DISTRIBUTIONAL STRATEGIES FOR GAMES WITH INCOMPLETE INFORMATION*†

PAUL R. MILGROM† and ROBERT J. WEBER‡

We study games with incomplete information from a point of view which emphasizes the empirical predictions arising from game-theoretic models. Using the notion of "distributional" strategies, we prove four main theorems: (i) a mixed-strategy Nash equilibrium existence theorem, (ii) a pure-strategy equilibrium existence theorem, (iii) a pure-strategy e-equilibrium existence theorem, and (iv) a theorem describing how the set of equilibria of a game varies with the parameters of the game.

1. Introduction. In 1961, William Vickrey introduced games with incomplete information into the mainstream of economic theory in a study of competitive bidding [30]. The variety of applications of these games has widened considerably over the years (as evidenced by [6], [13], [18], [21]-[23], [25], [26], [31]), and many contributions have been made to the underlying theory of information [15]-[17], [20].

Despite these advances, the most fundamental questions which arise in applications remain unanswered. Do Nash equilibria exist for general games with incomplete information? When will such games have equilibria in pure strategies? How sensitive are equilibrium outcomes to modeling assumptions? For example, can a small variation in the assumed information structure lead to a large change in the equilibrium strategies?

Our aim is to provide partial answers to all of these questions, for one-stage simultaneous-move games. We prove an equilibrium existence theorem for a broad class of these games. We also prove that, with the appropriate concepts of closeness for information structures, payoff functions, and strategies, the correspondence that maps the specifications of a game into its set of Nash equilibria is upper-semicontinuous. The ideas underlying this continuity theorem have been used elsewhere [13], [21] to simplify and unify the solutions of certain bidding games and to gain insights into the nature of the equilibria of these games; here we use them to study a game of timing that can serve as a model of competition between two animals, or of strike behavior. For games in which the players' informational variables have atomless distributions, we show that each player's set of pure strategies is dense in his complete set of strategies. For such games, mixed strategies are empirically indistinguishable from pure strategies and so the common objection that "one never observes people adopting mixed strategies" has no force. Finally, we identify a large class of games in which every mixed strategy equilibrium has a "purification", i.e., a pure-strategy equilibrium at which each player has the same expected payoff and the same distribution of

*Received October 13, 1980; revised August 3, 1984.
OR/MS Index 1978 subject classification. Primary: 238 Games/group decisions/noncooperative.
Key words. Incomplete information, Nash equilibrium, distributional strategies.

†This research was supported in part by U.S. Office of Naval Research contracts ONR-N000-14-77-C-0518 and ONR-N000-14-79-C-0685, National Science Foundation grants SOC-77-06609-A1 and SES-80-01932, and the Center for Advanced Studies in Managerial Economics at Northwestern University.
‡Yale University.
¶Northwestern University.
We take up the matter of pure-strategy equilibria in §5. Two theorems and a corollary summarize our results. Theorem 3 (the Denominator Theorem) states that a player’s set of pure strategies is dense in his set of distributional strategies whenever his type has an atomless distribution. A corollary asserts the existence of approximate equilibria in pure strategies. To study exact equilibria, we need the following definition. A mixed strategy is said to have a purification if there is a pure strategy that is equidistant to all others and for which the expected payoff of each type of each opposing player is the same against each player. For whatever strategies the other players may adopt, it is always true that the expected payoff of each type of each opposing player is the same against each player. For whatever strategies the other players may adopt, the joint distribution of the state is the same whether or not the player played the played strategy. Theorem 4 (The Purification Theorem) gives sufficient conditions to ensure that every mixed strategy has a purification. When the conditions are satisfied, every mixed-strategy equilibrium corresponds to some pure-strategy equilibrium in which each player faces the same decision problem (at equilibrium) and earns the same expected payoff as in the mixed-strategy equilibrium. §6 indicates how the assumption that the type spaces are metric can be relaxed and how the case of “inconsistent beliefs” can be treated.

2. An example: The war of attrition. In this section, we present an example which illustrates the notions of “closeness” of information and payoff structures, and of strategies (possibly different) games with incomplete information, notions which are central to the results of this paper. The game we analyze is known as the “War of Attrition," and has been used to study conflict among animals [6]. It can also serve to model other conflicts, such as labor-management disputes involving strikes.

In the animal-conflict interpretation, two animals face one another in competition for a valuable prize, such as food or the opportunity to mate. The competition is nonviolent (for example, the animals engage in a display ritual); the cost of competition is proportional to the time spent in conflict (time which alternatively could be employed in the search for a comparable prize elsewhere, time during which both competitors are vulnerable to attack from predators). Let \( s_i = 1.2 \) be the greatest length of time animal \( i \) would be willing to compete for the prize, if he were certain to receive the prize at the end of that time. We take the payoff functions of the game to be of the form

\[
U(t_i, s_2, a_2, a_1) = \begin{cases} 
  s_i - a_1 & \text{if } a_1 > a_2, \\
  -a_2 & \text{otherwise},
\end{cases}
\]

where \( i = (1, 2) \), and \( a_1 \) and \( a_2 \) are the lengths of time for which the two animals choose to compete. Suppose that to both competitors, the game has a unique symmetric Nash equilibrium point; the strategy of each at equilibrium is to choose the time at which he will concede according to the exponential distribution with mean \( \theta \). On the other hand, if \( F \) is a continuous distribution with density \( f \) it is known that there is a unique symmetric Nash equilibrium point, in which each uses a pure strategy which calls for competition until time \( b(s) \), where

\[
b(s) = \int_0^s \frac{a(t)}{1 - F(t)} \, dt.
\]
How do the equilibria in these two cases compare? Let \( G \) be the distribution of quitting times in the second case; then

\[
G(a) = \Pr(e(\xi) < a) = \Pr\left( \int_0^a \frac{f(t)}{1 - F(t)} \, dt < a \right).
\]

