JOURNAL OF ECONOMIC THEORY **45**, 370–391 (1988)

The Bayesian Foundations of Solution Concepts of Games*

TOMMY CHIN-CHIU TAN

University of Chicago, Chicago, Illinois 60637

AND

SÉRGIO RIBEIRO DA COSTA WERLANG[†]

IMPA and EPGE-FGV, Rio de Janeiro, Brazil

Received April 8, 1986; revised June 23, 1987

We transform a noncooperative game into a Bayesian decision problem for each player where the uncertainty faced by a player is the strategy choices of the other players, the priors of other players on the choice of other players, the priors over priors, and so on. We provide a complete characterization between the extent of knowledge about the rationality of players and their ability to successively eliminate strategies which are not best responses. This paper therefore provides the informational foundations of iteratively undominated strategies and rationalizable strategic behavior (B.D. Bernheim, *Econometrica* **52** (1984), 1007–1028; D. Pearce, *Econometrica* **52** (1984), 1029–1050). Sufficient conditions are also found for Nash equilibrium behavior and a result akin to R. J. Aumann (*Econometrica* **55** (1987), 1–18), on correlated equilibria, is derived with different hypotheses. *Journal of Economic Literature* Classification Numbers: 020, 022, 026. © 1988 Academic Press, Inc.

1. INTRODUCTION

This paper studies non-cooperative games from a Bayesian point of view. A given normal form game is transformed into a Bayesian decision problem in the sense of Savage [21]. The basic uncertainty a player faces in a game is the strategic choice of the other players. Each player therefore

---

* This paper is a substantially expanded and revised version of "The Bayesian Foundations of Rationalizable Strategic Behavior and Nash Equilibrium Behavior." We acknowledge the tremendous help and useful discussions we have had with Roger Myerson who first brought our attention to the Armbruster and Böge [1] paper. Discussions with David Hirshleifer, José Alexandre Scheinkman, and Hugo Sonnenschein have also helped to clarify our ideas.

[†] This author thanks CNPq and IMPA/CNPq for financial support.

0022-0531/88 $3.00

Copyright © 1988 by Academic Press, Inc.
All rights of reproduction in any form reserved.

has a prior over the strategy sets of the other players. In addition, each player is also uncertain about the priors on strategies of the other players and must therefore have priors on their priors, and so on. Hence, beginning with a game in normal form, the Bayesian approach leads to the study of infinite recursion of beliefs for the players—a mathematical object which has been investigated by Armbruster and Böge [1], Böge and Eisele [9], and Mertens and Zamir [15].

Once the transformation of a game into a decision problem has been completed, solution concepts may be derived axiomatically. The infinite recursion of beliefs facilitate explicit axioms on (i) the motivation of players (e.g., they are Bayesian rational); (ii) their beliefs about the motivations about the other players; and (iii) the extent to which players know each others' beliefs (e.g., "Bayesian rationality is common knowledge" or "player i believes that the other players are acting in a correlated manner, but player j does not"). After a set of such axioms has been imposed, the implications of the behavior of the Bayesian player may be derived.

The paper applies the framework to investigation of several different sets of axioms which generate four existing solution concepts—iterative elimination of dominated strategies, rationalizable strategies of Bernheim [5] and Pearce [19], Nash equilibria, and correlated equilibria. The primary aim of this paper is to demonstrate that the infinite recursion of beliefs can be used to derive solution concepts axiomatically and to compare solution concepts by their underlying assumptions on beliefs. Armbruster and Böge [1] pioneered this approach and it is hoped that the application of the framework to four familiar solution concepts will increase the accessibility of this material.

Clearly, a full development of this approach would be to apply infinite recursions to modeling of the decisions faced by each player at each information set in the extensive form. This may lead to axioms and solution concepts which resemble those of the recent refinements of Nash equilibrium such as Cho and Kreps [11]. The more ambitious aim of this line of research is to eventually provide a systematic framework for comparing the many solution concepts currently available as well as deriving new solution concepts.

Section 2 reviews the related literature. Section 3 provides the infinite recursion of beliefs as well as the Bayesian framework and Section 4 begins the Bayesian foundation of solution concepts and imposes rationality as an axiom. Section 5 studies the relationship between axioms on beliefs (such as Bayesian rationality, common knowledge of rationality, and the independence of players) and two solution concepts (rationalizability and the iterative elimination of strictly dominated strategies). It also provides a complete characterization of the implications of "everyone (believes that everyone)^l is rational." Section 6 investigates several different sets of axioms on beliefs which generate Nash equilibrium behavior. Of some novelty are the exchangeability axiom and the example which shows that in a three player game, even though the prior of each player is common knowledge, the resultant behavior may still not be consistent with Nash equilibrium behavior. Aumann's [4] result on the common priors axiom and correlated equilibrium is formulated in Section 7. The hypotheses are slightly different from those of the original and our proof exploits the "common knowledge of a common prior."

2. DISCUSSION OF RELATED LITERATURE

Harsanyi [12] first introduced the infinite recursion of beliefs in his analysis of incomplete information games and it was formally developed in the first three papers mentioned in Section 1. Armbruster and Böge [1], Böge and Eisele [9], and this paper focus on including the strategic choice of the other players in the basic uncertainty faced by each player.

Bernheim [6] discusses the four solution concepts investigated here in very much the same spirit. Although his approach does not examine the knowledge and the layers of knowledge, some of the axioms he examines have a direct analogy with the axioms used below. Bernheim [5] and Pearce [19] provide the basis for the results on rationalizable equilibria in Section 5.

More recently, Brandenburger and Dekel [10] have also applied the Bayesian framework games. They study the relationship between correlated equilibria, subjective correlated equilibria, and rationalizable equilibria.

Reny [20] presents a first step in applying the framework to extensive form games. He points out that the major difficulty in extensive form games is that "rationality is common knowledge" cannot be maintained at every information set as some of these are reached only if the hypothesis is violated. Binmore [8] is also motivated by the proliferation of solution concepts. His use of Turing machines and finite automata offers a contrast to the Bayesian approach pursued here.

3. GAMES AS BAYESIAN DECISION PROBLEMS

This section introduces the Bayesian framework for analyzing the following complete information simultaneous game:

DEFINITION 3.1. A *game with n players*, Γ, is a $2n$-tuple $(A_1, ..., A_n, \bar{U}_1, ..., \bar{U}_n)$, where:

(i) Each A_i, the set of strategies (pure or mixed—the distinction is not important here) available to player i, is a compact metric space.

(ii) If one denotes $A = \prod_{j=1}^{n} A_j$, $\bar{U}_i$ is a function from A to R, which gives the payoffs to player i, for each of the possible combinations of strategies of all players. $\bar{U}_i$ is assumed to be continuous for each i. (When $\{x_i\}_{i=1}^{n}$ is of interest, we let $x_{-i} = \prod_{j \neq i} x_j$.)

The remainder of this section will convert the game Γ into a Bayesian decision problem for player i. [If S is compact and metric, let $\Delta(S)$ be the set of probability measures on S endowed with the Borel σ-algebra. Furthermore, $\Delta(S)$, endowed with the topology of weak convergence of measures, is compact and metrizable—see Billingsley [7] or Hildenbrand [13]].

DEFINITION 3.2. A *Bayesian decision problem* for player i is given by (i) S_i, a compact metric probability space endowed with the Borel σ-algebra (S_i represents all the elements which are uncertain to player i); (ii) A_i, a compact set of actions available to player i; (iii) $U_i\colon A_i \times S_i \to R$, his subjective utility function; and (iv) $P_i \in \Delta(S_i)$, his subjective prior on S_i.

Given a decision problem, one can derive the structure above from more basic axioms, as in Savage [21]. It is important to note that U_i and P_i characterize player i.

DEFINITION 3.3. Let $V_i\colon A_i \times \Delta(S_i) \to R$ be the *expected subjective utility* for player i, when he takes an action a_i, and has prior P_i: $V_i(a_i, P_i) = \int_{S_i} U_i(a_i, s_i)\, dP_i(s_i)$. To avoid unnecessary notation, we will simply write $V(a_i, P_i)$ instead of $V_i(a_i, P_i)$.

