MANAGEMENT SCIENCE
Vol. 14, No. 7, March, 1968
Printed in U.S.A.

GAMES WITH INCOMPLETE INFORMATION PLAYED BY 'BAYESIAN' PLAYERS, PART III. THE BASIC PROBABILITY DISTRIBUTION OF THE GAME*†[1]

JOHN C. HARSANYI

University of California, Berkeley

Parts I and II of this paper have described a new theory for the analysis of games with incomplete information. Two cases have been distinguished: consistent games in which there exists some basic probability distribution from which the players' subjective probability distributions can be derived as conditional probability distributions; and inconsistent games in which no such basic probability distribution exists. Part III will now show that in consistent games, where a basic probability distribution exists, it is essentially unique.

It will also be argued that, in the absence of special reasons to the contrary, one should try to analyze any given game situation with incomplete information in terms of a consistent-game model. However, it will be shown that our theory can be extended also to inconsistent games, in case the situation does require the use of an inconsistent-game model.

12.

We now propose to turn back to the questions raised in Section 5 of Part I of this paper. Given any arbitrarily chosen I-game G, is it always possible to construct some Bayesian game G^* Bayes-equivalent to G? And in cases where this is possible, is this Bayesian game G^* always unique? As we have seen in Section 5, these questions are equivalent to asking whether, for any arbitrarily chosen n-tuple of subjective probability distributions $R_1(c^1 \mid c_1), \cdots, R_n(c^n \mid c_n)$, there always *exists* some probability distribution $R^*(c_1, \cdots, c_n)$ satisfying the functional equation(5.3), and whether this distribution R^* is always *unique* in cases in which it does exist. The answers to these two questions are given by Theorem III below. But before stating the Theorem we shall have to introduce a few definitions.

In Section 3 of Part I we formally defined the range space $C_i = \{c_i\}$ of each attribute vector c_i $(i = 1, \cdots, n)$ as being the *whole* Euclidean space E^r of the required number of dimensions. For convenience we shall now relax this definition, and shall allow the range space C_i to be restricted to any proper *subset* C_i^* of the space E^r. (Obviously it makes little difference whether we say that in a given game G the vector c_i ranges only over some subset C_i^* of the Euclidean space E^r, or whether we say that formally the range of this vector c_i is the whole

* Received June 1965 and revised May 1966.

† Parts I and II of "Games with Incomplete Information Played by 'Bayesian' Players" have appeared in the preceding two issues of *Management Science: Theory*. A "Table of Contents" and a "Glossary" appeared with Part I.

[1] The author wishes again to express his indebtedness to the persons and institutions listed in Footnote 1 of Part I of this paper.

space E^r but that outside this set C_i^* the vector c_i has zero probability density everywhere.)

A given subset D of the range space $C = \{c\}$ for the composite vector $c = (c_1, \cdots, c_n)$ is called a *cylinder* if it is a Cartesian product of the form $D = D_1 \times \cdots \times D_n$, where $D_1, \cdots, D_n$ are subsets of the range spaces $C_1 = \{c_1\}, \cdots, C_n = \{c_n\}$ of the vectors $c_1, \cdots, c_n$. These sets $D_1, \cdots, D_n$ are called the *factor sets* of cylinder D. Two cylinders D and D' are said to be *separated* if all their corresponding factor sets D_i and D'_i ($i = 1, \cdots, n$) are disjoint.

Suppose that in a given I-game G (as assessed by player j) each of the range spaces C_i ($i = 1, \cdots, n$) can be partitioned into two or more non-empty subsets $D_i^1, D_i^2, \cdots$, such that each player i can always infer with probability one that every other player's information vector c_k has taken a value $c_k = \gamma_k$ from the set D_k^m ($m = 1, 2, \cdots$) whenever his own information vector c_i has taken a value $c_i = \gamma_i$ from set D_i^m. In symbols this means that

$$R_i(c_k \in D_k^m \mid c_i) = 1 \quad \text{for all} \quad c_i \in D_i^m, \tag{12.1}$$

and for all pairs of players i and k. In this case we shall say that game G can be *decomposed* into the component games $G^1, G^2, \cdots$, in which G^m ($m = 1, 2, \cdots$) is defined as the game resulting from G if the range space C_i of each information vector c_i ($i = 1, \cdots, n$) is restricted to the set $C_i^m = D_i^m$. Accordingly, the range space C of the composite vector $c = (c_1, \cdots, c_n)$ in each component game G^m is restricted to the cylinder $C^m = D^m = D_1^m \times \cdots \times D_n^m$. We shall call D^m the *defining cylinder* of game G^m and shall write $G^m = G(D^m)$.

If G can be decomposed into two or more component games $G^1, G^2, \cdots$ then it is said to be *decomposable*, while in the opposite case it is said to be *indecomposable*.

This terminology is motivated by the fact that whenever a given decomposable game G is being played, the players in actual fact will always play one of its component games $G^1, G^2, \cdots$. Moreover, they will always know, *before* making any move in the game, and before choosing their strategies s_i, *which* of these component games has to be played on that particular occasion, since each player can always observe whether his own information vector c_i lies in set D_i^1, or D_i^2, etc.

We shall use the same terminology also in the case of C-games G^* given in standard form, except that condition (12.1) will then have to be replaced by its analogue:

$$R^*(c_k \in D_k^m \mid c_i) = 1 \quad \text{for all} \quad c_i \in D_i^m. \tag{12.2}$$

The concept of decomposition suggests the following postulate.

Postulate 3. Decomposition of games. Suppose that a given I-game G (as assessed by player j)—or a given C-game G^*—can be decomposed into two or more component games $G^1 = G(D^1), G^2 = G(D^2), \cdots$. Let $\sigma^1, \sigma^2, \cdots$ be the solutions (in terms of any particular solution concept) of $G^1, G^2, \cdots$ if they are regarded as independent games. Then the solution σ of the composite game G—or G^*—

itself will be equivalent to using solution σ^1 whenever the players' information vectors c_i take values from the sets D_i^1, to using solution σ^2 whenever these vectors c_i take values from the sets D_i^2, etc.