Assume that \( F \) concentrates its mass on a neighborhood \((\bar{e} - \epsilon, \bar{e} + \epsilon)\). Then the integrand in the expression above is nonzero only if \( s < \bar{e} - \epsilon \), and so

\[
G(a) > \Pr\left( \int_0^a \frac{f(t)}{1 - F(t)} \, dt < a \right) = \Pr(F(\xi) < 1 - \exp[-a/(\bar{e} + \epsilon)]) = 1 - \exp[-a/(\bar{e} + \epsilon)],
\]

since the random variable \( F(\xi) \) has the uniform distribution on \((0,1)\). Similarly, since the integrand is only nonzero if \( s > \bar{e} - \epsilon \), \( G(a) < 1 - \exp[-a/(\bar{e} - \epsilon)] \). From these two bounds on \( G(a) \), it is clear that, as \( \epsilon \to 0 \), \( G \) converges weakly to the exponential distribution with mean \( \bar{e} \). That convergence is comforting to the behavioral theorist, since it would be quite unsettling if a slight change in the specifications of the model led to a large change in predicted behavior.

We shall next present a general analysis of the War of Attrition, taking a point of view which is focused on the equilibrium distribution of observed behavior. Let \( t_i \) and \( t_j \) be independent and uniformly distributed on \((0,1)\), and let \( v = F^{-1} \). (Here we denote by \( F^{-1} \) the generalized inverse of \( F \), defined by \( F^{-1}(x) = \inf\{s : F(s) < x\} \).) Then we can equivalently view the game as one in which the value of the prize to animal \( i \) is the random variable \( v(t_i) \) (i.e., \( v(t_i) \) has the same distribution as \( \xi_i \)), and in which the quitting time chosen by animal \( i \) depends only on its "type" \( t_i \). Let \( \sigma: (0,1) \to [0,\infty) \) be a symmetric equilibrium strategy of the game. Without loss of generality, we may take \( \sigma \) to be nondecreasing;\(^3\) note that \( \sigma^{-1} \) is the distribution of quitting times for each animal. If animal \( i \), with type \( t_i \), competes until time \( a \), his expected payoff is

\[
\int_0^a \sigma(t_i - s) \, ds - \sigma(1 - \sigma^{-1}(a)).
\]

The first-order optimality condition is then

\[
\frac{1}{\sigma(t_i)} = \frac{d\sigma^{-1}/da}{1 - \sigma^{-1}(a)},
\]

and equilibrium requires that \( a = \sigma(t_i) \), or equivalently, that \( t_i = \sigma^{-1}(a) \). Substituting the latter equality into the first-order condition yields:

\[
\frac{1}{\sigma(\sigma^{-1}(a))} = \frac{d\sigma^{-1}/da}{1 - \sigma^{-1}(a)}.
\]

The left-hand side of this equation is a nonincreasing function of \( \sigma \), and the right-hand side is the hazard rate of the distribution of individual quitting times at equilibrium. From this we obtain a new result: The hazard rate of the duration of conflict is nonincreasing at equilibrium. This is an empirical prediction of the model which is independent of the specification of \( e \), and hence of \( F \).

Substituting \( a = \sigma(t_i) \) directly into the first-order condition yields a differential equation which, combined with the necessary boundary condition that \( \sigma(0) = 0 \), in turn yields the equilibrium strategy:

\[
\sigma(t) = \int_0^t \frac{e(s)}{1 - s} \, ds.
\]

Note that this analysis, which is focused on the distribution of equilibrium behavior, offers a unified treatment of the game, independent of the character of \( F \). Furthermore, because \( \sigma^{-1} \) is precisely the distribution of choices predicted at equilibrium, the analysis makes it relatively easy to deduce the empirical implications of the model. (This fact has already been illustrated by our discovery of the declining-hazard-rate property.) Finally, in this "distributional" form the dependence of equilibrium behavior on the specifications of the model can be clearly seen. For example, we can observe that \( \sigma \) increases monotonically with the function \( e \), i.e., the distribution \( \sigma^{-1} \) of quitting times increases stochastically with the distribution \( e^{-1} \) of values. Also, \( \sigma \) varies continuously with \( v \) (in the sense of almost-everywhere convergence), so \( \sigma^{-1} \) varies continuously with \( F \) (in the sense of weak convergence of probability distributions).

These relations are less transparent in the traditional nondistributional analysis.

This model possesses a special monotonic structure, which we exploited in concluding that animals with higher types would choose later quitting times. In more general models, a description of the equilibrium distributions of actions is not sufficient to specify the strategies of the players: It is necessary to indicate which types will take the various actions. In order to preserve the analytical advantages noted in the preceding paragraph, in the following sections we will represent the strategies of players by the joint distributions of their types and actions, and will endow each space of "distributional" strategies with the topology of weak convergence of probability measures. In this setting, the convergence of some sequences of pure-strategy games to mixed strategies emerges naturally, and implies convergence of the associated distributions of actions.\(^4\)

From the characterization of equilibrium strategies given above, it can be shown that the game has a symmetric equilibrium point in pure strategies if and only if the distribution \( F \) is atomless. We shall further explore the relationship between atomless information structures and the existence of pure-strategy equilibria in \S 5.

3. The formal model and the existence theorem. There are six formal elements in our model. The first four are:

(i) the set of players: \( N = \{1, 2, \ldots, n\} \).

(ii) the set of types for each player: \( \{T_i\}_{i \in N} \). Each \( T_i \) is a complete, separable metric space.

(iii) the set of actions available to each player: \( \{A_i\}_{i \in N} \). Each \( A_i \) is a compact metric space.

(iv) the set of possible states: \( \Omega \), a complete, separable metric space.