DEFINITION 3.4. Player i is *Bayesian rational* when, faced with a Bayesian decision problem, he chooses an action $\bar{a}_i \in A_i$ such that

$$V(\bar{a}_i, P_i) \geq V(a_i, P_i), \qquad \forall a_i \in A_i.$$

The simultaneity of the game Γ implies that player i chooses a strategy in ignorance of the strategies chosen by the other players. Hence the basic uncertainty that player i faces is:

DEFINITION 3.5. Let $S_i^0 = \prod_{j=1}^{N} A_j \equiv A$.

Remark. Player i's own strategy set is included in his basic uncertainty because the study of correlated equilibria requires allowing the actions of all the players to be coordinated by some random device, so that player i's own actions may be uncertain to himself before the realization of the random device. However, player i's prior on his own actions will be ignored until Section 7.

Player i, who is Bayesian, must have a prior on this basic uncertainty. Such a prior is a point in the set of all probability measures on A:

DEFINITION 3.6. Let $S_i^1 = \Delta(A) = \Delta(S_i^0)$.

S_i^0 does not exhaust the uncertainty faced by player i. He realizes that player j must have a prior on S_j^0 as well, and this prior or the first layer belief of player j—a point in S_j^1—is unknown to player i. Consequently, $\prod_{j \neq i} S_j^1$, the first layer beliels of the other players, is also part of the uncertainty faced by player i. Thus, i must have a prior in this as well and such a prior would be a point in:

DEFINITION 3.7. $S_i^2 = \Delta(A \times \prod_{j \neq i} S_j^1) = \Delta(S_i^0 \times \prod_{j \neq i} S_j^1)$.

Notice that we have included A, the basic uncertainty space, in the domain of the second layer beliefs. We are permitting player i to believe that strategies of the players would be correlated with some first layer beliefs of the other players.

Just as S_j^1 was uncertain to player i, so is S_j^2 and so on. Hence each of these layers of beliefs of player j is uncertain to player i and i must successively have priors in:

DEFINITION 3.8. $S_i^l = \Delta(A \times \prod_{j \neq i} S_j^{l-1}) = \Delta(S_i^0 \times \prod_{j \neq i} S_j^{l-1})$.

As before, we are allowing player i to believe that the strategies of the other players may be correlated with their $(l-1)$st layer beliefs. Beginning with the basic uncertainty A, player i's stream of priors is an infinite recursion of beliefs. We assume that a Bayesian player i is completely characterized by his infinite recursion of beliefs (equivalently, his type or psychology) which is a point in:

DEFINITION 3.9. $S_i^\infty = \{(s_i^1, s_i^2, ...) \in \prod_{l \geq 1} S_1^l : (s_i^1, s_i^2, ...)$ satisfies the minimum consistency requirement$\}$.

The minimum consistency requirement, discussed in detail in the Appendix, simply requires that if the probability of an event E may be computed using i's lth layer belief, s_i^l, or his mth layer belief, s_i^m, they must give the same number.

The reader may wonder if we have exhausted all the uncertainty of player i. By extending our earlier argument, it may seem that since S_j^∞ is unknown to player i, he must have a prior on $A \times \prod_{j \neq i} S_j^\infty$. Such a prior would be a point in $\Delta(S_i^0 \times \prod_{j \neq i} S_j^\infty)$ as before. That is, player i should have a prior on the infinite recursion of beliefs of the other players as well. The answer to this question is given by the fundamental mathematical result proven in Armbruster and Böge [1], Böge and Eisele [9], and Mertens and Zamir [15]:

THEOREM 3.1. *$\forall i$, S_i^∞ is compact and metric in the topology induced by the product topology on $\prod_{l \geq 1} S_i^l$. Moreover, $\forall i$ there exists a canonical homeomorphism, ϕ_i: $S_i^\infty \to \Delta(A \times \prod_{j \neq i} S_j^\infty)$. One property of this canonical homeomorphism is that*

$$\operatorname*{marg}_{S_i^0 \times \prod_{j \neq i} S_j^{l-1}} [\phi_i(s_i^\infty)] = s_i^l.$$

Proof. See one of the above sources. Q.E.D.

Remark. The content of this theorem is that the two, S_i^∞ and $\Delta(S_i^0 \times \prod_{j \neq i} S_j^\infty)$, are of the "same size" and although one might consider another layer of priors on $S_i^0 \times \prod_{j \neq i} S_j^\infty$, it would be redundant as this information is already contained in S_i^∞. The theorem allows us to define the universal domain of uncertainty for player i, based on the basic uncertainty space A, to be $A \times \prod_{j \neq i} S_j^\infty$. This space exhausts all that is uncertain to agent i. Consequently, the standard Bayesian approach can be applied by taking the space $S_i = A \times \prod_{j \neq i} S_j^\infty$ as the given domain of uncertainty and then the Bayesian player i must have a prior on this space.

The canonical homeomorphism allows us to speak of s_i^∞, the type of the agent i, and his prior on $A \times \prod_{j \neq i} S_j^\infty$ interchangeably. If an agent is of type s_i^∞, his prior on $A \times \prod_{j \neq i} S_j^\infty$ is given by $\phi_i(s_i^\infty)$. Consequently, his prior on A can be recovered using the canonical homeomorphism and taking the marginal of $\phi_i(s_i^\infty)$ on A. Another feature of $\phi_i(s_i^\infty)$ which we use extensively below is the marginal of $\phi_i(s_i^\infty)$ on S_j^∞. This is player i's prior on the possible types of player j.

We are now in a position to define:

DEFINITION 3.10. Given a game Γ, we define the *Bayesian decision problem associated with Γ* when player i's beliefs are given by $s_i^\infty \in S_i^\infty$ as (i) $S_i = A \times S_{-i}^\infty$; (ii) A_i is the same as A_i for Γ; (iii) $U_i(a_i, s_i) = \bar{U}_i(a_i, \mathrm{Proj}_{A_{-i}}(s_i))$; and (iv) $P_i \in \Delta(S_i)$ is given by $\phi_i(s_i^\infty)$, where ϕ_i is the canonical homeomorphism between S_i^∞ and $\Delta(A \times S_{-i}^\infty) = \Delta(S_i)$.

Remark, (i) It seems that the only relevant probability distribution for player i is the $\mathrm{marg}_{A_{-i}}[\phi_i(s_i^\infty)]$, since $V(a_i, \phi_i(s_i^\infty)) = \int_{S_i} U_i(a_i, s_i)\, d[\phi_i(s_i^\infty)](s_i) = \int_{A_{-i}} \bar{U}_i(a_i, a_{-i})\, d[\mathrm{marg}_{A_{-i}}[\phi_i(s_i^\infty)]]$. This seems to tell us that the only important part of s_i^∞ is the first order belief. The next sections show how one can obtain several different kinds of behavior and solution concepts by imposing restrictions on high orders of beliefs.

(ii) Notice also that since the integral is taken with respect to $\mathrm{Proj}_{A_{-i}}(s_i)$, player i's own belief about his own actions is not relevant in this definition. Hence, in Definition 3.10, only the prior on the other's actions enter into the expected utility calculations. The player's own beliefs about his own action will become relevant only in Section 7 when restrictions are imposed on these beliefs. By an abuse in notation, we define

Definition 3.11. $V(a_i, \phi_i(s_i^\infty)) = V(a_i, s_i^\infty)$.

In summary, the central assumption of the Bayesian approach is:

Axiom [B]. Player i in a game Γ of Definition 3.1 is Bayesian and is characterized by an infinite recursion of beliefs $s_i^\infty \in S_i^\infty$. Given these beliefs, he considers the decision he faces in Γ to be the same as the Bayesian decision problem of Definition 3.10.

4. THE BAYESIAN FOUNDATION OF SOLUTION CONCEPTS

Holding the strategy sets of each player constant, a game Γ is simply given by the payoff function. Hence the space of games is the space of all n-tuples of payoff functions. A solution concept then maps each game into subsets of the strategy sets of players.

Definition 4.1. *Solution concepts.* Let G be the n-product of the space of all continuous functions from $\prod_{i=1}^n A_i$ into the real line. G is the space of games, holding the strategy sets constant. A solution concept Σ is then a correspondence $\Sigma: G \twoheadrightarrow \prod_{i=1}^n A_i$.

