Any distributed lag will do; and each market maker can use a different distribution and the price process can start anywhere. The process will always converge to the competitive equilibrium.

THEOREM 12. *For any family of probability distributions G_n ($n = 1, \ldots, N$) on $[0,\infty)$ and any $p(0)$, $q(0) \in \mathbf{R}^N$, if*

$$\dot{p}_n(t) = \int_0^t q_n[p_n(t), p_{-n}(t - s)]dG_n(s) + (1 - G_n(t))q_n(0) \qquad (*)$$

and if demand satisfies the assumptions stated above, then $p(t)$ converges to the unique competitive equilibrium.

Proof. Fix ε and define $T(\hat{t}, t) \equiv \{p(s) \mid \hat{t} \leq s < t\}$ and $D(\hat{t}, t) = U^{\varepsilon}(T(\hat{t}, t))$. Since $q_n(p_n, p_{-n})$ is nonincreasing in p_n, $D(\hat{t}, t)$ is an interval which we may write as $[\underline{p}(\hat{t}, t), \bar{p}(\hat{t}, t)]$. Since $T(\hat{t}, t)$ grows with increasing t and since U^{ε} is monotone, $D(\hat{t}, t)$ grows with increasing t. Hence, $\underline{p}(\hat{t}, t)$ can only decrease and $\bar{p}(\hat{t}, t)$ increase with t.

Suppose that $p_n(t) \notin D_n(\hat{t}, t)$, for example, $p_n(t) > \bar{p}_n(\hat{t}, t)$. Then $q_n(p_n(t), p_{-n}) < -\varepsilon$ for all $p \in T(\hat{t}, t)$. So, by (*), there is a $\bar{t}$ such that if $t > \bar{t}$ and $p_n(t) \notin D_n(\hat{t}, t)$, then $\dot{p}_n(t) < -\varepsilon/2$. Hence, $(\exists \bar{\bar{t}})(\forall t > \bar{\bar{t}})\ p_n(t) < \bar{p}_n(\hat{t}, t)$. Similarly, $(\exists \bar{\bar{t}})(\forall t > \bar{\bar{t}})\ p_n(t) > \underline{p}_n(\hat{t}, t)$. These imply that $p_n(t) \in D_n(\hat{t}, t)$, so $\{p(t)\}$ is consistent with adaptive learning using U^{ε}. The desired conclusion then follows from Theorem 7. ■

Thus, for example, if information processing is delayed so that the market can adjust current prices only on the basis of last quarter's demand, prices will nevertheless converge to the competitive equilibrium levels.

Extensions of Theorem 11 are easily found. The lag coefficients $G_n(s)$ can be replaced by nonstationary coefficients $H_n(s, t)$, provided $\inf_t H_n(\cdot, t)$ is a probability distribution. This extension allows the possibility that the demand data used for price adjustment in the various goods markets become available only at nonsynchronized discrete time intervals. Similarly, various transformations of the demand data in the integrand and of the rate of adjustment can be accommodated. For example, $\dot{p}_n(t)$ might be

set equal to $+1$ or -1, according to whether the integral is positive or negative.

4. Discussion

In general games, our conditions of consistency with adaptive (or sophisticated) learning impose a joint restriction on the game and the players' learning processes. It is hardly a surprise that one can express a necessary condition for convergence to equilibrium in this general form. It is rather more surprising that the condition is sufficient, by itself, to imply "convergence" into the set of serially undominated strategies and especially that it is so easy to prove that best reply dynamics, fictitious play, and various other processes are consistent with adaptive learning for *all* normal-form games in the broad class that we have studied. A second surprise is that for certain (nongeneric) examples of games that have historically attracted the interest of economists, the necessary learning conditions are also sufficient to imply that behavior converges to the unique equilibrium.