Let \( \Omega \equiv \Omega_1 \times \cdots \times \Omega_n \) and let \( A \equiv A_1 \times \cdots \times A_n \). Then the last two elements are:

(v) the payoff functions: \( \{U_i\}_{i \in N} \). Each \( U_i \) is a bounded, measurable function from \( T \times A \) into \( \mathbb{R} \).

(vi) the information structure: \( \eta \), a probability measure on the Borel subsets of \( \Omega \), associated with the information structure \( \eta \) is a marginal distribution on each \( T_i \).

\(^3\) Other instances of the "point-opening" approach which unifies pure and mixed strategies are reported in [13] and in [21, footnote 8].
How do the equilibria in these two cases compare? Let $G$ be the distribution of quitting times in the second case; then

$$G(a) = \Pr(\sigma(t) < a) = \int_0^a \frac{f(t)}{1 - F(t)} \, dt < a.$$

Assume that $F$ concentrates its mass on a neighborhood $(\tilde{\delta} - \epsilon, \tilde{\delta} + \epsilon)$. Then the integrand in the expression above is nonzero only if $t < \tilde{\delta} + \epsilon$ and so

$$G(a) > \Pr \left( (\tilde{\delta} + \epsilon) \int_0^a \frac{f(t)}{1 - F(t)} \, dt < a \right)$$

$$= \Pr \left( F(\tilde{\delta}) < 1 - \exp \left( -a / (\tilde{\delta} + \epsilon) \right) \right)$$

$$= 1 - \exp \left( -a / (\tilde{\delta} + \epsilon) \right),$$

since the random variable $F(\tilde{\delta})$ has the uniform distribution on $(0, 1)$. Similarly, since the integrand is only nonzero if $t > \tilde{\delta} - \epsilon$, $G(a) < 1 - \exp(-a / (\tilde{\delta} - \epsilon))$. From these two bounds on $G(a)$, it is clear that, as $\epsilon \to 0$, $G$ converges weakly to the exponential distribution with mean $\tilde{\delta}$. That convergence is comforting to the behavioral theorist, since it would be quite unsettling if a slight change in the specifications of the model led to a large change in predicted behavior.

We shall next present a general analysis of the War of Attrition, taking a point of view which is focused on the equilibrium distribution of observed behavior. Let $t_i$ and $t_j$ be independent and uniformly distributed on $(0, 1)$, and let $\sigma \equiv F^{-1}$. (Here we denote by $F^{-1}$ the generalized inverse of $F$, defined by $F^{-1}(t) = \sup\{s | F(s) < t\}$.) Then we can equivalently view the game as one in which the value of the prize to animal $i$ is the random variable $\sigma(t)$ (i.e., $\sigma(t)$ has the same distribution as $t_i$), and in which the quitting time chosen by animal $i$ depends directly on its "type" $t_i$. Let $\sigma:(0,1) \to [0,\infty]$ be a symmetric equilibrium strategy of the game. Without loss of generality, we may take $\sigma$ to be nondecreasing; 3 note that $\sigma^{-1}$ is the distribution of quitting times for each animal. If animal $i$, with type $t_i$, competes until time $a$, his expected payoff is

$$\int_0^a (\nu(t_i) - s) \sigma^{-1}(t) - a(1 - \sigma^{-1}(a)).$$

The first-order optimality condition is then

$$\frac{1}{\nu(t_i)} = \frac{\sigma^{-1}(a)}{1 - \sigma^{-1}(a)},$$

and equilibrium requires that $\sigma = \sigma(t_i)$, or equivalently, that $t_i = \sigma^{-1}(a)$. Substituting the latter equality into the first-order condition yields:

$$\frac{1}{\nu(\sigma^{-1}(a))} = \frac{\sigma^{-1}(a)}{1 - \sigma^{-1}(a)}.$$

The left-hand side of this equation is a nonincreasing function of $a$, and the right-hand side is the hazard rate of the distribution of individual quitting times at equilibrium. From this we obtain a new result: The hazard rate of the duration of conflict is nonincreasing at equilibrium. This is an empirical prediction of the model which is independent of the specification of $\sigma$, and hence of $F$.

Substituting $\sigma = \sigma(t_i)$ directly into the first-order condition yields a differential equation which, combined with the necessary boundary condition that $\sigma(0) = 0$, in turn yields the equilibrium strategy:

$$\sigma(t_i) = \int_0^t \frac{f(s)}{1 - a} \, ds.$$

Note that this analysis, which is focused on the distribution of equilibrium behavior, offers a unified treatment of the game, independent of the character of $F$. Furthermore, because $\sigma^{-1}$ is precisely the distribution of choices predicted at equilibrium, the analysis makes it relatively easy to deduce the empirical implications of the model. (This fact has already been illustrated by our discovery of the declining-hazard-rate property.) Finally, in this "distributional" form the dependence of equilibrium behavior on the specifications of the model can be clearly seen. For example, we can observe that $\sigma$ increases monotonically with the function $\nu$, i.e., the distribution $\sigma^{-1}$ of quitting times increases stochastically with the distribution $\sigma^{-1} = F$ of values. Also, $\sigma$ varies continuously with $\nu$ (in the sense of almost-everywhere convergence), so $\sigma^{-1}$ varies continuously with $F$ (in the sense of weak convergence of probability distributions). These relations are less transparent in the traditional nondistributional analysis.

This model possesses a special monotonic structure, which we exploited in concluding that animals with higher types would choose later quitting times. In more general models, a description of the equilibrium distributions of actions is not sufficient to specify the strategies of the players: It is necessary to indicate which types will take the various actions. In order to preserve the analytical advantages noted in the preceding paragraph, in the following sections we will represent the strategies of players by the joint distributions of their types and actions, and will endow each space of "distributional" strategies with the topology of weak convergence of probability measures.