Hence given a game $\Gamma \in G$, $\Sigma(\Gamma) \subset \prod_{i=1}^n A_i$ is the solution concept. For example, Σ could be the Nash equilibrium mapping and $\Sigma(\Gamma)$ the set of Nash equilibria of the game Γ.

Since we have taken as an axiom that the decision faced by players in Γ is the same as the Bayesian decision problem, we are led to investigate the beliefs of player i (i.e., restrictions on s_i^∞) and how he behaves given these beliefs. In particular, given a game Γ and a solution concept Σ, what assumptions would result in player i taking an action consistent with the solution concept—that is, choose a strategy in $\mathrm{Proj}_{A_i} \Sigma(\Gamma)$? Similarly, what assumption on beliefs and behavior for all the players would lead them to choose a vector of strategies which lies in $\Sigma(\Gamma)$?

We impose the following axiom from the beginning:

Axiom [B.R.]. In the Bayesian decision problem associated with Γ, when player i has beliefs s_i^∞, he is Bayesian rational in the sense of Definition 3.4.

5. THE BAYESIAN FOUNDATIONS OF ITERATIVELY UNDOMINATED STRATEGIES AND RATIONALIZABLE BEHAVIOR

We consider two axioms on the beliefs of Bayesian rational players and derive solution concepts which are consistent with them.

The first axiom [R] is that player i believes that Bayesian rationality of all the players is common knowledge. In the absence of any other restriction, we show that the player must then take a strategy which is rationalizable against correlated strategies of other players. This is equivalent to iterative elimination of strictly dominated strategies. We also show that every action which is rationalizable against correlated strategies can be supported by an infinite recursion of beliefs for player i which is consistent with "rationality is common knowledge." Consequently, this assumption alone restricts the solution concept to be the iterative elimination of strictly dominated strategies, but no more.

The second axiom [I] is "players act independently is common knowledge." This leads to rationalizability in the sense of Bernheim [5] and of Pearce [19] when combined with Axioms [B], [B.R.], and [R]. Each rationalizable strategy can also be supported by beliefs consistent with the four axioms.

DEFINITION 5.1. *Iteratively undominated strategies or rationalizable strategies against correlated strategies of other players.* Let

$$A_i^0 = A_i$$

$$A_i^l = \left\{ a_i \in A_i \colon \text{There exists } \mu \in \Delta\left(\prod_{j \neq i} A_j^{l-1}\right), \text{ s.t.} \right.$$

$$\left. \int_{A_{-i}} \bar{U}_i(a_i, a_{-i})\, d\mu = \max_{\hat{a}_i \in A_i} \int_{A_{-i}} \bar{U}_i(\hat{a}_i, a_{-i})\, d\mu \right\}.$$

The iteratively undominated strategies of player i are $\bigcap_{l \geqslant 0} A_i^l$.

DEFINITION 5.2. *Knowledge of Bayesian rationality.* Let K_i^l denote the event that i knows everyone (knows everyone)$^{l-1}$ is rational. That is, for $l=1$, $K_i^1 = \{s_i^\infty \colon (a_j, s_j^\infty) \in \operatorname{supp} \operatorname{marg}_{A_j \times S_j^\infty} [\phi_i(s_i^\infty)] \Rightarrow V(a_j, s_j^\infty) = \max_{\hat{a}_j \in A_j} V(\hat{a}_j, s_j^\infty)\}$ and for $l > 1$, $K_i^l = \{s_i^\infty \in K_i^{l-1} \colon s_j^\infty \in \operatorname{supp} \operatorname{marg}_{S_j^\infty} [\phi_i(s_i^\infty)] \Rightarrow s_j^\infty \in K_j^{l-1}\}$.

Remark. When player i has beliefs s_i^∞, then $\phi_i(s_i^\infty)$ is his prior on the strategies and infinite recursion of beliefs of the other players by Axiom [B]. $s_i^\infty \in K_i^1$ implies that if player i believes that it is possible for player j to take strategy a_j when he has beliefs s_i^∞ (that is, $(a_j, s_j^\infty) \in$

supp $\text{marg}_{A_j \times S_j^\infty}[\phi_i(s_i^\infty)]$), then a_j must maximize player j's expected subjective utility given j's prior $\phi_j(s_j^\infty)$. In other words, player i believes that player j is Bayesian rational or satisfies Axioms [B] and [B.R.]. Higher levels of knowledge of rationality are simply defined inductively. Hence $s_i^\infty \in K_i^l$ implies that player i believes that player j's beliefs are in K_j^{l-1}. That is, i believes that j believes rationality is known up to $l-1$ layers. This definition captures the intuitive notion of common knowledge. A more thorough discussion of this and a comparison with Aumann [3] are provided in Tan and Werlang [22].

Axiom [R]. *Bayesian rationality is common knowledge.* $\forall i$, $s_i^\infty \in \bigcap_{l \geqslant 1} K_i^l$.

Lemma 5.1 states that if player i believes that everyone (knows everyone)$^{l-1}$ is rational, i.e., $s_i^\infty \in K_i^l$, then any of the strategies he believes the other player may play—strategies in the support of his prior on the strategies of the other players ($a_{-i} \in \text{supp marg}_{A_{-i}}[\phi_i(s_i^\infty)]$)—must survive l rounds of elimination of strictly dominated strategies. Moreover, Theorem 5.1 states that rational player i then chooses a strategy which survives $l+1$ rounds of elimination of strictly dominated strategies.

LEMMA 5.1. $\forall l$, $\forall i$, $s_i^\infty \in K_i^l \Rightarrow \text{supp marg}_{A_{-i}}[\phi_i(s_i^\infty)] \subset A_{-i}^l$.

Proof of Lemma 5.1. *By induction.* Let $s_i^\infty \in K_i^1$. Then

$$(a_j, s_j^\infty) \in \text{supp marg}_{A_j \times S_j^\infty}[\phi_i(s_i^\infty)] \Rightarrow V(a_j, s_j^\infty) = \max_{A_j} V(\cdot, s_j^\infty)$$

$$\Rightarrow a_j \in A_j^1$$

since $\text{marg}_{A_{-j}}[\phi_j(s_j^\infty)]$ is the required $\mu \in \Delta(A^0_{-j})$ in the definition A_j^1. Since this is true for all $j \neq i$, then $a_{-i} \in \text{supp marg}_{A_{-i}}[\phi_i(s_i^\infty)] \Rightarrow a_{-i} \in A^1_{-i}$. Hence $\text{supp marg}_{A_{-i}}[\phi_i(s_i^\infty)] \subset A^1_{-i}$. Assume that $s_i^\infty \in K_i^l \Rightarrow \text{supp marg}_{A_{-i}}$ $[\phi_i(s_i^\infty)] \subset A^l_{-1}$. We shall show that $s_i^\infty \in K_i^{l+1} \Rightarrow \text{supp marg}_{A_{-i}}[\phi_i(s_i^\infty)] \subset A_{-i}^{l+1}$.

Let $s_i^\infty \in K_i^{l+1}$. Then for $(a_j, s_j^\infty) \in \text{supp marg}_{A_j \times S_j}[\phi_i(s_i^\infty)]$, we know by the definition of K_i^{l+1} that $s_j^\infty \in K_j^l$. Hence by the inductive hypothesis, $\text{supp marg}_{A_{-j}}[\phi_j(s_j^\infty)] \subset A^l_{-j}$. Now, $s_i^\infty \in K_i^{l+1} \subset K_i^1$ by definition, so that $V(a_j, s_j^\infty) = \max_{A_j} V(\cdot, s_j^\infty)$. Hence, $a_j \in A_j^{l+1}$ since $\text{marg}_{A_{-j}}[\phi_j(S_j^\infty)] \in \Delta(A^l_{-j})$ is the required $\mu \in \Delta(A^l_{-j})$ in the definition of A_j^{l+1}. Since this holds for all j, $a_{-i} \in \text{supp marg}_{A_{-i}}[\phi_i(s_i^\infty)] \Rightarrow a_{-i} \in A_{-i}^{l+1}$. $\therefore$ $\text{supp marg}_{A_{-i}}$ $\{\phi_i(s_i^\infty)] \subset A_{-i}^{l+1}$. Q.E.D.