In other words, playing game G—or G^*—in practice always means playing one of its component games G^m. Hence in each case the players will use strategies appropriate for that particular component game G^m, and their strategy choices will be unaffected by the fact that, had their information vectors c_i taken different values, they might now have to play a different component game $G^{m'} \neq G^m$. (To repeat, the justification of this postulate is wholly dependent on the fact that, under our definition of games G and G^*, at the time of their strategy choices the players will always be assumed to know which particular component game G^m they are actually playing. If the players could choose their strategies *before* they were informed about the values of their own information vectors c_i—and so before they knew which component game G^m they would be playing—then Postulate 3 would be inappropriate.)

Let $R_1 = R_1(c^1 \mid c_1), \cdots, R_n = R_n(c^n \mid c_n)$ be the subjective probability distributions that player j uses in analyzing a given I-game G. We shall say that these distributions $R_1, \cdots, R_n$ are *mutually consistent* if there exists some probability distribution $R^* = R^*(c)$ [not necessarily unique] satisfying equation (5.3) with respect to $R_1, \cdots, R_n$. In this case, we shall say that game G (as assessed by player j) is itself *consistent*. Moreover, each probability distribution R^* satisfying (5.3) will be called *admissible*.

On the basis of these definitions we can now state the following theorem.

Theorem III. Let G be an I-game (as assessed by player j) in which the player's subjective probability distributions are discrete or absolutely continuous or mixed (but have no singular components). Then three cases are possible:

1. Game G may be inconsistent, in which case no admissible probability distribution R^* will exist.

2. Game G may be consistent and indecomposable, in which case exactly one admissible probability distribution R^* will exist.

3. Game G may be consistent and decomposable. In this case the following statements hold.

(a) Game G can be decomposed into a finite or infinite number of component games $G^1 = G(D^1)$, $G^2 = G(D^2)$, $\cdots$, themselves indecomposable.

(b) There will be an infinite set of different admissible probability distributions R^*.

(c) But for each indecomposable component game $G^m = G(D^m)$, every admissible distribution R^* will generate the same uniquely determined conditional probability distribution

$$R^m = R^m(c) = R^*(c \mid c \in D^m), \tag{12.3}$$

whenever this conditional probability distribution R^m generated by this admissible distribution R^* is well defined.

(d) Every admissible probability distribution R^* will be a convex combination of the conditional probability distributions $R^1, R^2, \cdots$, and different admissible distributions R^* will differ only in the probability weights $R^*(D^m) = R^*(c \in D^m)$ assigned to each component R^m. Conversely, every convex combination of the distributions $R^1, R^2, \cdots$ will be an admissible distribution R^*.

13.

We shall prove the theorem only for games where the distributions $R_1, \cdots, R_n$ are discrete (so that all admissible distributions R^*, if there are any, will also be discrete).[2] For games with absolutely continuous probability distributions the proof is essentially the same, except that the probability mass functions used in the proof have to be replaced by probability density functions. Once the theorem has been proven for both discrete and continuous distributions, the proof for mixed distributions is quite straightforward. Before proving the theorem itself, we shall prove a few lemmas.

Let G be an I-game (as assessed by player j) where all n players i have discrete subjective probability distributions $R_i = R_i(c^i \mid c_i)$. For each probability distribution R_i we can define the corresponding probability mass function r_i as

$$r_i = r_i(\gamma^i \mid \gamma_i) = R_i(c^i = \gamma^i \mid c_i = \gamma_i) \qquad i = 1, \cdots, n, \tag{13.1}$$

where $\gamma^i = (\gamma_1, \cdots, \gamma_{i-1}, \gamma_{i+1}, \cdots, \gamma_n)$ and γ_i stand for specific values of the vectors c^i and c_i, respectively. Likewise, for each admissible probability distribution R^* (if there is one) we can define the corresponding probability mass function r^* as

$$r^* = r^*(\gamma) = R^*(c = \gamma) \tag{13.2}$$

where $\gamma = (\gamma_1, \cdots, \gamma_n)$ stands for specific values of vector c. If R^* is admissible then the corresponding probability mass function r^* will also be called admissible. The set of all admissible functions r^* will be called $\mathcal{R}$.

Equation (5.3) now can be written in the form

$$r^*(\gamma) = r_i(\gamma^i \mid \gamma_i) \cdot r^*(\gamma_i) \qquad \text{for all } r^* \in \mathcal{R} \tag{13.3}$$

and for all values $c = \gamma$, $c^i = \gamma^i$, and $c_i = \gamma_i$. Here $r^*(\gamma_i)$ denotes the marginal probability

$$r^*(\gamma_i) = \textstyle\sum_{(\gamma^i \varepsilon C^i)} r^*(\gamma_i, \gamma^i). \tag{13.4}$$

We shall call any point $c = \gamma = (\gamma_i, \gamma^i)$ of the range space C a *null point* if one or more of the expressions $r_1(\gamma^1 \mid \gamma_1), \cdots, r_n(\gamma^n \mid \gamma_n)$ are zero. We shall call a point $c = \gamma$ a *nonnull point* if all these expressions are positive.

In a consistent game G a given value $c_i = \gamma_i$ of vector c_i will be called an *impossible value* if every admissible function r^* assigns to this value $c_i = \gamma_i$ the marginal probability $r^*(\gamma_i) = 0$. We shall adopt the convention that in any consistent game G each range space $C_i = \{c_i\}$, $i = 1, \cdots, n$, will be defined with

[2] Section 13 contains only the proof of Theorem III and may be omitted without loss of continuity.

the exclusion of all impossible values $c_i = \gamma_i$ of vector c_i, from C_i. Therefore, we shall assume that

$$\text{(13.5)} \qquad \gamma_i \in C_i \quad \text{implies} \quad r^*(\gamma_i) > 0 \quad \text{for } some \quad r^* \in \mathcal{R}.$$

Consequently at a null point $c = \gamma$ *all n* expressions $r_1(\gamma^1 \mid \gamma_1), \cdots, r_n(\gamma^n \mid \gamma_n)$ will be zero because, in view of (13.3) and (13.5), $r_i(\gamma^i \mid \gamma_i) = 0$ implies $r_k(\gamma^k \mid \gamma_k) = 0$ for $k = 1, \cdots, n$.