In this setting, the convergence of some sequences of pure strategies to mixed strategies emerges naturally, and implies convergence of the associated distributions of actions. 4 From the characterization of equilibrium strategies given above, it can be shown that the game has a symmetric equilibrium point in pure strategies if and only if the distribution $F$ is atomless. We shall further explore the relationship between atomless information structures and the existence of pure-strategy equilibria in §5.

3 Indeed, for every strategy of one competitor, the other has a nondecreasing best response. In [6], the assumption that $\nu$ is increasing is motivated by that statement: "an animal is sometimes hungry and sometimes less so. It is common sense that it should be willing to compete more strongly for food when hungry."

4 Other instances of the "point-opening" approach which unifies pure and mixed strategies are reported in [12] and in [21, footnote 8].

---

**DISTRIBUTIONAL STRATEGIES FOR GAMES WITH INCOMPLETE INFORMATION**

---

323
A pure strategy is a measurable function \( \rho_i: T_i \to A_i \). In the case where \( T_i \) is uncountable, it was observed by Aumann [2] that a mixed strategy cannot be accepted by definition as a measure on the set of pure strategies. To find a more appropriate definition, he reasoned as follows.

Let us recall the intuitive meaning of a mixed strategy: it is a method for choosing a pure strategy by the use of a random device. Physically, one tosses a coin, and according to which side comes up chooses a corresponding pure strategy; or, if one wants to randomize over a continuum of pure strategies, one uses a continuous roulette wheel. Mathematically, the random device—the sides of the coin, or the set of points on the edge of the roulette wheel—constitutes a probability measure space, sometimes called a sample space; a mixed strategy is a function from this sample space to the set of all pure strategies. In other words, what we have here is precisely a random variable whose values are pure strategies. We previously attempted to work with something corresponding to the distribution of this random variable; now we propose to use the random variable itself.

It is this idea which underlies the definition of a mixed strategy for player \( i \) as a measurable function \( \sigma_i: [0, 1] \times T_i \to A_i \). Our approach of defining a (distributional) strategy as a measure on \( T_i \times A_i \) provides another way of avoiding measurability problems.

**Definition.** A distributional strategy for player \( i \) is a probability measure \( \mu_i \) on the subsets of \( T_i \times A_i \), for which the marginal distribution of \( T_i \) is \( \eta_i \). Formally, this restriction on the marginal distribution is that for all \( S \subseteq T_i \), \( \mu_i(S \times A_i) = \eta_i(S) \). When the players adopt distributional strategies \( \mu_1, \ldots, \mu_n \), the expected payoff \( \pi_i \) to player \( i \) is defined to be:

\[
\pi_i(\mu_1, \ldots, \mu_n) = \int U(t, a) \mu_1(da_1 | t_1) \cdots \mu_n(da_n | t_n) \eta(dt).
\]

There is a simple correspondence between a player's behavioral strategies and his distributional strategies. Given a behavioral strategy \( \beta_i \), the corresponding distributional strategy \( \mu_i \) is defined for each \( S \times B \subseteq T_i \times A_i \) (and, hence, for all Borel subsets of \( T_i \times A_i \)) by

\[
\mu_i(S \times B) = \int \beta(B, t) \eta(dt).
\]

In the reverse direction, for any given distributional strategy \( \mu_i \), the corresponding behavioral strategies are the regular conditional distributions (see [7] for definitions):

\[
\beta_i(B, t) = \mu_i(B | t).
\]

Aumann [2] has shown that there is many-to-one mapping from mixed to behavioral strategies that preserves the players' expected payoffs. We have just seen that there is another many-to-one payoff-preserving mapping from behavioral strategies to distributional strategies, i.e., each distributional strategy corresponds to an equivalence class of behavioral strategies. Since any pair of distinct distributional strategies will generally lead to distinct payoffs and since distinct distributional strategies represent different predictions about a player's behavior in the game (when his type and selected action are observable), distributional strategies give the most parsimonious representation possible of a player's meaningful strategic options.

Consider the following regularity conditions for the games we are studying.

R1: Equicontinuous Payoffs. For every player \( i \) and every \( \varepsilon > 0 \), there is a subset

\[ E \] of \( T \) such that \( \eta(E) > 1 - \varepsilon \) and such that the family of functions \( \{ U_i(t, a) | t \in E \} \) is equicontinuous.

R2: Absolutely Continuous Information. The measure \( \eta \) is absolutely continuous with respect to the measure \( \eta \equiv \eta_1 \times \cdots \times \eta_n \). We denote the density of \( \eta \) with respect to \( \eta \) by \( f \).

A principal requirement imposed by R1 is that for any \( t \) the players' payoffs must be continuous functions of their actions. This aspect of the condition is genuinely restrictive: it rules out many bidding games and games of timing. One cannot, however, prove an existence theorem without some such restriction. For example, there are bidding games for which no equilibrium exists.

The following proposition indicates that a large number of models do meet the requirements of R1.

**Proposition 1.** Each of the following three conditions is sufficient to imply R1:

(a) For each \( i \), \( A_i \) is finite.

(b) For each \( i \), \( U_i: T \times A \to \mathbb{R} \) is a uniformly continuous function.

(c) For each \( i \), and for each \( t \in T \), \( U_i(t, \cdot) \) is (uniformly) continuous with modulus of continuity \( B(t, \cdot) \), and for every \( \varepsilon > 0 \), \( B(t, \cdot) \) is measurable.

The following important consequence of the equicontinuous payoffs condition can be proved using Lusin's Theorem.

**Proposition 2.** In a game with equicontinuous payoffs, the following condition is satisfied:

\[ R^*: \] For every player \( i \) and every \( \varepsilon > 0 \), there are a continuous function \( V_i: T \times A \to \mathbb{R} \) and a subset \( K \) of \( T \) such that if \( \eta(K) > 1 - \varepsilon \), (ii) \( V_i \) has the same bound as \( U_i \), and (iii) \( U_i \) and \( V_i \) agree on \( K \times A \).