THEOREM 5.1. *If player i is Bayesian rational and $s_i^\infty \in K_i^l$, then he chooses an action $a_i \in A_i^{l+1}$.*

Proof of Theorem 5.1. $s_i^\infty \in K_i^l \Rightarrow \operatorname{supp} \operatorname{marg}_{A_{-i}}[\phi_i(s_i^\infty)] \subset A^l_{-i}$. Hence $\operatorname{marg}_{A_{-i}} \phi_i(s_i^\infty) \in \Delta(A^l_{-i})$. Since i is rational, he choose an action a_i s.t.

$$V(a_i, s_i^\infty) = \max_{A_i} V(\cdot, s_i^\infty) = \int_{A_{-i}} \bar{U}_i(\cdot, a_{-i})\, d[\operatorname{marg}_{A_{-i}} \phi_i(s_i^\infty)]$$

$$= \int_{A^l_{-i}} \bar{U}_i(\cdot, a_{-i})\, d[\operatorname{marg}_{A_{-i}} \phi_i(s_i^\infty)]$$

since the support of $\operatorname{marg}_{A_{-i}} \phi_i(s_i^\infty) \subset A^l_{-i}$ by Lemma 5.1. Hence $a_i \in A_i^{l+1}$. Q.E.D.

The following is immediate:

THEOREM 5.2. *Under Axioms* [B], [B.R.], *and* [R], *player i chooses an action* $a_i \in \bigcap_{l \geqslant 1} A_i^l$.

We now show that rationality does not impose more restrictions than what is implied by Theorem 5.2. The proofs of these two results, as well as several technical lemmata on measurability which are required for their proof, are available in Tan and Werlang [23] and Werlang [24]. The technical lemmata may be of interest for others who wish to apply this framework.

LEMMA 5.2. $\forall i, \forall l, \quad \forall a_i \in A_i^{l+1}$, *there exists* $s_i^\infty \in K_i^l$ *such that* $a_i \in \arg\max_{\hat{a}_i \in A_i} V(\hat{a}_i, s_i^\infty)$.

Proof. See Tan and Werlang [23] or Werlang [24]. Q.E.D.

THEOREM 5.3. *Let* $a_i \in \bigcap_{l \geqslant 1} A_i^l$. *Then there exists* $s_i^\infty \in S_i^\infty$ *such that* $s_i^\infty \in \bigcap_{l \geqslant 1} K_i^l$ *and* $V(a_i, s_i^\infty) = \max_{\hat{a}_i \in A_i} V(\hat{a}_i, s_i^\infty)$.

Proof. See Tan and Werlang [23] or Werlang [24]. Q.E.D.

We now add independence to rationality and state the analogous results pertaining to rationalizable strategies in the sense of Bernheim [5] and Pearce [19]. It should be clear that in a precise sense, rationalizability requires more assumptions about the beliefs about each player than iterative dominance. We let $\otimes_{i=1}^N \Delta(A_i)$ denote the set of probability measures, on $\prod_{i=1}^N A_i$, which are products of measures on each A_i.

DEFINITION 5.3. *Rationalizable strategies* (Bernheim [5] and Pearce [19]). Let $R_i^0 = A_i$ and for $l \geqslant 1$, let

$$R_i^l = \left\{ a_i \in A_i : \text{there exists } \mu \in \bigotimes_{j \neq i} \Delta(R_j^{l-1}) \text{ s.t.} \right.$$

$$\left. \int_{A_{-i}} \bar{U}_i(a_i, a_{-i})\, d\mu = \operatorname*{Max}_{\hat{a}_i \in A_i} \int_{A_{-i}} \bar{U}_i(\hat{a}_i, a_{-i})\, d\mu \right\}.$$

The *rationalizable strategies* of player i are $\bigcap_{l \geqslant 0} R_i^l$.

DEFINITION 5.4. *Knowledge of independence.* Let I_i^l denote the event "player i believes (everyone knows)$^{l-1}$ acts independently." That is, for $l = 1$, $I_i^1 = \{s_i^\infty \in S_i^\infty : \text{marg}_{A_{-i}}[\phi_i(s_i^\infty)] \in \otimes_{j \neq i} \Delta(A_j)\}$ and for $l > 1$, $I_i^l = \{s_i^\infty \in S_i^\infty : s_j^\infty \in \text{supp marg}_{S_j^\infty}[\phi_i(s_i^\infty)] \Rightarrow s_j^\infty \in I_j^{l-1}\}$.

Axiom [I]. *Independence is common knowledge* if $\forall i$, $s_i^\infty \in \bigcap_{l \geqslant 1} I_i^l$.

LEMMA 5.3. *Axioms* [B], [B.R.], *and* $s_i^\infty \in K_i^l \cap I_i^l$ *imply that player i takes an action in* R_i^{l+1}.

THEOREM 5.4. *Axioms* [B], [B.R.], [R], *and* [I] *imply that player i takes a rationalizable strategy.*

LEMMA 5.4. $\forall i$, $\forall l$, $\forall a_i \in A_i^{l+1}$, *there exists* $s_i^\infty \in K_i^l \cap I_i^l$ *such that* $a_i \in \arg\max_{\hat{a}_i \in A_i} V(\hat{a}_i, s_i^\infty)$.

THEOREM 5.5. *Let* $a_i \in \bigcap_{l \geqslant 1} R_i^l$. *Then there exists* $s_i^\infty \in \bigcap_{l \geqslant 1} (K_i^l \cap I_i^l)$ *such that* $V_i(a_i, s_i^\infty) = \max_{\hat{a}_i \in A_i} V(\hat{a}_i, s_i^\infty)$.

The proofs of these four results are straightforward adaptations of the proofs for the analogous results for iterative dominance.

6. THE BAYESIAN FOUNDATIONS OF NASH EQUILIBRIUM

This section is taken from Werlang [24].

The Nash equilibrium is by far the most widely accepted solution concept. It has frequently been argued that it should be a necessary property of any solution concept. We saw in Section 5 that the four axioms we discussed earlier are not sufficient to generate Nash behavior.

6.1. *Coordination and Nash Behavior*

Our first result on Nash behavior is just a formalization of the usual justification for the Nash concept. This is expressed in the classical quote below, taken from Luce and Raiffa [14, p. 173]:

> Nonetheless, we continue to have one very strong argument for equilibrium points: if our non-cooperative theory is to lead to an n-tuple of strategy choices, and if it is to have the property that knowledge of the theory does not lead one to make a choice different from that dictated by the theory, then the strategies isolated by the theory must be equilibrium points.

This justification is a simple restatement of the definition of a Nash equilibrium: they are the only n-tuple of actions which are consistent with common knowledge of the actions taken, as well as of rationality.

Fix a game $\Gamma \in G$. The formalization of the knowledge of a theory (equivalently, solution concept) by the players is simply that the actions this theory predicts are the only actions which are considered possible by the players. The notation is the same as in Sections 3 and 4. In particular, if one wants to refer to "knowledge of a theory $\Sigma(\Gamma)$," where $\Sigma(\Gamma)$ is contained in $A = A_1 \times \cdots \times A_n$, we have:

DEFINITION 6.1.1. Given $\Sigma(\Gamma)$ contained in A, a theory, we say that *player i knows a theory $\Sigma(\Gamma)$* when $s_i^\infty \in \Sigma_i^1 = \{s_i^\infty \in S_i^\infty : \mathrm{Proj}_{A_{-i}} \Sigma(\Gamma) \supset \mathrm{supp}\, \mathrm{marg}_{A_{-i}} [\phi_1(s_i^\infty)]\}$. In other words, player i knows a theory when he thinks other players are going to fulfill their role in this theory.

DEFINITION 6.1.2. *A theory $\Sigma(\Gamma)$ is common knowledge in the eyes of player i* if: $s_i^\infty \in \bigcap_{l \geqslant 1} \Sigma_i^l$, where Σ_i^1 is given above, and $\forall l \geqslant 2 : \Sigma_i^l = \{s_i^\infty \in \Sigma_i^{l-1} : \forall k \neq i : s_k^\infty \in \mathrm{supp}\, \mathrm{marg}_{S_k^\infty} [\phi_i(s_i^\infty)] \Rightarrow s_k^\infty \in \Sigma_k^{l-1}\}$.