Lemma 1. Suppose that in a consistent game G at a given point $c = \gamma$ all admissible functions r^* identically take the value

$$\text{(13.6)} \qquad r^*(\gamma) = 0.$$

This will be the case if and only if $c = \gamma$ is a null point.

The lemma again follows from (13.3) and (13.5).

Let $c = \alpha = (\alpha_1, \cdots, \alpha_n)$, $c = \beta = (\beta_1, \cdots, \beta_n)$, $c = \gamma = (\gamma_1, \cdots, \gamma_n)$, and $c = \delta = (\delta_1, \cdots, \delta_n)$ be four nonnull points, such that $\alpha_i = \beta_i$ and $\gamma_i = \delta_i$, whereas $\alpha_k = \gamma_k$ and $\beta_k = \delta_k$. Let r^* be an admissible function taking a positive value at the point $c = \alpha$. Then, by equation (13.3), the quantities $r^*(\alpha_i) = r^*(\beta_i)$ and $r^*(\alpha_k) = r^*(\gamma_k)$ will also be positive, and, therefore, the same will be true of the quantities $r^*(\beta)$ and $r^*(\gamma)$. This in turn implies that the quantities $r^*(\beta_k) = r^*(\delta_k)$ and $r^*(\gamma_i) = r^*(\delta_i)$, as well as the quantity $r^*(\delta)$, will also be positive. Consequently in view of equation (13.3) we have:

$$\text{(13.7)} \qquad \frac{r^*(\alpha)}{r^*(\delta)} = \frac{r^*(\alpha)}{r^*(\beta)} \cdot \frac{r^*(\beta)}{r^*(\delta)} = \frac{r_i(\alpha^i \mid \alpha_i)}{r_i(\beta^i \mid \beta_i)} \cdot \frac{r_k(\beta^k \mid \beta_k)}{r_k(\delta^k \mid \delta_k)},$$

but also

$$\text{(13.8)} \qquad \frac{r^*(\alpha)}{r^*(\delta)} = \frac{r^*(\gamma)}{r^*(\delta)} \cdot \frac{r^*(\alpha)}{r^*(\gamma)} = \frac{r_i(\gamma^i \mid \gamma_i)}{r_i(\delta^i \mid \delta_i)} \cdot \frac{r_k(\alpha^k \mid \alpha_k)}{r_k(\gamma^k \mid \gamma_k)}.$$

If the probability mass functions r_i and r_k—or, equivalently, the probability distributions R_i and R_k—are chosen arbitrarily, then equations (13.7) and (13.8) may become inconsistent, so that no admissible probability mass function r^* taking a positive value $r^*(\gamma) > 0$ at the nonnull point $c = \gamma$ will exist. But then, by Lemma 1, the game will be inconsistent. Thus, we can state:

Lemma 2. If the subjective probability distributions $R_1, \cdots, R_n$ of a given game G are chosen arbitrarily, then they may turn out to be mutually inconsistent, so that no admissible probability distribution R^* will exist.

Now suppose again that G is a consistent game with discrete probability distributions. Any pair of nonnull points $c = \gamma$ and $c = \delta$ will be called *similar* if all admissible probability mass functions r^* yield the same unique ratio $r^*(\gamma)/r^*(\delta)$ whenever this ratio is well defined. (That is, this must be true for all admissible functions r^* not taking zero values $r^*(\gamma) = r^*(\delta) = 0$ at *both* of these two points.) Clearly the relation of similarity partitions the set C^* of all nonnull points into equivalence classes, which we shall call *similarity classes.* All points $c = \gamma$ belonging to the same similarity class E will be similar, and all points $c = \gamma$ and $c = \delta$ belonging to two different similarity classes E and E' will be dissimilar.

Lemma 3. Let $c = \gamma$ and $c = \delta$ be two nonnull points of a consistent game G, agreeing in their ith component $c_i = \gamma_i = \delta_i$. Then γ and δ will belong to the same similarity class E.

Proof. By Lemma 1, we can find some admissible function $r^* = r^0$ such that

$$(13.9) \qquad r^0(\gamma) > 0.$$

In view of (13.3) this implies that

$$(13.10) \qquad r^0(\gamma_i) = r^0(\delta_i) > 0.$$

As both γ and δ are nonnull points we also have

$$(13.11) \qquad r_i(\gamma^i \mid \gamma_i) > 0 \quad \text{and} \quad r_i(\delta^i \mid \delta_i) > 0.$$

Now let $r^* = r^{00}$ be any other admissible function for which the ratio $r^{00}(\delta)/r^{00}(\gamma)$ is well defined. In view of (13.3) and (13.10), we can write

$$(13.12) \qquad r^0(\delta)/r^0(\gamma) = r_i(\delta^i \mid \delta_i)/r_i(\gamma^i \mid \gamma_i) = r^{00}(\delta)/r^{00}(\gamma).$$

By (13.9) and (13.11), the first two ratios in the equation will also be well defined. As this equation will hold for *any* admissible function $r^* = r^{00}$ for which the last ratio is well defined, we can infer that the points δ and γ will belong to the same similarity class E, as desired.

For a given set $E \subseteq C$, let D_i be the set of all vectors $c_i = \gamma_i$ occurring as the i^{th} component of some vector(s) $c = \gamma = (\gamma_1, \cdots, \gamma_i, \cdots, \gamma_n)$ in set E. Then D_i is called the *projection* of set E on vector space C_i. Moreover, if $D_1, \cdots, D_n$ are the projections of set E on the vector spaces $C_1, \cdots, C_n$ then the cylinder $D = D_1 \times \cdots \times D_n$ is called the cylinder *spanned* by set E. Obviously, $E \subseteq D$.

Lemma 4. Let $D^1, D^2, \cdots$ be the cylinders spanned by the similarity classes $E^1, E^2, \cdots$ of a given consistent game G. Then these cylinders $D^1, D^2, \cdots$ will be mutually separated.