Condition R2 is a fairly weak requirement on the joint information of the players. It is always satisfied when the variables \( t_1, \ldots, t_n \) are independent, as well as when \( T \) is finite. It is also satisfied in many applied models. Nevertheless, R2 is a potent assumption. It allows us to express the players' expected payoffs in a convenient manner:

\[
\pi_i(\mu_1, \ldots, \mu_n) = \int_{T \times A} U(t, a) P(t) \, d\mu_1 \cdots d\mu_n. \tag{3.1}
\]

The frequent applicability of R2 is emphasized by the following proposition.

**Proposition 3.** Each of the following three conditions is sufficient to imply R2:

(a) For each \( i \), \( T_i \) is finite or countable.

(b) The variables \( t_1, \ldots, t_n \) are independent.

\[ 1 \text{family of functions \( \{ U_i \} \) is equicontinuous if for every \( x \) and every \( \varepsilon > 0 \), there is a \( \delta > 0 \) such that if } |x - y| < \delta \text{ implies } |U_i(x) - U_i(y)| < \varepsilon \text{ for every } a. \]

\[ 2 \text{a measure } \eta \text{ is absolutely continuous with respect to another measure } Q \text{ if for every } S, Q(S) = 0 \text{ implies } \eta(S) = 0. \]

\[ 3 \text{consider the two-person game in which } T_1 = \{10 \}, T_2 = \{10, 20\}, A_1 = A_2 = [0, 30], \text{and } \]

\[ U(10, a_1) = \begin{cases} 10 - a_1 & \text{if } a_1 > 10, \\ 10 - a_1 & \text{if } a_1 < 10, \\ 0 & \text{otherwise}. \end{cases} \]

Suppose \( \text{Pr}(t_2 = 20) = \text{Pr}(t_2 = 10) = \frac{1}{2} \). This simple bidding game has no Nash equilibrium.

\[ 4 \text{for an example in which R2 does not hold, let } P \text{ be the uniform distribution on the unit square and let } \]

\( Q \) be the uniform distribution on the diagonal of the square. Then \( \eta = (P + Q)/2 \) is not an absolutely continuous information structure.
(c) There exists some product measure \( \tilde{\nu} = \lambda_0 \times \cdots \times \lambda_n \) on \( T_0 \times \cdots \times T_n \) such that \( \eta \) is absolutely continuous with respect to \( \tilde{\nu} \).

**THEOREM 1** (Existence Theorem). If a game has equicontinuous payoffs and absolutely continuous information (i.e., it satisfies R1 and R2), then there exists an equilibrium point in distributional strategies.

**PROOF.** We verify that conditions hold that are sufficient for the application of Glicksberg's existence theorem.11

In view of the tightness32 of \( \eta \) and hence of each \( \eta_i \) and the compactness of the action spaces, each player's set of distributional strategies is a tight set of probability measures; also, it is easy to check that the set is closed in the weak topology. By Prohorov's Theorem,33 it follows that the strategy sets are compact metric spaces in the weak topology. Convexity of these sets is also easy to check.

Since the density \( f \) with respect to \( \tilde{\nu} \) is \( \tilde{\nu} \)-integrable, there exists a sequence \( \{ \tilde{f}_n \} \) of bounded continuous functions such that

\[
\int f(t) - \tilde{f}_n(t) \, \tilde{\nu}(dt) \to 0.
\]

Also, using R1*, we can approximate any \( U_i \) by a continuous function \( V_i \). Let \( b \) be a bound on \( U_i \) and let \((\mu_1^0, \ldots, \mu_n^0)\) be a sequence of strategy n-tuples converging to \((\mu_1, \ldots, \mu_n)\). Then using (3.1),

\[
\sigma_1(\mu_1^0, \ldots, \mu_n^0) = \int_{T^X} V_i(t, \sigma) \tilde{f}_n(t) \eta(t, dt) \tilde{\nu}(dt) \to \int_{T^X} V_i(t, \sigma) \tilde{f}(t) \eta(t, dt) \tilde{\nu}(dt) + 2K^B.
\]

An expression similar to (3.2) can be written for \( \sigma_i(\mu_1, \ldots, \mu_n) \). Since the integrand in (3.2) is bounded and continuous, it follows for all pairs \((b, \epsilon)\) that

\[
\limsup_{k \to \infty} |\sigma_i(\mu_1^k, \ldots, \mu_n^k) - \sigma_i(\mu_1, \ldots, \mu_n)| < 2B \int |f(t) - \tilde{f}_n(t)| \tilde{\nu}(dt) + 4eB.
\]

For large \( b \) and small \( \epsilon \), this bound approaches zero. Hence, \( \sigma_i \) is continuous. From (3.1), \( \sigma_i \) is linear.

In summary, when distributional strategies are topologized by weak convergence, the players' strategy sets are compact, convex metric spaces and the payoff functions are continuous and linear. By Glicksberg's theorem, an equilibrium exists.

4. The convergence theorem. Having proved the existence of a Nash equilibrium, we turn our attention to sequences of games to study how variations in the specifications of a game affect the game's equilibria. Throughout the analysis, we hold the type space \( T \) fixed and assume that R1 and R2 hold. We index games in the sequence by \( k \). In the \( k \)-th game, \( \eta^k \) is the distribution of types, and we define \( \tilde{\nu}^k = \eta^k \times \cdots \times \eta^k \) and \( f^k = f \, d\tilde{\nu}^k \). The set of actions available to player 1 is a compact set \( A_1^k \) and his payoff function is \( U_1^k \). The corresponding items in the \( \ast \)-game are \( \eta^\ast, \tilde{\nu}^\ast, f^\ast, A_1^\ast \) and \( U_1^\ast \). Let \((\mu_1^k, \ldots, \mu_n^k)\) be an equilibrium point of the \( k \)-th game.