Axiom [T]. $\forall i$, the theory $\Sigma(\Gamma)$ is common knowledge in the eyes of player i.

THEOREM 6.1.1. *Assume that $\Sigma(\Gamma) = \{(\tilde{a}_1, ..., \tilde{a}_n)\}$, that is to say, Σ is a single-valued solution concept for Γ under Axiomas* [B], [B.R.], *and* [T]. *Then $\Sigma(\Gamma)$ is a Nash equilibrium. Moreover, any Nash equilibrium is compatible with common knowledge of the theory and common knowledge of rationality.*

Proof. Since player i knows that player k is rational and player k knows the theory, it follows that $\tilde{a}_k$ is a best response to $\tilde{a}_{-k}$, for all $k \neq i$. To check that $\tilde{a}_i$ is a best response to $\tilde{a}_{-i}$, it is enough to carry the same argument above one more layer. Observe that it was necessary to use only $s_i^\infty \in K_i^2 \cap \Sigma_i^3$. The second part of the theorem is immediate. Q.E.D.

This result gives one set of assumptions on beliefs and behavior which justifies the Nash equilibrium concept. However, we feel that the theorem above also shows the weakness of the concept. The Nash equilibrium is played when the actions which are going to be taken are common knowledge, *before they have been taken.* It shows the strong need for coordination in obtaining Nash behavior.

From the proof of Theorem 6.1.1, it was not necessary to assume common knowledge of rationality and the solution concept. Rationality known to two layers and the solution concept known to three were sufficient to show that the solution concept had to be a Nash equilibrium. This is slightly stronger than the "knowledge of the theory" in the quote from Luce and Raiffa above. See Tan and Werlang [23] for the implications of a weaker requirement that "players know that other *may* play the Nash equilibrium" on oligopoly.

6.2. *Knowledge of other Players and Nash Equilibrium*

We now proceed to investigate other, possibly weaker, sets of assumptions which would lead the players to play a Nash equilibrium. First, let us reinterpret a Nash equilibrium in mixed strategies. The traditional view of a mixed strategy is literally that each player chooses a randomization device which "plays" the mixed strategy. In the Bayesian framework, an alternative interpretation of a mixed strategy for player j is that it is player i's prior on the strategies of player j. Hence, in a mixed strategy Nash equilibrium, instead of each player actually randomizing according to their equilibrium strategies, one could have player i's prior be the equilibrium mixed strategies of the other players. Hence, mixed strategies are subjective priors:

DEFINITION 6.2.1. Let $(\mu_1, ..., \mu_n)$ be a mixed strategy Nash equilibrium for the game Γ, where $\mu_i \in \Delta(A_j)$. We say that *the n-tuple of types* $(s_1^\infty, ..., s_n^\infty)$ *plays the Nash equilibrium* $(\mu_1, ..., \mu_n)$ if for all i: $\mathrm{marg}_{A_{-i}}[\phi_i(s_i^\infty)] = \otimes_{k \neq i} \mu_k = \mu_1 \otimes \cdots \otimes \mu_{i-1} \otimes \mu_{i+1} \otimes \cdots \otimes \mu_n$.

DEFINITION 6.2.2. Given an n-tuple $(s_1^\infty, ..., s_n^\infty)$ of beliefs, we say that *player i knows the other players* if: $\mathrm{supp}\ \mathrm{marg}_{S_{-i}} \{\phi_i(s_i^\infty)] = \{s_{-i}^\infty\}$.

The following theorem is a characterization of Nash equilibria in two-person games. The first part of the theorem below is in Armbruster and Böge [1].

THEOREM 6.2.1. *Let Γ be a two-person game. Suppose rationality is common knowledge, and that player* 1 *knows player* 2 *and player* 2 *knows player* 1. *Then they play a mixed strategy Nash equilibrium of the game Γ in the sense of Definition* 6.2.1. *Conversely, if (μ_1, μ_2) is a mixed strategy Nash equilibrium of Γ, there are beliefs (s_1^∞, s_2^∞) such that rationality is common knowledge, and each player knows each other, with the property that (s_1^∞, s_2^∞) plays (μ_1, μ_2) in the sense of Definition* 6.2.1.

Proof. Consider the pair $\mu_1 = \text{marg}_{A_1}[\phi_2(s_2^\infty)]$ and $\mu_2 = \text{marg}_{A_2}[\phi_1(s_1^\infty)]$. We know that $\forall a_1 \in \text{supp}\,\mu_1$, a_1 is a best response to μ_2, since player 2 thinks player 1 is rational, and $\text{marg}_{S_1}[\phi_2(s_2^\infty)] = \{s_1^\infty\}$. Similarly, $\forall a_2 \in \text{supp}\,\mu_2$, a_2 is a best response to μ_1. Thus (μ_1, μ_2) is a mixed strategy equilibrium of the game Γ. Conversely, suppose (μ_1, μ_2) is a mixed strategy Nash equilibrium of the game u. One can construct the infinite hierarchies of beliefs (s_1^∞, s_2^∞) which will play (μ_1, μ_2) by rationalizing in each round every point in the support of one of the mixed strategies by the mixed strategies of the opponent. These infinite hierarchies of beliefs will obviously satisfy the requirements of the theorem. Q.E.D.

This result is appealing in the sense that if two players in a game are so familiar with each other that they know each other's beliefs completely, then if rationality is also common knowledge, each player's prior on the other player's strategic choice must be one of the other player's Nash equilibrium mixed strategies in a two-person game. Moreover, the two priors come from the same Nash equilibrium. Each player would then take a strategy in the support of his own equilibrium mixed strategy.

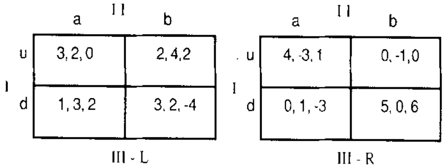
Unfortunately, the result above is not true for games with more than two players. Consider a situation with three players. Each player has beliefs over the actions of the other players. Suppose these beliefs satisfy the following condition: for each player i, the support of the beliefs on the actions of player k $(k \neq i)$ is contained in the set of best responses of player k against player k's beliefs over actions of players who are not k. If there were only two players, the condition above would imply that the two players were playing a mixed strategy Nash equilibrium, according to Definition 6.2.1. With three players, it is not necessarily true that these players have a common prior. Thus, even when all the three players know each other, it is possible that they are not at a Nash equilibrium: this because they may hold prior about the actions of others which are not consistent with a common prior. The next example will illustrate this point.

EXAMPLE 6.2.1. *Common knowledge of rationality and knowledge of each other does not imply Nash behavior in three-person games.* There are three players. The pure strategy sets are $A_1 = \{u, d\}$, $A_2 = \{a, b\}$, and $A_3 = \{L, R\}$. The payoffs are given by the two matrices shown in Fig. 1. The matrix on the left corresponds to player 3 playing L; the matrix on the right, to player 3 playing R. Define $\mu_{ij} \in \Delta(A_j)$, for $i \neq j$, and $i, j = 1, 2, 3$, by

$$\mu_{12} = (1/2a, 1/2b), \qquad \mu_{13} = (1/2L, 1/2R);$$

$$\mu_{21} = (1/3u, 2/3d), \qquad \mu_{23} = (1/3L, 2/3R);$$

$$\mu_{31} = (2/3u, 1/3d), \qquad \mu_{32} = (2/3a, 1/3b).$$



		II	
		a	b
I	u	3,2,0	2,4,2
	d	1,3,2	3,2,-4

III - L

		II	
		a	b
I	u	4,-3,1	0,-1,0
	d	0,1,-3	5,0,6

III - R

FIGURE 1

Then, we have

Set of best responses to $\mu_{12} \otimes \mu_{13} \equiv v_1$ is A_1;

Set of best responses to $\mu_{21} \otimes \mu_{23} \equiv v_2$ is A_2;

Set of best responses to $\mu_{32} \otimes \mu_{31} \equiv v_3$ is A_3.