Proof. All we have to prove is that for any given value of i $(i = 1, \cdots, n)$ the projections D_i^m and $D_i^{m'}$ of two different similarity classes E^m and $E^{m'}$ are always disjoint. Now suppose this would not be the case. Then D_i^m and $D_i^{m'}$ would have some common point $c_i = \gamma_i$. This in turn would imply that γ_i would be the i^{th} component of some point $c = \gamma = (\gamma_i, \gamma^i)$ in set E^m and also of some point $c = \delta = (\delta_i, \delta^i) = (\gamma_i, \delta^i)$ in set $E^{m'}$. But then, by Lemma 3, γ and δ would belong to the same similarity class, and could not belong to two different similarity classes E^m and $E^{m'}$, which is inconsistent with the assumptions made. Hence, D^m and $D^{m'}$ cannot have any common point $c_i = \gamma_i$.

Lemma 5. If the nonnull points of a given consistent game G can be partitioned into two or more similarity classes $E^1, E^2, \cdots$, then game G itself can be decomposed into corresponding component games $G^1 = G(D^1)$, $G^2 = G(D^2), \cdots$, where the defining cylinder D^m $(m = 1, 2, \cdots)$ of component game $G^m = G(D^m)$ is the cylinder spanned by the similarity class E^m.

Proof. Let $c_i = \alpha_i$ and $c_k = \beta_k$ be specific values of vectors c_i and c_k, such that $\alpha_i \in D_i^m$ but $\beta_k \notin D_k^m$. Let $c = \gamma = (\gamma_1, \cdots, \alpha_i, \cdots, \beta_k, \cdots, \gamma_n)$ be any point having $\gamma_i = \alpha_i$ as its i^{th} component and having $\gamma_k = \beta_k$ as its k^{th} component.

Then, by Lemma 4, γ will be a null point so that

$$r_i(\gamma^i \mid \alpha_i) = r_i(\beta_k, \gamma^{ik} \mid \alpha_i) = 0, \tag{13.13}$$

where γ^{ik} is the vector which remains if the i^{th} and k^{th} components are both omitted from $\gamma = (\gamma_1, \cdots, \gamma_n)$. Therefore,

$$r_i(\beta_k \mid \alpha_i) = R_i(c_k = \beta_k \mid c_i = \alpha_i) = 0 \tag{13.14}$$

whenever $\alpha_i \in D_i^m$ but $\beta_k \notin D_k^m$. Consequently,

$$R_i(c_k \notin D_k^m \mid c_i) = 0 \qquad \text{if } c_i \in D_i^m \tag{13.15}$$

and so

$$R_i(c_k \in D_k^m \mid c_i) = 1 \quad \text{if} \quad c_i \in D_i^m. \tag{13.16}$$

Hence, for each cylinder D^m $(m = 1, 2, \cdots)$, every pair of factor sets D_i^m and D_k^m $(i, k = 1, \cdots, n)$ will satisfy condition (12.1), and so game G can be decomposed into the component games $G^m = G(D^m)$, $m = 1, 2, \cdots$.

Lemma 6. A given consistent game G is decomposable if and only if it has more than one admissible probability distribution R^*.

Proof of the "if" part of the lemma. If all nonnull points of game G belong to the same similarity class E, then, for all pairs of nonnull points γ and δ, the ratios $r^*(\gamma)/r^*(\delta)$ are uniquely determined. Hence, the game can have only one admissible probability mass function r^* and only one admissible probability distribution R^*.

Thus, if game G does have more than one admissible probability distribution R^*, then it must contain two or more similarity classes E^1, E^2, $\cdots$; and, therefore, by Lemma 5 it will be decomposable.

Proof of the "only if" part of the lemma. We have to show that G will have more than one admissible probability distribution R^* (or more than one admissible probability mass function r^*) if it is a decomposable game.

Suppose to the contrary that a given game G can be decomposed into the component games $G^1 = G(D^1)$, $G^2 = G(D^2)$, $\cdots$ but has only one admissible probability mass function $r^* = r^0$. Then, by Lemma 1, this function r^0 must take a positive value $r^0(\gamma) > 0$ at each nonnull point $c = \gamma$ of G. Let r^m $(m = 1, 2, \cdots)$ be the conditional probability mass function

$$\begin{aligned} r^m(\gamma) &= r^0(\gamma \mid \gamma \in D^m) \\ &= r^0(\gamma) / \textstyle\sum_{(c \in D^m)} r^0(c) && \text{for all } \gamma \in D^m, \\ r^m(\gamma) &= 0 && \text{for all } \gamma \notin D^m. \end{aligned} \tag{13.17}$$

Since $r^0(c) > 0$ at all nonnull points c, this function r^m will be everywhere well defined. In view of (13.3) and (13.17) we can write

$$r^m(\gamma) = r_i(\gamma^i \mid \gamma_i) \cdot r^m(\gamma_i), \tag{13.18}$$

which means that the function $r^* = r^m$ $(m = 1, 2, \cdots)$ satisfies condition (13.3). Hence, the functions r^1, r^2, $\cdots$ are just as much admissible probability mass functions as r^0 itself is, which contradicts the assumption that $r^* = r^0$ is the only admissible probability mass function of the game. This completes the proof of Lemma 6.

Lemma 7. A consistent game G is indecomposable if and only if all its nonnull points belong to the same similarity class E.

Proof of the "if" part. If all nonnull points are similar than the ratios $r^*(\gamma)/r^*(\delta)$ are uniquely determined for all nonnull points γ and δ. Hence, the game can have only one admissible probability mass function r^*, and so, by Lemma 6, it is indecomposable.

Proof of the "only if" part. If game G is indecomposable, then, by Lemma 6, it will have only one admissible function r^*. Hence, the ratios $r^*(\gamma)/r^*(\delta)$ will be uniquely determined for all pairs of nonnull points γ and δ, and so all of these points will be similar.