**THEOREM 2** (Convergence Theorem). Suppose that each game has equicontinuous payoffs (R1) and absolutely continuous information (R2). If for all \( i \in N \),

(i) \( \{ \mu_i^k \} \) converges weakly to \( \mu_i^\ast \), and hence \( \{ \tilde{\nu}^k \} \) converges weakly to \( \tilde{\nu}^\ast \),

(ii) \( \{ U_i^k \} \) converges uniformly to \( U_i^\ast \),

(iii) \( \{ f^k \} \) converges uniformly to \( f^\ast \) on every compact subset of \( T \),

(iv) \( U_i^\ast \) is continuous on \( T \times A \) and \( f^\ast \) is continuous almost everywhere \( \tilde{\nu}^\ast \), and

(v) \( \{ A_i^k \} \) converges in the Hausdorff metric to \( A_i^\ast \), then \((\mu_1^\ast, \ldots, \mu_n^\ast)\) is an equilibrium of the \( \ast \)-game.

**PROOF.** Suppose, contrary to the theorem, that player 1 has a pure strategy \( \sigma^\ast \) in the \( \ast \)-game which raises his expected payoff by some positive amount \( \alpha \) over his payoff from playing \( \mu_1^\ast \). Notice that a pure strategy in distributional form is simply a probability measure concentrated on the graph of a classical pure strategy. Then, by condition (v) of the theorem, there exists a sequence \( \{ \sigma_i^k \} \) of pure strategies, viewed as functions, that converges uniformly to \( \sigma^\ast \), where \( \sigma^\ast \) is a feasible strategy in the \( k \)-th game.

Arguing as in the proof of the Existence Theorem, one can show that: 

(a) \( \lim_{k \to \infty} \sigma_i(\mu_1^k, \ldots, \mu_n^k) = \sigma_i(\mu_1^\ast, \ldots, \mu_n^\ast) \), and

(b) \( \lim_{k \to \infty} \sigma_i(\epsilon, \mu_1^k, \ldots, \mu_n^k) = \sigma_i(\epsilon, \mu_1^\ast, \ldots, \mu_n^\ast) \).

Also, by assumption,

(c) \( \sigma_i(\sigma^\ast, \mu_1^k, \ldots, \mu_n^k) > \sigma_i(\epsilon, \mu_1^k, \ldots, \mu_n^k) + \alpha \).

From (a), (b), and (c) it follows that for all sufficiently large \( k \), the strategy \( \sigma_i^k \) is better than \( \mu_1^\ast \) in the \( k \)-th game, contradicting our hypothesis that each \((\mu_1^k, \ldots, \mu_n^k)\) is an equilibrium point.

Condition (iii) of the theorem is noteworthy: it is not sufficient that the \( \eta^k \)'s converge weakly to \( \eta^\ast \), as the following example shows.

**Example 1.** A Bayesian statistical decision problem is a game pitting one strategic player (the statistician) against Nature. We pose the standard estimation problem in which the statistician must estimate an unknown parameter \( \theta \). Let \( T = [0, 1] \), \( T_1 = T_2 = [0, 1] \). In this problem, one often supposes that there is a quadratic loss function: \( U(\theta, a) = (\theta - a)^2 \). We define a sequence of games, in which the information structure for the \( k \)-th game is concentrated on 2k points:

\[
\Pr(\theta^k = 0, \theta^k = j/k) = 1/(2k^2) \quad \text{for} \quad j = 1, \ldots, k,
\]

\[
Pr(\tilde{\theta}^k = 1, \tilde{\theta}^k = (2j - 1)/(2k)) = 1/(2k^2) \quad \text{for} \quad j = 1, \ldots, k.
\]

The information structure for each game conveys perfect information about \( \theta^k \). If \( 2k \tilde{\theta}_1^0 \) is even, then \( \tilde{\theta}_0 = 0 \); if it is odd, then \( \tilde{\theta}_0 = 1 \). Obviously, the optimal strategy in the \( \ast \)-game is:

\[
\sigma^\ast(\tilde{\theta}_1) = \begin{cases} 0 & \text{if } 2k \tilde{\theta}_1 \text{ is even}, \\ 1 & \text{if } 2k \tilde{\theta}_1 \text{ is odd}. \end{cases}
\]

Passing to the weak limit, the information structure becomes:

\[
Pr(\tilde{\theta}^0 = 0, \tilde{\theta}^0 < a) = Pr(\tilde{\theta}^0 = 1, \tilde{\theta}^0 < a) = a/2.
\]

For this information structure, \( \tilde{\theta}_0 \) and \( \tilde{\theta}_1 \) are independent. Thus, \( \tilde{\theta}_1 \) conveys no information about \( \tilde{\theta}_0 \). The optimal strategy under this null information structure is:

11We refer to the following result, which can be extracted from [14]; related results appear in [9] and [19].

Let the players' strategy spaces be nonempty, compact, convex subsets of convex Hausdorff linear topological spaces. Let the payoff functions be continuous on the product of the strategy sets, and let each player's payoff function be quasiconcave in his strategy. Then an equilibrium point exists.

12A set of probability measures on a metric space is called tight if for every \( \epsilon > 0 \) there is a compact set \( K \) such that for every \( \mu \in \mathcal{M} \) in the set of measures, \( \mu(K) > 1 - \epsilon \). Any single measure \( \mu \) on a complete separable metric space is tight. See [5, Theorem 1.4].

13See [5, Theorem 6, p. 246].

14It suffices to consider sequences (rather than nets) because the domain of \( \eta \) is a finite product of metric spaces.