We now construct three infinite hierarchies of beliefs $(s_1^\infty, s_2^\infty, s_3^\infty)$ such that for every i: $\text{marg}_{A_j \times A_k}[\phi_i(s_i^\infty)] = \mu_{ij} \otimes \mu_{ik}$ for $j, k \neq i$, with $j \neq k$. These hierarchies of beliefs will be such that rationality is common knowledge and for all i: $\text{supp marg}_{S_i \times S_k}[\phi_i(s_i^\infty)] = \{(s_i^\infty, s_k^\infty)\}$ for $j \neq k$, and $j, k \neq i$ (this means that each player knows the other two players). The constructions are simultaneous. The first order beliefs, s_1^1, s_2^1, s_3^1, are given by v_1, v_2, v_3, respectively. The higher order beliefs will all be constructed in the same fashion as the second order beliefs. For example, $s_1^2 \in \Delta(A_2 \times A_3 \times S_2^1 \times S_3^1)$ is given by $s_{21} = s_{11} \otimes \delta\{(s_2^1, s_3^1)\}$, where $\delta\{\cdot\}$ is the probability measure which puts mass 1 on the set $\{\cdot\}$. The hierarchies of beliefs thus built are clearly consistent and satisfy the properties required above. However, $\mu_{21} \neq \mu_{31}$, $\mu_{12} \neq \mu_{32}$, $\mu_{13} \neq \mu_{23}$. Therefore the triple $(s_1^\infty, s_2^\infty, s_3^\infty)$ does not play a mixed strategy Nash equilibrium.

6.3. *The Exchangeability Axiom and the Nash Hypothesis*

A final axiom we study in relation to Nash behavior is exchangeability. Although it is less intuitive than the earlier assumptions, it is weaker and the first result we obtain requires no knowledge of Nash theory as in contrast to the earlier sections. It is therefore somewhat appealing since it derives Nash behavior from more basic assumptions. However, exchangeability itself is a strong assumption.

Axiom [E]. *Exchangeability is common knowledge among players*: $\forall i$: $s_i^\infty \in \bigcap_{l \geqslant 1} E_i^l$, where $\forall i$,

for $l = 1$: $E_i^1 = \{s_i^\infty \in S_i^\infty; \forall j \text{ supp marg}_{(A_j \times S_j^\infty)}[\phi_i(s_i^\infty)] = C_j \times D_j$ for some $C_j \subset A_j$ and $D_j \subset S_j^\infty\}$;

for $l \geqslant 2$: $E_i^l = \{s_i^\infty \in E_i^{l-1}; \forall j, s_j^\infty \in \text{supp marg}_{S_j^\infty}[\phi_i(s_i^\infty)] \Rightarrow s_j^\infty \in E_j^{l-1}\}$.

[In words: the exchangeability hypothesis being common knowledge simply means that it is common knowledge among players that if an action by player j, $a_j \in A_j$, is considered possible by player i, then he also considers it possible when player j is of any of the types s_j^{∞} he believes player j can be.]

Certain games have rationalizable strategies which are rationalized by a player i believing that player j is persistently wrong about himself (i). We require another restriction on beliefs to eliminate these completely inconsistent beliefs. The following assumption requires that i believes that j may be correct about i:

DEFINITION 6.3.2. Player i is said to be *conjecturally consistent* if $\forall j \neq i$: $\forall s_j^{\infty} \in \operatorname{supp} \operatorname{marg}_{S_j^{\infty}} [\phi_i(s_i^{\infty})] \Rightarrow s_i^{\infty} \in \operatorname{supp} \operatorname{marg}_{S_i^{\infty}} [\phi_j(s_j^{\infty})]$.

THEOREM 6.3.1. *Suppose player i satisfies Axioms* [B], [B.R.], [R], *and* [E]. *Let player i be conjecturally consistent. Also let Γ be a two-person game. Then, for $j \neq i$,* $\operatorname{marg}_{A_j}[\phi_i(s_i^{\infty})]$ *is a Nash equilibrium mixed strategy for player j.*

Proof. Let us fix $i = 1$, without loss of generality. Then, if $a_2 \in \operatorname{supp} \operatorname{marg}_{A_2}[\phi_1(s_i^{\infty})]$, we must have, by the fact that $s_1^{\infty} \in K_1^{\infty} \cap E_1^{\infty}$, that a_2 is a best response against $\operatorname{marg}_{A_1}[\phi_2(s_2^{\infty})]$, for every $s_2^{\infty} \in \operatorname{supp} \operatorname{marg}_{S_2^{\infty}} [\phi_1(s_1^{\infty})]$. But by conjectural consistency we have $s_1^{\infty} \in \operatorname{supp} \operatorname{marg}_{S_1^{\infty}} [\phi_2(s_2^{\infty})]$. As player 1 thinks player 2 thinks he is rational ($s_1^{\infty} \in K_1^2$), we must have that $(a_1, s_1^{\infty}) \in \operatorname{supp}[\phi_2(s_2^{\infty})]$, then a_1 is a best response to $\operatorname{marg}_{A_2}[\phi_1(s_1^{\infty})]$. By exchangeability this must be true for every $a_1 \in \operatorname{supp} \operatorname{marg}_{A_1}[\phi_2(s_2^{\infty})]$. Consider the pair $\mu_1 = \operatorname{marg}_{A_1}[\phi_2(s_2^{\infty})]$ for any $s_2^{\infty} \in \operatorname{marg} \operatorname{supp}_{S_{2\eta}^{\infty}} [\phi_1(s_1^{\infty})]$, and $\mu_2 = \operatorname{marg}_{A_2}[\phi_1(s_1^{\infty})]$. Then they form a Nash equilibrium, because $\forall a_1 \in \operatorname{supp} \mu_1$, a_1 is the best response against μ_2, and $\forall a_2 \in \operatorname{supp} \mu_2$, a_2 is the response against μ_1. Notice that we used in the proof only that $s_i^{\infty} \in E_i^2 \cap K_i^2$. Q.E.D.

This result cannot be generalized to n-person games unfortunately. See Werlang [24] for an example and Tan and Werlang [23] for other results involving exchangeability.

7. CORRELATED EQUILIBRIUM AND COMMON PRIORS—A RESULT OF AUMANN

Aumann [4] examines the concept of correlated equilibrium in very much the same spirit as our analysis of iterated dominance. He shows that rationality and a common prior being common knowledge gives rise to a correlated equilibrium. We provide that result in our framework below; the hypotheses and proof here are different from Aumann's and makes clearer the relationship between the assumptions and the result.

It should be noticed that the common prior assumption is stronger than the earlier axioms we have examined. It resuires the highest degree of coordination among the solution concepts. A correlated eqilibrium is perhaps more likely in situations where players had discussed the game and agreed to coordinate their actions with a randomization device before actually playing. Such an agreement may be the basis of a common prior among players being common knowledge. In other situations where such coordination and centralization are unlikely, a correlated equilibrium might be less likely to hold. (See, however, Aumann [4].)

DEFINITION 7.1. *A correlated equilibrium* is given by $P \in \Delta(\prod_{j=1}^{n} A_j)$, a regular probability measure, with well defined conditional probability measures given each a_i such that for every $a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P$,

$$a_i \in \arg\max_{\hat{a}_i \in A_i} \int_{A_{-i}} \bar{U}(\hat{a}_i, a_{-i})\, dP(a_{-i} \mid a_i).$$

The interpretation of this definition (which by Aumann [2] is equivalent to other formulations) is that the realization of the randomizing device tells player i to play a_i. The probability distribution of the randomization is P. Hence, given the recommendation a_i, player i's conditional distribution on the strategies of the other players is $P(\cdot \mid a_i)$. Hence if P is a correlated equilibrium, it must be the case that a_i is a best response to the conditional distribution $P(\cdot \mid a_i)$.

Several interpretations exist for a player's belief about his own action (Definition 7.2(i)). One, as mentioned above, is that it is the recommendation of an agreed upon randomization device. Another is that each player is born knowing himself completely and knowing what he would do in different situations. Hence his belief about his own action is a reflection of his self-knowledge. In this interpretation, Nature is the randomization device. See Aumann [4] for a more complete discussion.