It had once been thought that convergence in these games that economists have studied depended very much on the algorithms that were used. For example, in the general equilibrium example, the analogy with known results about optimization algorithms inspired some to think that the damping imposed by the continuous time dynamics played an important role in facilitating convergence (see the discussion in Arrow and Hurwicz (1958)). One of the lessons of our analysis is that convergence can be almost solely a property of the game being studied: "nearly everything" converges to equilibrium in the three economic applications that we studied.

Generic games do not enjoy the special structure that we have exploited in our analysis for applications. It is still important, therefore, to explore processes that are robustly convergent for a wider class of games, and perhaps for all games. The hope is that such a theory would be analogous to the theory of "refinements," predicting the relative likelihood of various equilibria as the outcome of learning in a way that might be tested in laboratory experiments. Jordan (1991) makes progress along that line, showing that in a model where uncertainty about competitors' play can be represented as due to differences in types, if all the players engage in Bayesian learning about each others' types, then play converges to a full information Nash equilibrium—a stronger conclusion than we have obtained for general learning models. Jordan's model provides an example of a setting in which one might usefully investigate which equilibria are most likely to arise as the result of adaptive learning.

APPENDIX: PROOF OF THEOREM 9

Let $M(x_n, t)$ be the number of experiments using x_n conducted by player n by time t and let $M(t)$ be the expected total number of experiments using all the player's strategies. Let $x_{-n}(t, \omega)$ be the strategy combination played at date t by n's competitors. Fix $\hat{t}$ and define $T = \{x(t) \mid t \geq \hat{t}, t \notin D\}$, where D is the set of dates at which experiments are conducted. Since the strategy sets are finite, there is some date $\bar{t} \geq \hat{t}$ by which all the strategies in T_m for all $m \neq n$ have already been played and the "justifiable" strategy profiles are eventually limited to those in $U(T)$.

Suppose that $x_n \notin U_n(T)$. To establish that the play is consistent with adaptive learning, we need to show that there is some date t' after which x_n is no longer played, except during periods of experimentation. Since S_n is finite and $x_n \notin U_n(T)$, $(\exists y_n \in S_n)(\exists \varepsilon > 0)(\forall z \in T)\ \pi_n(x_n, z_{-n}) < \pi_n(y_n, z_{-n}) + 2\varepsilon$. Let $P(x_n, t, \omega)$ be the total payoff at experimental dates when x_n is played and define $P(y_n, t, \omega)$ similarly. Suppressing ω from the notation, we are given that n will not play x_n at any nonexperimental date where $P(x_n, t)/M(x_n, t) < P(y_n, t)/M(y_n, t)$, so it suffices to show that, with probability one, there exists a date after which this inequality always holds. By the Strong Law of Large Numbers, $M(x_n, t)/M(t)$ and $M(y_n, t)/M(t)$ both converge almost surely to $1/|S_n|$ (where $|S_n|$ is the cardinality of S_n). Consequently, it suffices to show that for large t, $P(x_n, t)/M(t) < P(y_n, t)/M(t)$ or that $P(x_n, t) < P(y_n, t) - \varepsilon M(t)/|S_n|$. Observe that $M(t) = \sum_{\tau \leq t} \varepsilon_{n\tau}$.

Let $\Delta/2$ be a bound for $|\pi_n|$ and let t' be large enough that for all player indexes m and all $t > t'$, $\sum_{m \in N} \varepsilon_{mt} < \varepsilon/\Delta$.

Let $\mathbf{F}_t$ be the history of play through time t. At any date $t + 1$, the choice by player n to experiment and, if so, which strategy to choose are (by assumption) independent of $x_{-n}(t + 1)$. So,

$$E[P(x_n, \tau + 1) - P(y_n, \tau + 1) \mid \mathbf{F}_\tau] - P(x_n, \tau) - P(y_n, \tau) = (\varepsilon_{n,\tau+1}/|S_n|) \cdot E[\pi_n(x_n, x_{-n}(\tau + 1)) - \pi_n(y_n, x_{-n}(\tau + 1)) \mid \mathbf{F}_\tau] < P(x_n, \tau) - P(y_n, \tau) - (2\varepsilon_{n,\tau+1}/|S_n|)\varepsilon.$$