15It is well known that in an estimation problem with a quadratic loss function, the optimal estimate is the posterior expectation of the unknown parameter (cf. [10, p. 228]).
The weak limit \( \sigma^* \) of the sequence \( \{\sigma^i\} \) is quite different (and nonoptimal), calling for the player to choose his estimate to be either 0 or 1, each with probability \( \frac{1}{2} \).

The following example highlights the role of assumption R2.

**Example 2.** Consider the following variant of the "Battle of the Sexes" game. Let \( T_1 = T_2 = [0,1] \) and let \( A_1 = A_2 = \{1,2\} \). Assume that the payoffs are independent of the types, and are given by the following table:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Suppose that the information structure is given by \( Pr(\tilde{t}_i < u, \tilde{t}_i < v) = min(u, v) \), where \( u \) and \( v \) are numbers in \([0,1] \). Thus, \( \tilde{t}_i = \tilde{t}_2 \), and these variables are uniformly distributed. Now consider the pure strategies

\[
\sigma^i(t) = \begin{cases} 
1 & \text{if the integer part of } k \tilde{t}_i \text{ is odd,} \\
2 & \text{otherwise.}
\end{cases}
\]

If both players adopt the strategy \( \sigma^i \), perfect coordination is achieved and the equilibrium pair is an equilibrium point. The limit of this sequence of pure strategies is the following distributional strategy for player \( i \):

\[
Pr(\tilde{t}_i < u, a_i = 1) = Pr(\tilde{t}_i < u, a_i = 2) = u/2.
\]

Equation (4.1) asserts that the player ignores his information and randomizes his choice of action, choosing each action with probability \( \frac{1}{2} \). This "limit" is not an equilibrium: a better response for player \( i \) would be to choose action \( i \) with certainty. Thus, the set of equilibria of this game is not closed in the weak topology, and hence the Convergence Theorem cannot apply to this game.

5. Pure strategies. Game-theoretic models are often criticized for their reliance on mixed-strategy equilibrium points. Critics argue that mixed strategies have no role in a behavioral theory: people do not base their decisions on the roll of a die or the toss of a coin.

There are several kinds of responses one might make to such criticisms. First, one can challenge the premise that mixed strategies are not actually observed. Close decisions are often made on the basis of minor distinctions or simple whimsy, factors which are hardly less random than roulette wheels. Second, one can claim that the critics have failed to show that there is any observable difference between mixed and pure strategic behavior. Third, models without pure-strategy equilibria may nevertheless have pure-strategy \( \epsilon \)-equilibria for every positive \( \epsilon \). If these "close" to the mixed-strategy equilibria in some appropriate sense, and if the \( \epsilon \)-equilibrium concept seems empirically justifiable, then mixed strategies can be viewed as a convenient technical device for behavioral modeling. Finally, one can concede to the critics and try to identify classes of games for which pure-strategy equilibria exist.

The two theorems that we offer in this section address this whole range of possible responses. The Dunschness Theorem asserts that if a player's type has an atomless distribution, then his set of pure strategies is dense in his entire strategy set. It then follows that if one can only observe points \( T_i \times A_i \) subject to some continuous measurement error, pure strategies and mixed strategies are empirically indistinguishable.

**Theorem 3 (Dunschness Theorem).** Suppose that \( \eta \) is atomless. The player \( i \)'s set of pure strategies is dense in his set of distributional strategies.

To prove this, fix a distributional strategy \( \mu_i \) for player \( i \), and fix \( \epsilon > 0 \). Since \( A_i \) is compact, there exists a finite partition \( B_1, \ldots, B_n \) of \( A_i \), i.e., a partition such that each \( B_j \) has radius less than \( \epsilon \). Since \( T_i \) is complete and separable, \( \eta \) is tight [5, Theorem 4.1]. Therefore \( T_i \) can be partitioned into \( \{K, S_j\} \), where \( K \) is compact and \( \eta(S_j) < \epsilon \). Also, \( K \) has a finite partition \( \{S_1, \ldots, S_n\} \). Since \( \eta \) is atomless, each \( S_j \) can in turn be partitioned into sets \( S_{j1}, \ldots, S_{ja} \), such that \( \eta(S_{ji})/\eta(S_j) = \mu_i(B_i|S_j) \) for \( i = 1, \ldots, k \) (cf. [12], or [8, §2.3, problem 23]).

Fix any points \( b_{1j}, \ldots, b_{aj} \) in \( B_{i1}, \ldots, B_{ia} \), and define a pure strategy \( \sigma_i : T_i \rightarrow A_i \), by \( \sigma_i(t) = b_{ij} \) for all \( t \) in \( S_{ij} \). It is routine to verify that as \( \epsilon \to 0 \), \( \sigma_i \) converges weakly to \( \mu_i \) (cf. [4, p. 603]).

In the statement and proof of the Dunschness Theorem, we have assumed neither equicontinuous payoffs nor absolutely continuous information. In the course of proving the Existence Theorem, these two conditions were shown to imply that each player's expected payoff is a continuous function of the \( n \)-tuple of strategies. Thus, these continuity conditions, together with the Dunschness Theorem, ensure that there are pure-strategy \( \epsilon \)-equilibrium points arbitrarily near any mixed-strategy equilibrium point. In view of the Existence Theorem, we have the following result.

**Corollary.** If a game satisfies the equicontinuous payoffs and absolutely continuous information conditions (R1 and R2), if each \( \eta_i \) is atomless, and if the action spaces are compact, then for every \( \epsilon > 0 \) there exists a pure-strategy \( \epsilon \)-equilibrium point.

Adapting terminology introduced by Radner and Rosenthal [27] to our model, we say that a pure strategy \( \sigma_i \) is a purification of the strategy \( \mu_i \), if two conditions are met:

- (5.1) For almost every \( t_i, \sigma_i(t) \) lies in the support of \( \mu_i(\cdot|t_i) \). (Consequently, if \( \mu_i \) is a best response to some \((n - 1)\)-tuple of strategies \((\mu_1, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_n) \) and if \( R_i \) holds, then \( \sigma_i \) is also a best response.)