DEFINITION 7.2. We say that *player i's beliefs come from a prior* $P \in \Delta(\prod_{j=1}^{n} A_j)$ if $s_i^{\infty} \in \mathrm{CP}_i^0 = \{s_i^{\infty} \in S_i^{\infty}$: there exists $a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P$ such that

(i) $\operatorname{supp} \operatorname{marg}_{A_i} [\phi_i(s_i^{\infty})] = \{a_i\}$,

(ii) $\operatorname{marg}_{A_{-i}} [\phi_i(s_i^{\infty})] = P(\cdot \mid a_i)\}$.

Player i believes that the other player's beliefs come from the common prior P if:

$$s_i^\infty \in \mathrm{CP}_i^1 = \{s_i^\infty \in S_i^\infty : \forall j \neq i, \forall (a_j, s_j^\infty) \in \operatorname{supp} \underset{A_j \times S_j^\infty}{\operatorname{marg}} [\phi_i(s_i^\infty)],$$

$$\underset{A_{-j}}{\operatorname{marg}} [\phi_j(s_j^\infty)] = P(\cdot \mid a_j)\}.$$

For $l \geqslant 2$, let

$$\mathrm{CP}_i^l = \{s_i^\infty \in \mathrm{CP}_i^{l-1} : \forall j \neq i, \forall s_j^\infty \in \operatorname{supp} \underset{S_j^\infty}{\operatorname{marg}} [\phi_i(s_i^\infty)], s_j^\infty \in \mathrm{CP}_j^{l-1}\}.$$

Axiom [CP]. The common prior $P \in \Delta(\prod_{j=1}^n A_j)$ is common knowledge: $\forall i$, $s_i^\infty \in \bigcap_{l \geqslant 0} \mathrm{CP}_i^l$.

Remark. (i) By constructing the beliefs of player i, one can show that for any $a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P$, there is $s_i^\infty \in \bigcap_{l \geqslant 0} \mathrm{CP}_i^l$ such that $\operatorname{supp} \operatorname{marg}_{A_i} [\phi_i(s_i^\infty)] = \{a_i\}$. In particular, for any i, it can be shown that

$$s_i^\infty \in \bigcap_{l \geqslant 0} \mathrm{CP}_i^l \operatorname{supp} \operatorname{marg}_{A_i} [\phi_i(s_i^\infty)] = \operatorname{supp} \operatorname{marg}_{A_i} P.$$

(ii) Notice that it is only here that player i's belief about his own strategies becomes important since it comes from a common prior.

(iii) In Definition 7.2 we could have defined $\mathrm{CP}_i^0 = \{s_i^\infty \in S_i^\infty : \operatorname{marg}_A \phi_i(s_i^\infty) = P\}$ to begin with and then redefine rationality to mean a_i in $\operatorname{supp} \operatorname{marg}_{A_i} [\phi_i(s_i^\infty)]$ implies that a_i is a best response against $\operatorname{marg}_{A_{-i}} [\phi_i(s_i^\infty)]$. This definition was used in Aumann [4]. However, this would make P a correlated equilibrium by definition. We show in Theorem 7.1 that Axioms [R] and [CP] can be exploited more fully.

(iv) Notice that the common prior is in $\Delta(\prod_{j=1}^n A_j)$ rather than in $\Delta(\prod_{j=1}^n S_j)$. Definition 7.2(i, ii) only restricts player i's first layer beliefs on actions but not higher order beliefs. It is only with common knowledge of a common prior—imposed in the remainder of Definition 7.2 and Axiom [CP] that higher order beliefs are restricted.

In Definition 7.2, there is no presumption of optimality or rationality yet. Definition 7.2(i) requires that i's belief on his own strategies be a unit mass point and that the belief on the actions of the other players be $P(\cdot \mid a_i)$. This we feel captures Aumann's intention that a player's beliefs be a posterior (given the recommendation a_i by the randomization device) derived from a common prior. However, there is no guarantee that i will play a_i. The optimality of a_i against $P(\cdot \mid a_i)$ is derived from a more subtle argument: i believes that j's belief comes from P, moreover, i believes that j believes i

is rational, so that since j may believe that i's action is a_i when i's belief on actions is $P(\cdot\,|\,a_i)$, then a_i must be optimal against $P(\cdot\,|\,a_i)$.

THEOREM 7.1 (Aumann). *Suppose Axioms* [B] *and* [B.R.] *hold. Consider a prior* $P \in \Delta(\prod_{j=1}^{n} A_j)$. *If for any* $(s_1^\infty, ..., s_n^\infty) \in \prod_{j=1}^{n}(\bigcap_{l=0}^{2} \mathrm{CP}_j^l)$, *Axiom* [R] *is satisfied, then* P *is a correlated equilibrium.*

Proof. We need to show that $\forall i$, $\forall a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P$,

$$a_i \in \underset{\hat{a}_i \in A_i}{\arg\max}\, V(\hat{a}_i, P(\cdot\,|\,a_i)). \tag{$*$}$$

Let $s_i^\infty \in \bigcap_{l=0}^{2} \mathrm{CP}_i^l$. Then for all $(a_j, s_j^\infty) \in \operatorname{supp} \operatorname{marg}_{A_j \times S_j^\infty} [\phi_i(s_i^\infty)]$, $\operatorname{marg}_{A_{-j}} [\phi_j(s_j^\infty)] = P(\cdot\,|\,a_j)$ since $s_i^\infty \in \mathrm{CP}_i^1$. Moreover, since $s_i^\infty \in K_i^1$, $a_j \in \arg\max_{\hat{a}_j \in A_j} V(\hat{a}_j, P(\cdot\,|\,a_j))$. This is true for any $a_j \in \operatorname{supp} \operatorname{marg}_{A_j} [\phi_i(s_i^\infty)]$. Moreover, by the remark following Definition 7.2,

$$\bigcup_{s_i^\infty \in \bigcap_{l=0}^{2} \mathrm{CP}_i^l} \operatorname{supp} \operatorname{marg}_{A_i} [\phi_i(s_i^\infty)] = \operatorname{supp} \operatorname{marg}_{A_i} P,$$

then

$$\begin{aligned}
&\bigcup_{s_i^\infty \in \bigcap_{l=0}^{2} \mathrm{CP}_i^l} \operatorname{supp} \operatorname{marg}_{A_j} [\phi_i(s_i^\infty)] \\
&\quad = \bigcup_{s_i^\infty \in \bigcap_{l=0}^{2} \mathrm{CP}_i^l} \operatorname{supp} \operatorname{marg}_{A_j} P(\cdot\,|\operatorname{supp} \operatorname{marg}_{A_i} [\phi_i(s_i^\infty)]) \\
&\quad = \bigcup_{a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P} \operatorname{supp} \operatorname{marg}_{A_j} P(\cdot\,|\,a_i) = \operatorname{supp} \operatorname{marg}_{A_j} P.
\end{aligned}$$

So that $(*)$ is true for $j \neq i$.

What remains is to show this for player i.

Since $s_i^\infty \in \mathrm{CP}_i^2 \cap K_i^2$, for any $s_j^\infty \in \operatorname{supp} \operatorname{marg}_{S_j^\infty} [\phi_i(s_i^\infty)]$, if $(a_i, s_i^\infty) \in \operatorname{supp} \operatorname{marg}_{A_i \times S_i^\infty} [\phi_j(s_j^\infty)]$, $\operatorname{marg}_{A_{-i}} [\phi_i(s_i^\infty)] = P(\cdot\,|\,a_i)$ and $a_i \in \arg\max_{\hat{a}_i \in A_i} V(\cdot, P(\cdot\,|\,a_i))$.

Since this is true for any $a_i \in \operatorname{supp} \operatorname{marg}_{A_i} [\phi_j(s_j^\infty)]$ and

$$\bigcup_{s_i^\infty \in \bigcap_{l=0}^{2} \mathrm{CP}_i^l} \; \bigcup_{s_j^\infty \in \operatorname{supp} \operatorname{marg}_{S_j^\infty} [\phi_i(s_i^\infty)]} \operatorname{supp} \operatorname{marg}_{A_i} (\phi_j(s_j^\infty)) = \operatorname{supp} \operatorname{marg}_{A_i} P,$$

it is true for any $a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P$. Hence, $\forall i$, $\forall a_i \in \operatorname{supp} \operatorname{marg}_{A_i} P$, $a_i \in \arg\max_{\hat{a}_i \in A_i} V(\cdot, P(\cdot\,|\,a_i))$. Q.E.D.