Lemma 8. Let G be a decomposable consistent game, and let $G^1 = G(D^1)$, $G^2 = G(D^2)$, $\cdots$ be its component games as defined by Lemma 5. Then

1. Each component game G^m will be itself an indecomposable game.
2. If $r^* = r^0$ and $r^* = r^{00}$ are two admissible probability mass functions for G then, for each component game G^m ($m = 1, 2, \cdots$), both $r^* = r^0$ and $r^* = r^{00}$ will generate the same conditional probability mass function

$$r^m(\gamma) = r^0(\gamma \mid \gamma \in D^m) = r^{00}(\gamma \mid \gamma \in D^m) \tag{13.19}$$

whenever these conditional probability mass functions generated by r^0 and r^{00} are well defined—that is, whenever the two functions r^0 and r^{00} assign a positive probability mass to the event $E = \{\gamma \in D^m\}$.

3. Every admissible probability mass function r^* for the game will be a convex combination of these functions $r^1, r^2, \cdots$ corresponding to the component games $G^1, G^2, \cdots$; and any convex combination of these functions $r^1, r^2, \cdots$ will be an admissible probability mass function r^*.

Proof. Parts 1 and 2 follow from Lemma 7 in view of the fact that all nonnull points γ lying in any given cylinder D^m belong to the same similarity class E^m. Part 3 follows from equations (13.17) and (13.3).

Lemmas 2, 5, 6, 7, and 8 together imply Theorem I as restricted to games with discrete probability distributions. We have already indicated (in the first paragraph of this section) how the proof can be extended to the general case.

14.

By Theorem III, if a given player j participating in some I-game G does not take special care to base his analysis of the game on mutually consistent subjective probability distributions $R_1, \cdots, R_j, \cdots, R_n$, then, in general, no admissible basic probability distribution R^* will exist; and, therefore, there will be no Bayesian game G^* Bayes-equivalent to G. In view of this fact we now propose the following postulate.

Postulate 4. Mutual consistency. Unless he has special reasons to believe that the subjective probability distributions $R_1, \cdots, R_j, \cdots, R_n$ used by the n players are, in fact, mutually inconsistent, every player j of an I-game G will always analyze the game in terms of a set of some mutually consistent probability distributions $R_1, \cdots, R_j, \cdots, R_n$.

Postulate 4 can be supported by the following considerations:

(i) Player j can greatly simplify his analysis of the game situation by using n mutually consistent distributions $R_1, \cdots, R_n$, because this will enable him to reduce the problem of analyzing the I-game G to the much easier problem of analyzing an equivalent C-game G^*. Therefore, player j will be well advised to assume n mutually consistent distributions $R_1, \cdots, R_n$, unless he feels that the information he has about the game situation is incompatible with this assumption.

(ii) Admittedly, if player j does not restrict his choice to mutually consistent distributions R_i, then he will have a freer hand in postulating suitable differences between the different players' subjective probability distributions R_i, and so will have a freer hand in providing mathematical representation for any differences he may wish to assume between different players i as to their beliefs and expectations about the game situation. However, even if he does restrict his choice to mutually consistent R_i's, this need not prevent him from finding suitable representation for such inter-player differences in his mathematical model of the game; for such differences can often be given fully adequate representation by assuming corresponding differences between the relevant players' attribute vectors c_i, and by choosing a basic probability distribution $R^* = R^*(c_1, \cdots, c_n)$ which is a sufficiently asymmetric function of the different players' attribute vectors c_i; so that the conditional probability distributions $R_i = R^*(c^i \mid c_i)$ for different players i will have appropriately different mathematical forms. Mutually inconsistent R_i's will have to be assumed only in cases in which player j comes to the definite conclusion that the inter-player differences in question could not be represented by mutually consistent R_i's.

(iii) According to the prior-lottery model (see Section 6 of Part I of this paper), every I-game G can be considered to be a result of a random social process which has selected the n actual players of the game from some hypothetical populations $\Pi_1, \cdots, \Pi_n$ of possible players, where the probability that this random process will select n players with any given combination of specific attribute vectors $c_1 = c_1^0, \cdots, c_n = c_n^0$ is governed by some (objective) probability distribution $R^* = R^*(c_1, \cdots, c_n)$. However, in general, the true distribution R^* governing this random process will be unknown to the players.

Let $R_i^* = R^*(c^i \mid c_i)$, for $i = 1, \cdots, n$, be the conditional probability distributions generated by this unknown distribution R^*. Clearly the n distributions $R_1^*, \cdots, R_n^*$ will always be mutually consistent, since by definition they are conditional distributions generated by the same basic distribution R^*. The n players' subjective probability distributions $R_1, \cdots, R_n$ can be regarded as their *estimates* of these mutually consistent, but unknown, conditional probability distributions $R_1^*, \cdots, R_n^*$.

Accordingly, it is natural to argue (at least in cases where player j has no special reasons for using inconsistent distributions R_i) that, instead of trying to estimate each player's subjective probability distribution R_i separately, player j

should rather try to estimate directly the objective probability distribution R^* governing the random social process in question; and then he should choose the n distributions R_i he will use in the analysis of the game situation, by setting $R_i = R^*(c^i \mid c_i)$, where R^* now refers to his *estimate* of the true distribution R^*. Obviously, if player j uses this estimation procedure, he will always come up with n mutually consistent distributions $R_1, \cdots, R_n$.

We have seen in Section 6 (Part I) that the prior-lottery model pictures the game situation in the way in which it would be seen by an intelligent and "properly informed" outside observer.[3] Consequently, when player j is trying to estimate the probability distribution R^*, he will have to ask himself the question, "What probability distribution R^* would be chosen by a randomly selected intelligent and properly informed outside observer, as his estimate of the true probability distribution governing the random social process underlying this game situation G?" (If player j feels that he cannot give a unique answer to this question because different intelligent and properly informed observers are likely to choose different probability distributions R^*, then he can construct his own probability distribution R^* by some averaging process over the various distributions R^* he thinks different observers might choose.)

To put it differently, in trying to estimate the probability distribution R^*, player j should try to use only the information *common* to all n players. Of course, in his analysis at some point or another he should make use of *all* information he has about the game situation, including any special information which may be available to him but may not be available to the other players. But under our model all such special information must be represented by incorporating it into player j's own information vector c_j. In contrast, the basic probability distribution R^* itself is meant to represent only information common to all n players. (The same is true for the n subjective probability distributions $R_1, \cdots, R_n$ and the n payoff functions $V_1, \cdots, V_n$: by definition they are meant to contain only such information that, in player j's opinion, is common property of the n players —see Section 3 in Part I.)