- (5.2) For every \( \epsilon > 1 \) and every \((n - 1)\)-tuple \((\mu_1, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_n) \) of strategies for players \( 1, \ldots, n \), substituting \( \sigma_i \) for \( \mu_i \), preserves \( i \)'s expected payoff: \( \pi_i(\mu_i, \sigma_{-i}) = \pi_i(\sigma_i, \mu_{-i}) \).

It is clear from the definition that if \( (\mu_1, \ldots, \mu_n) \) is an equilibrium point and \( \sigma_i \) is a purification of \( \mu_i \), then \( (\sigma_1, \ldots, \sigma_n) \) is also an equilibrium point. Radner and Rosenthal have shown that if (i) the players' types are mutually independent, (ii) each \( \eta_i \) is atomless, (iii) each player's payoff depends only on his own type \( t_i \) and the list of actions \( a \) (that is, \( U_i(U_i(a)) \)), and (iv) the action spaces are all finite, then each strategy of each player has a purification.

In a paper studying statistical decision problems, Dvoretzky, Wald and Wolfowitz [11] proved that if \( T_i \) is a finite set and \( \eta_i \) is atomless, then for every strategy \( \mu_i \) there is a pure strategy \( \sigma_i \) satisfying condition (5.2) and the following condition:

- (5.3) Conditional on any \( t_i \), the distributions induced on \( A_i \) by \( \mu_i \) and \( \sigma_i \) are identical, i.e., for any subset \( B \) of \( A_i \),

\[
\eta(\sigma_i^{-1}(B)|t_i) = \mu_i(B|t_i).n_i(\cdot|t_i) \]
As a corollary to the Dvořák–Wald–Wolffowitz result and to the Existence Theorem, we obtain the Purification Theorem.

**Theorem 4 (Purification Theorem).** If (i) conditional on \( t_0 \), the players’ types are independent, (ii) each \( \eta \) is atomless, (iii) each player’s payoff depends only on the state variable \( t_0 \), his own type \( t \), and the list of actions \( a \) (that is, \( U = U(t_0, t, a) \)), (iv) each player’s action set \( A \) is finite, (v) payoffs are equicontinuous (R1 holds), and (vi) \( T_k \) is a finite set, then each strategy of each player has a purification satisfying conditions (5.1), (5.2), and (5.3). Furthermore, the game has an equilibrium point, and hence has an equilibrium point in pure strategies.

**Proof.** It is direct to verify that conditions (i), (iii), (iv), and (5.3) imply (5.2), so the existence of purifications follows from the Dvořák–Wald–Wolffowitz theorem. Also, it is direct to show that conditions (i) and (vi) imply R2, so existence follows from the Existence Theorem.

Theorem 4 extends the Radner–Rosenthal purification result [27] to allow some players to have information about variables that appear in other players’ payoff functions. Models with this latter feature are known as “adverse selection” models, and play an important role in information economics.

6. Complements and comments. Our formulation of games with incomplete information contains the assumption that an exogenously-specified metric on the type space \( T \) is available. It might appear preferable to simply treat the players’ types as points in a general measurable space, without assuming any topological structure. Indeed, the critical conditions of equicontinuous payoffs (R1) and absolutely continuous information (R2) depend only on measure-theoretic properties of \( T \). Yet the topology on types was necessary in order to define the weak topology on distributional strategies, and this topology played a crucial role in the Existence, Convergence, and Denseness Theorems.

How might we have proceeded, if only a measurable structure on \( T \) had been given?

A natural approach would have been to define endogenously a metric on \( T \) which reflects the nature of the game. In general, a player’s type has two aspects. First, it influences his payoffs, as well as the payoffs of others. Additionally, it affects his beliefs about the types of his competitors, and hence about their behavior. As noted, for example, in [20] and [24], both of these effects are measurable. We here define two metrics (actually, pseudometrics) on \( T \) which correspond to the two effects. For the sake of expository simplicity, our analysis will be in terms of the canonical form of the game (cf. footnote 1; we assume that the state has been integrated out of the payoff functions, and that \( T = T_1 \times \cdots \times T_n \)).

Assume that the players’ payoff functions are bounded, and are continuous on \( A \) for each \( i \) in \( T \). For any player \( i \) and types \( t_i \) and \( t'_i \) in \( T \), define

\[
d'_i(t'_i, t_i) = \sum_{j=1}^n \sum_{a \in A_i} \sup_{t_j \in T_j} |U_j(t'_i, t_{-i}, a) - U_j(t_i, t_{-i}, a)|,
\]

and for \( t \) and \( t' \) in \( T \), define \( d(t', t) = \sum_{i=1}^n d'_i(t'_i, t_i) \). With respect to the product topology on \( T \times A \) induced by this metric on \( T \) and the originally-given topology on \( A \), all of the players’ payoff functions are continuous. (Of course, this statement is trivial if, for example, the metric \( d' \) induces the discrete topology on \( T \).

For any player \( i \) and type \( t_i \) in \( T_i \), let \( \eta_i(.|t_i) \) denote the conditional distribution.

19Yakar Kannai has provided an example which, when suitably extended, shows that this condition cannot be weakened to allow arbitrary compact action spaces.

We assume that a regular conditional distribution exists—as it does, for example, if \( T \) is a Borel space (see [7, Chapter 4]).

**References**


MILGROM: SCHOOL OF ORGANIZATION AND MANAGEMENT, YALE UNIVERSITY, BOX 1A, NEW HAVEN, CONNECTICUT 06520
WEBER: J. L. KELLOGG GRADUATE SCHOOL OF MANAGEMENT, NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS 60201