Remark. Notice that since we required only $s_i^\infty \in \bigcap_{l=0}^{2} \mathrm{CP}_i^l \cap K_i^l$ it is weaker than requiring that the prior P and rationality be common knowledge. In Aumann [4], the model implcitily had common knowledge of rationality and the prior built in at every state of the world.

This theorem is really another version of Theorem 6.1.1 in disguise. Correlated equilibria are Nash equilibria of a game with an expanded strategy space.

Aumann [4] discusses the concept of subjective correlated equilibrium. This is the case where players are permitted to have different priors and may believe that other players have different priors. This is, by Theorem 5.2, the same as the iteratively undominated strategies.

APPENDIX

In the definition of S_i^l one could include correlation among all the previous layers of beliefs. This is the approach followed by Mertens and Zamir [15], but given the consistency requirements they have (as we do below), they show this is equivalent to the framework we use here.

Observe that an arbitrary lth order belief contains information about all beliefs of order less than l. An obvious requirement that should hold is that the first order beliefs of player i should be the marginal of his second order belief on his basic uncertainty. We will construct a way of determining the lower order beliefs, given a belief of a certain order. This is the approach of Myerson [17, 18]. However, as it is very enlightening, we think it is worth going through it. Let us impose on a player's beliefs the *minimal consistency requirement*: that if it is possible to evaluate the probability of an event through his lth order beliefs and his kth order beliefs, with $l \neq k$, then both the probabilities agree. Define inductively the functions that will recover the $(l-1)$st order beliefs, given lth order beliefs, by:

For $l \geqslant 2$: $\Psi_i^{l-1}: S_i^l \to S_i^{l-1}$.

For $l = 2$: $\Psi_i^1(s_i^2)(E) = s_i^l(E \times S_{-i}^1)\ \forall E \subset S_i^0$.

For $l \geqslant 3$: by induction on l we assume $(\Psi_j^{l-2})_{j=1}^n$ defined, and $\Psi_i^{l-1}(s_i^l)(E) = s_i^l(\{(s_i^0, (s_j^{l-1})_{j=i}) \in S_i^0 \times S_{-i}^{l-1};\ \ (s_i^0,\ \ (\Psi_j^{l-2}(s_j^{l-1}))_{j \neq i}) \in E\})$, $\forall E \subset S_i^0 \times S_{-i}^{l-2}$.

We have then:

PROPOSITION A.1. *Suppose all players are aware that each of them satisfy the minimum consistency requirement. Then* $\forall i, \forall l \geqslant 2$: $\Psi_i^{l-1}(s_i^l) = s_i^{l-1}$.

Proof. Let us prove the proposition by induction. For $l = 2$, let $E \subset S_i^0$. Then the event E (event = measurable set) is evaluated by s_i^1 as $s_i^1(E)$. However, E is the same as $E \times S_{-i}^1$ evaluated by s_i^2. Therefore by the consistency of the players: $s_i^1(E) = s_i^2(E \times S_{-i}^1) = \Psi_i^1(s_i^2)$. They also know $s_j^1 = \Psi_j^1(s_j^2)$, because they know the others are consistent. Let us assume it is true for $l \geqslant 2$. We will prove that it is true for $l+1$. If it is true for l, we know that $\forall j=1, ..., n$, $s_j^{l-1} = \Psi_j^{l-1}(s_j^l)$. Given the event $E \subset S_i^0 \times S_{-i}^{l-1}$, define $E^* \subset S_i^0 \times S_{-i}^l$ by

$$E^* = \{(s_i^0, (s_j^l)_{j \neq i}) \in S_i^0 \times S_{-i}^l; (s_i^0, (\Psi_j^{l-1}(s_j^l))_{j \neq i}) \in E\}.$$

By the induction hypothesis we have that

$$E^* = \{(s_i^0, (s_j^l)_{j \neq i}) \in S_i^0 \times S_{-i}^l; (s_i^0, (s_j^{l-1})_{j \neq 1}) \in E\}.$$

Therefore E^* and E are the "same" events (same in the sense used before: one is true if and only if the other is). Hence by the hypothesis $s_i^l(E) = s_i^{l+1}(E^*)$. But $\Psi_i^l(s_i^{l+1})(E) = s_i^{l+1}(E^*)$, and so the result follows. Q.E.D.

Given the proposition above, we will restrict ourselves to consistent beliefs. Therefore the set of all possible beliefs for player i is

$$S_i^\infty = \left\{(s_1^1, s_i^2, ...) \in \prod_{l \geqslant 1} S_i^l; \forall l: \Psi_i^l(s_i^{l+1}) = s_i^l\right\}.$$

REFERENCES

1. W. ARMBRUSTER AND W. BÖGE, Bayesian game theory, *in* "Game Theory and Related Topics" (O. Moeschlin and D. Pallaschke, Eds.), North-Holland, Amsterdam, 1979.
2. R. J. AUMANN, Subjectivity and correlation in randomized strategies, *J. Math. Econ.* **1** (1974), 67–96.
3. R. J. AUMANN, Agreeing to disagree, *Ann. Statist.* **4** (1976), 1236–1239.
4. R. J. AUMANN, Correlated equilibrium as an expression of bayesian rationality, *Econometrica* **55** (1987), 1–18.
5. B. D. BERNHEIM, Rationalizable strategic behavior, *Econometrica* **52** (1984), 1007–1028.
6. B. D. BERNHEIM, Axiomatic characterization of rational choice in strategic environments, *Scan. J. Econ.* **88** (1986), 473–488.
7. P. BILLINGSLEY, "Convergence of Probability Measures," Wiley, New York, 1968.
8. K. BINMORE, "Modelling Rational Players," Caress Working Paper No. 85–36, University of Pennsylvania, 1985.
9. W. BÖGE AND T. H. EISELE, On solutions of bayesian games, *Int. J. Game Theory* **8** (1979), 193–215.
10. A. BRANDENBURGER AND E. DEKEL, Rationalizability and correlated equilibria (mimeo, 1985), *Econometrica*, in press.
11. I. K. CHO AND D. KREPS, Signalling games and stable equilibria (mimeo, 1985), *Quart. J. Econ.*, in press.
12. J. C. HARSANYI, Games with incomplete information played by "bayesian" players, Part I–III, *Manag. Sci.* **14** (1967–1968), 159–182, 320–334, 486–502.
13. W. HILDENBRAND, "Core and Equilibria of a Large Economy," Princeton Univ. Press, Princeton, NJ, 1984.
14. D. LUCE AND H. RAIFFA, "Games and Decisions," Wiley, New York, 1957.

15. J. F. MERTENS AND S. ZAMIR, Formalization of Harsanyi's notion of "type" and "consistency" in games with incomplete information, *Int. J. Game Theory* **14** (1985), No. 1, 1–29.
16. H. MOULIN, Dominance solvable voting schemes, *Econometrica* **47** (1979), 1337–1351.
17. R. MYERSON, Bayesian equilibrium and incentive compatibility: An introduction, *in* "Social Goals and Social Organization (L. Hurwicz, D. Schmeidler, and H. Sonnenschein, Eds.), pp. 229–259, Cambridge Univ. Press, Cambridge, 1985.
18. R. MYERSON, private communication, 1984.
19. D. PEARCE, Rationalizable strategic behavior and the problem of perfection, *Econometrica* **52** (1984), 1029–1050.
20. P. RENY, "Rationality, Common Knowledge and the Theory of Games," Discussion Paper, Princeton University, 1985.
21. L. J. SAVAGE, "The Foundations of Statistics," Wiley, New York, 1954.
22. T. TAN AND S. R. C. WERLANG, "On Aumann's Notion of "Common Knowledge—An Alternative Approach," Working Paper Series in Economics and Econometrics, No. 85-26, University of Chicago, 1985.
23. T. TAN AND S. R. C. WERLANG, "The Bayesian Foundations of Solution Concepts of Games," Working Paper Series in Economics and Econometrics, No. 86-34, University of Chicago, 1986.
24. S. R. C. WERLANG, "Common Knowledge and Game Theory," Ph. D. thesis, Princeton University, 1986.