To be sure, if player j follows the suggested estimation procedure this will guarantee only that his own subjective probability distribution R_j and the subjective probability distributions $R_1, \cdots, R_{j-1}, R_{j+1}, \cdots, R_n$ he will *ascribe* to the other $(n - 1)$ players will satisfy equation (5.3). But, of course, it will not guarantee that his own subjective probability distribution R_j and the subjective probability distributions $R_1, \cdots, R_{j-1}, R_{j+1}, \cdots, R_n$ the other $(n - 1)$ players will in *actual fact* use in the game will satisfy equation (5.3). In other words, the suggested estimation procedure will guarantee *internal consistency* among the n probability distributions $R_1, \cdots, R_n$ that player j himself will use in analyzing the game, but will not and cannot guarantee *external consistency* among the probability distributions used by the different players. Indeed, so long as each player has to choose his subjective probabilities (probability estimates)

[3] We have defined a "properly informed" observer as a person possessing all information *shared* by all n players but having no extra information unavailable to the players, or available to some particular player(s) yet unavailable to the other player(s).

independently of the other players, no conceivable estimation procedure can ensure consistency among the different players' subjective probabilities.

More particularly, there is no way of ensuring that the different players will choose the *same* subjective probability distribution as their estimate of the objective probability distribution R^* underlying the game situation. For, by the very nature of subjective probabilities, even if two individuals have exactly the same information and are at exactly the same high level of intelligence, they may very well assign different subjective probabilities to the very same events.

At the same time, while the suggested estimation procedure does not in any way guarantee external consistency among the different players' probability estimates, we feel it does go as far as an estimation procedure can go in promoting such external consistency. By asking each player to take the point of view of an outside observer in estimating the objective probability distribution R^*, it asks him to choose an estimate as close as possible to the estimates other intelligent people might be expected to choose, and to make his estimate as independent as possible of his own personal prejudices and idiosyncrasies.

15.

So far our analysis has been restricted to the case in which player j has no special information suggesting mutual inconsistencies among the probability distributions $R_1, \cdots, R_j, \cdots, R_n$ of the players. Now what happens if player j does have such information? For instance, what happens if in an international conflict situation country j obtains quite convincing reports from its diplomatic representatives, intelligence agencies, newspaper correspondents, etc., to the effect that in analyzing the international situation country i is using a very different political and social philosophy, and a very different model of the world, from country j's own; and that as a result the subjective probability distribution R_i used by country i is clearly inconsistent with the subjective probability distribution R_j used by country j?

In a situation such as this, one obvious possibility is to take this information at face value and to conclude that the players' subjective probability distributions *are* in fact inconsistent, which means that every player j simply has to resign himself to the necessity of analyzing the game situation in terms of n mutually inconsistent subjective probability distributions $R_1, \cdots, R_j, \cdots, R_n$.

As Reinhard Selten has pointed out,[4] his posterior-lottery model (see Section 6, Part I of this paper) can be extended to this case of mutually inconsistent subjective probability distributions. All we have to do is to assume that when all K players have chosen their strategies there will be a *separate* lottery $L(i_m)$ for every player i_m, instead of there being merely *one* grand lottery L^* for all K players. For each player i_m in role class i his lottery $L(i_m)$ will choose $(n - 1)$ players [one player from each of the $(n - 1)$ role classes $1, \cdots, i - 1, i + 1, \cdots, n$] as his "partners" in the game. If player i_m himself has the attribute vector $c_i = c_i^0$, then the probability of his $(n - 1)$ partners' having any specific

[4] In private communication (see Footnote 1 in Part I).

combination of attribute vectors $c_1 = c_1^0; \cdots ; c_{i-1} = c_{i-1}^0 ; c_{i+1} = c_{i+1}^0 ; \cdots ; c_n = c_n^0$ will be governed by the subjective probability distribution $R_i = R_i(c^i \mid c_i = c_i^0)$ which player i himself entertains. Player i_m will receive the payoff:

$$x_i = V_i(s_1^0, \cdots, s_i^0, \cdots, s_n^0; c_1^0, \cdots, c_i^0, \cdots, c_n^0) \tag{15.1}$$

where $s_1^0, \cdots, s_n^0$ are the strategies chosen by player i_m and by his partners, whereas $c_1^0, \cdots, c_n^0$ are the attribute vectors of these same n players.

Note that, under this model, "partnership" is not necessarily a symmetric relationship. If lottery $L(i_m)$ of player i_m chooses a given player k_r as a partner for player i_m, then it does not follow that lottery $L(k_r)$ of player k_r will likewise choose player i_m as a partner[5] for player k_r. As partnership in general will not be a symmetric relationship, and as the lotteries associated with different players will be quite independent, this model does not presuppose any mutual consistency in the sense of Postulate 4 among the n probability distributions $R_1 = R_1(c^1 \mid c_1), \cdots, R_n = R_n(c^n \mid c_n)$ governing the lotteries conducted for the players in the n role classes $1, \cdots, n$.

Thus, a Selten game G^{**} will exist even for an inconsistent I-game G, for which no Bayesian game G^* exists.

16.

It is, however, always questionable whether any given information suggesting inconsistencies among the different players' subjective probability distributions $R_1, \cdots, R_n$ should really be taken at face value.

First of all, information about other players' assessment of probabilities (and indeed about their internal beliefs and attitudes in general) tends to be very unreliable, because the players will often have a vested interest in misleading the other players about their real ways of thinking. But even in the empirical facts themselves about the other players' probability judgments are quite correct, they will be usually open to alternative interpretations.