Taking expectations and summing over τ, we get a telescoping series which yields

$$E[P(x_n, t) - P(y_n, t)] < -(2\varepsilon/|S_n|) \sum_{\tau=1}^{t} \varepsilon_{n\tau} = -(2\varepsilon/|S_n|) \cdot M(t).$$

Since $|\pi_n| < \Delta/2$ and the probability of a change in $[P(x_n, t) - P(y_n, t)]$ from date $t - 1$ to date t is $2\varepsilon_{t+1}/|S_n|$,

$$\text{Var}[P(x_n, t) - P(y_n, t)] \leq (2\Delta^2\varepsilon/|S_n|) \sum_{\tau=1}^{t} \varepsilon_{n\tau} = (2\Delta^2\varepsilon/|S_n|) \cdot M(t).$$

If follows from a supermartingale convergence theorem (Breiman, 1968, Theorem 5.23) that $P(x_n, t) - P(y_n, t) + \varepsilon M(t)/|S_n|$ is a supermartingale converging to $-\infty$.

REFERENCES

ARROW, K., BLOCK, H. D., AND HURWICZ, L. (1959). "On the Stability of Competitive Equilibrium II," *Econometrica* **27**, 82–109.

ARROW, K., AND HURWICZ, L. (1958). "On the Stability of Competitive Equilibrium I," *Econometrica* **26**, 522–552.

AUMANN, R. (1987). "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica* **55**, 1–18.

BERNHEIM, B. D. (1984). "Rationalizable Strategic Behavior," *Econometrica* **52**, 1007–1028.

BREIMAN, L. (1968). *Probability*. Reading, MA: Addison–Wesley.

BROWN, G. W. (1951). "Iterative Solution of Games by Fictitious Play," in *Activity Analysis of Production and Allocation*. New York: Wiley.

COURNOT, A. (1960). *Recherches sur les Principes Mathématiques de la Théorie des Richesses*, 1838, translated by N. T. Bacon as *Researches into the Mathematical Principles of the Theory of Wealth*. London: Hafner.

FUDENBERG, D., AND KREPS, D. (1988). "A Theory of Learning, Experimentation, and Equilibrium in Games," unpublished monograph.

GUL, F. (1990). "Rational Strategic Behavior and the Notion of Equilibrium," mimeo, Stanford University.

JORDAN, J. (1991). "Bayesian Learning in Normal Form Games," *Games Econ. Behav.* **3**, 60–81.

KREPS, D. (1990). *Game Theory and Economic Modelling*. Oxford: Oxford Univ. Press.

MARIMON, R., MCGRATTAN, E., AND SARGENT, T. (1989). "Money as a Medium of Exchange in an Economy with Artifically Intelligent Agents," mimeo.

MILGROM, P., AND ROBERTS, J. (1990). "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, in press.

MILGROM, P., AND SHANNON, C. (1991). "An Extended Theory of Games with Strategic Complementarities," in preparation.

MOULIN, H. (1986). *Game Theory for the Social Sciences*, 2nd ed. New York: New York Univ. Press.

PEARCE, D. (1984). "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica* **52**, 1029–1050.

ROBINSON, J. (1951). "An Iterative Method of Solving a Game," *Ann. Math.* **54**, 296–301.

SHAPLEY, L. (1964). "Some Topics in Two-Person Games," in *Advances in Game Theory*, *Annals of Mathematical Studies*, Vol. 5, pp. 1–28.

TOPKIS, D. (1979). "Equilibrium Points in Non-zero Sum n-Person Submodular Games," *SIAM J. Control Optim.* **17**, 773–787.

VIVES, X. (1990). "Nash Equilibrium with Strategic Complementarities," *J. Math. Econ.* **19**, 305–321.