More particularly, any inconsistency among the various players' subjective probability distributions R_i is always a result of discrepancies among the basic probability distributions $R^* = R^{(i)}$ used by different players i (see Theorem IV below). On the other hand, these discrepancies among the probability distributions $R^{(i)}$ will themselves often admit of explanation in terms of the *differences in the information* available to different players i—in which case, as we shall see, the

[5] This fact does not give rise to any logical difficulty. It only means that, owing to the outcome of lottery $L(i_m)$, player i_m's payoff may come to depend on the strategy s_k^0 and the attribute vector c_k^0 of some player k_r—even though, owing to the outcome of lottery $L(k_r)$, player k_r's own payoff will not be dependent on the strategy s_i^0 and the attribute vector c_i^0 of player i_m, but rather will be dependent on the strategy s_i^{00} and the attribute vecto- c_i^{00} of another player $i_t \neq i_m$ in role class i. (Of course, when the players choose their strater gies they will not know who will be whose partner and whether any such relation of partnership will be reciprocal or not.)

game can be *reinterpreted* as a game involving mutually *consistent* subjective probability distributions $R_1, \cdots, R_n$ on the part of the n players.

We now propose to state the following simple theorem.

Theorem IV. Let

$$R^*(c_1, \cdots, c_i, \cdots, c_n) = R^{(i)}(c_1, \cdots, c_i, \cdots, c_n) \tag{16.1}$$

be player i's basic probability distribution ($i = 1, \cdots, n$), defined as his estimate of the probability distribution governing the random social process which gave rise to the current game situation. Suppose that

$$R^{(1)} = \cdots = R^{(n)} = R^0. \tag{16.2}$$

Then the n players' subjective probability distributions $R_1, \cdots, R_n$ will be mutually consistent.

Proof. Suppose player i thinks that the social process in question is governed by the probability distribution $R^* = R^{(i)}$, and knows that his own attribute vector has the value $c_i = c_i^0$. Then he must use the conditional probability distribution $R^*(c^i \mid c_i) = R^{(i)}(c^i \mid c_i)$, generated by this distribution $R^* = R^{(i)}$, for evaluating the probability that the other $(n - 1)$ players will have any specific combination of attribute vectors $c_1 = c_1^0; \cdots; c_{i-1} = c_{i-1}^0; c_{i+1} = c_{i+1}^0; \cdots; c_n = c_n^0$. Hence, player i's subjective probability distribution R_i will be defined by this conditional distribution[6]

$$R_i = R_i(c^i \mid c_i) = R^*(c^i \mid c_i) = R^{(i)}(c^i \mid c_i). \tag{16.3}$$

This will be true for all players $i = 1, \cdots, n$. Consequently, in view of equation (16.2) we can write

$$R_i = R^0(c^i \mid c_i) \qquad i = 1, \cdots, n. \tag{16.4}$$

Hence, all n functions $R_1, \cdots, R_n$ will have the nature of conditional probability distributions generated by the same basic probability distribution R^0, which means that they will be mutually consistent.

By Theorem IV, if the functions $R_1, \cdots, R_n$ are in fact mutually inconsistent, then the basic probability distributions $R^{(1)}, \cdots, R^{(n)}$ used by the n players cannot be all identical; and it will be the discrepancies existing among these basic probability distributions $R^{(1)}, \cdots, R^{(n)}$ that account for the inconsistency among the subjective probability distributions $R_1, \cdots, R_n$. Now, our contention is that the discrepancies existing among the different players' basic probability distributions $R^{(i)}$ will often admit of explanation in terms of differences in the information that different players have about the nature of the social process underlying the current game situation.

For instance, in our previous example, if the subjective probability distributions R_i and R_j used by countries i and j are mutually inconsistent, this will indi-

[6] In the case $k \neq i$, the relationship $R_i = R^{(k)}(c^i \mid c_i)$ will hold in general only if $R^{(k)} = R^{(i)}$ as required by equation (16.2). However, in the case $k = i$, the relationship will always hold, by the very definitions of the subjective probability distributions R_i and of the basic probability distributions $R^* = R^{(i)}$.

cate that they assess probabilities in terms of two quite different basic probability distributions $R^{(i)}$ and $R^{(j)}$. More particularly, *e.g.*, country i may assign much higher probabilities than country j does to the success of violent social revolutions in various parts of the world. Now it will often be a very natural hypothesis that these differences in the two countries' assessment of probabilities are due to differences in the information they have acquired about the world as a result of their divergent historical experiences—more particularly as a result of their divergent experience about the prospects of violent social revolutions on the one hand, and about the chances of achieving social reforms by peaceful nonrevolutionary methods on the other hand.

Of course, if player j makes the assumption that the discrepancies among the probability distributions $R^{(i)}$ used by different players i are due to differences in the information available to them, this will mean under our terminology that these distributions $R^{(i)}$ are not really the *basic* probability distributions of these players, because in this case, obviously, these distributions $R^{(i)}$ cannot be assumed any longer to represent exclusively the information *common* to all n players. Instead, each distribution $R^{(i)}$ will now have to be interpreted as a *conditional* probability distribution, conditioned by some special information available to player i himself but not available to all n players.

On the other hand, if player j is willing to assume that *all* discrepancies among the distributions $R^{(i)}$ used by different players i are due to these differences in the information available to them, then he can consider all these distributions $R^{(i)}$ to be conditional distributions derived from *one* and the *same* unconditional distribution $R^{*\prime}$, which means that this distribution $R^{*\prime}$ will be the true basic distribution of the game.

In symbols, let d_i be a vector summarizing all special information about the relevant variables which makes any given player i adopt the probability distribution $R^{(i)} = R^{(i)}(c_1, \cdots, c_i, \cdots, c_n)$—instead of adopting some probability distribution $R^{(k)} = R^{(k)}(c_1, \cdots, c_i, \cdots, c_n)$ used by another player k. Then player i's true information vector (or attribute vector) will not be the vector c_i but rather will be the larger vector $c'_i = (c_i, d_i)$ summarizing the information contained in *both* vectors c_i and d_i.

Consequently, the true basic probability distribution $R^{*\prime}$ of the game will be the joint probability distribution of these larger vectors c'_i, and so will have the form

$$(16.5) \qquad R^{*\prime} = R^{*\prime}(c'_1, \cdots, c'_n) = R^{*\prime}(c_1, d_1; \cdots; c_n, d_n).$$

In contrast, each distribution $R^* = R^{(i)}$ will have the nature of a conditional probability distribution (or more exactly of a conditional marginal probability distribution) of the form

$$(16.6) \quad R^{(i)} = R^{(i)}(c_1, \cdots, c_i, \cdots, c_n) = R^{*\prime}(c_1, \cdots, c_i, \cdots, c_n \mid d_i).$$

In other words, player j will find that, as soon as he redefines the players' attribute vectors as the larger vectors $c'_i = (c_i, d_i)$, then the information he has about the other players' subjective probabilities will become compatible with the hypothesis that all n players are actually using the *same* basic probability distribution $R^{*\prime} = R^{*\prime}(c'_1, \cdots, c'_j, \cdots, c'_n)$. Consequently, by Theorem IV, this information will be compatible also with the hypothesis that the n players are using *mutually consistent* subjective probability distributions $R'_1, \cdots, R'_j, \cdots, R'_n$ of the form

$$(16.7) \qquad R'_i = R'_i((c^i)' \mid c'_i) = R^{*\prime}((c^i)' \mid c'_i) \qquad i = 1, \cdots, j, \cdots, n,$$

where

$$(c^i)' = (c'_1, \cdots, c'_{i-1}, c'_{i+1}, \cdots, c'_n).$$

To sum up, if the information player j has about the other players' subjective probabilities seems to suggest that the n players are using mutually *inconsistent* subjective probability distributions $R_1, \cdots, R_j, \cdots, R_n$, then more careful reanalysis of the game situation may very well lead him to the conclusion that, in actual fact, the players' subjective probability distributions can be better represented by another set of probability distributions $R'_1, \cdots, R'_j, \cdots, R'_n$, mutually *consistent* with one another.

Note that, in reanalyzing the game situation in this way, player j must use not only the information he has obtained independently about the game, but also any information he can infer from what he knows about the other players' assessment of probabilities. For instance, suppose that player j himself would be inclined to assign the probability $P_j(e)$ to a given event e, but then obtains the information that another player i assigns a very different probability $P_i(e) \neq P_j(e)$ to this event. Then player j cannot rest the matter with the conclusion that he and player i have apparently been assessing the probability of this event on the basis of two rather different sets of information. Instead, he also has to reach a decision on whether his own assessment of this probability is likely to be based on more correct and more complete information than player i's assessment is, or whether the opposite is the case; and, in particular, he has to decide whether player i's quite different assessment of this probability is or is not a good enough reason for him himself also to change his own assessment of this probability from $P_j(e)$ to some number closer to the probability $P_i(e)$ which player i is assigning to event e.

Of course, this method of reconciling the different players' assessments of probabilities in terms of the same basic probability distribution $R^{*\prime}$ will work only if player j feels that the empirical facts which he knows about the other players' probability assessments are favorable to, or at least are compatible with, the hypothesis that the discrepancies among the various players' probability judgments can be reasonably explained in terms of differences in their information. In our own view, in most cases the empirical facts will be at least compatible with this hypothesis. But if in any given I-game G player j reaches the opposite conclusion, then he will have to fall back upon Selten's model (see Section 15), which permits mutually inconsistent subjective probability distributions on the part of the players.

17.

According to Theorem III (Section 12), even if player j uses mutually consistent subjective probability distributions $R_1, \cdots, R_n$ in his analysis of a given I-game G, he may still have the problem of there being an infinite set of equally admissible probability distributions R^* satisfying equation (5.3) [Part I of this paper]; which means that there will be an infinite number of Bayesian games G^* Bayes-equivalent to this I-game G.

However, by Postulate 3, this fact will not give rise to any difficulties because, if player j uses the solution σ of *any* of these games G^* as the solution of game G, he will always obtain the *same* solution σ. This is so because, by Theorem III, even if there are many Bayesian games G^* Bayes-equivalent to G, there will always be only *one* Bayesian game $(G^m)^*$ Bayes-equivalent to any given indecomposable *component game* G^m of G (since G^m will have only one unique admissible probability distribution $R^*(G^m) = R^m$. Hence, no ambiguity will result if he uses the solution σ^m of this Bayesian game $(G^m)^*$ as the solution of this component game G^m. On the other hand, by Postulate 3, the solutions σ^m of these component games G^m will completely determine the solution σ of the whole game G.

Yet, even though the possible multiplicity of Bayesian games G^* Bayes-equivalent to a given I-game G causes no real problem, for some purposes it will be convenient if we can make the assumption that there is always only *one* Bayesian game G^* Bayes-equivalent to each I-game G. We can achieve this by using the fact that, under the procedure proposed in Section 14, each player j will start his analysis of a given I-game G by first choosing a basic probability distribution $R^* = R^*(c_1, \cdots, c_n)$, and then will define the n probability distributions $R_1 = R_1(c^1 \mid c_1), \cdots, R_n = R_n(c^n \mid c_n)$ as the conditional distributions $R_1 = R^*(c^1 \mid c_1); \cdots; R_n = R^*(c^n \mid c_n)$ derived from this distribution R^*. This make it natural to define the I-game G itself (as assessed by player j) in terms of this one distribution R^* chosen by player j at the beginning of his analysis, rather than in terms of the n distributions $R_1, \cdots, R_n$ derived from this distribution R^*.

Formally, this will mean that not only a C-game G^* in standard form (i.e., a Bayesian game G^*), but also any consistent I-game G in standard form, will now be defined as the ordered set

$$G = G^* = \{S_1, \cdots, S_n; C_1, \cdots, C_n; V_1, \cdots, V_n; R^*\}, \tag{17.1}$$

instead of being defined as the ordered set

$$G = \{S_1, \cdots, S_n; C_1, \cdots, C_n; V_1, \cdots, V_n; R_1, \cdots, R_n\} \tag{17.2}$$

in accordance with equation (3.18), which was our original suggestion. (Of course, in the case of an inconsistent I-game G we have to go on using equation (17.2) as our formal definition of the game.)

This now completes our general discussion of incomplete-information games G and of their Bayesian-game analogues G^*, which represent games with complete but imperfect information involving certain chance moves. We have discussed the actual solution concepts our approach selects for these games only in the case of a few simple illustrative examples (see Sections 9–11, Part II). More systematic discussion of these solution concepts for a wider range of games with incomplete information must be left for other forthcoming publications.