THE IMPOSSIBILITY OF BAYESIAN GROUP DECISION MAKING WITH SEPARATE AGGREGATION OF BELIEFS AND VALUES

BY AANUND HYLLAND AND RICHARD ZECKHAUSER[1]

Bayesian theory for rational individual decision making under uncertainty prescribes that the decision maker define independently a set of beliefs (probability assessments for the states of the world) and a system of values (utilities). The decision is then made by maximizing expected utility. We attempt to generalize the model to group decision making. It is assumed that the group's belief depends only on individual beliefs and the group's values only on individual values, that the belief aggregation procedure respects unanimity, and that the entire procedure guarantees Pareto optimality. We prove that only trivial (dictatorial) aggregation procedures for beliefs are possible.

1. INTRODUCTION

MANY DECISIONS MADE under uncertainty, indeed many important ones, are made by a group, be it a collection of friends, the Congress of the United States, or the participants in a corporate bureaucracy. The members of such a group may differ in both beliefs (probability assessments) and values (utilities).

Models for individual decision making under uncertainty are often based on the Bayesian paradigm.[2] Following it, probability assessments and utility functions are defined independently. It seems reasonable to suggest that groups adopt a similar approach.[3] That is, the group should deal separately with the two areas of potential disagreement. By some prescribed procedure, the group aggregates the individual probability assessments into a group probability assessment. Similarly, a group utility function is constructed by aggregating individual utility functions. The group probability assessment and utility function are multiplied in the usual manner to give expected utilities. The action offering the highest expected utility is chosen.[4]

Most real world decision processes, we recognize, make less than conscientious attempts to separate beliefs from values, either in debate or at time of decision.[5]

[1] Hylland's work was supported by a fellowship from the Norwegian Research Council for Science and the Humanities, and Zeckhauser's work by NSF grants SOC77-16602 and SER76-17502. Howard Raiffa, Thomas Schelling, and the referees gave us helpful comments.

[2] It is assumed that the reader is familiar with the basics of Bayesian decision making under uncertainty; in particular, the concepts of subjective probability assessment and (von Neumann–Morgenst[illegible]) utility functions, their use, and the consistency or rationality conditions required to prove their existence. For an introduction to these subjects, see, for example, Howard Raiffa [1].

[3] In this paper, we are not going to address the question of whether the Bayesian theory actually is a reasonable one; that issue is discussed at length by others. We simply observe that the Bayesian model is often used in connection with individual decision making under uncertainty, and ask if it can be generalized to the case of group decision making.

[4] Note the introspective setting; we regard the theory as a prescriptive guide to rational action. A descriptive theorist who observes a person's actions and tries to compute the underlying probability assessments and utilities will have an identification problem.

[5] Undoubtedly, political discussion could often gain if more effort were made to separate facts and values. When the structure of a problem is ill-defined, it is not impossible, however, that such an attempt will systematically favor one point of view and therefore in itself can be used in the political struggle. These interesting issues are not the subject of this paper.

But these may be merely the flaws of practice. Here we investigate the theoretical ideal.

Following the efficiency dictates of welfare economics, the chosen action should also be weakly Pareto optimal. That is, there should be no alternative action that offers higher expected utility for all individuals.

Essentially, we prove in this paper that no group decision procedure can satisfy these requirements. In the context of the analysis, a finite group of individuals shall make a decision. There is a fixed and finite set of possible actions, one of which must be chosen, and a fixed and finite[6] set of potential states of the world, one of which will occur. Each combination of action and state leads to a specific outcome. There may be disagreement about the probabilities of the various states of the world, and the individuals may evaluate the outcomes differently.[7]

If the aggregation procedures for the group probability assessment and the group utility function were simple averages, then it is easy to find examples where the chosen action will be Pareto inferior. That is, it is possible for all members of the group, based on their own probability assessments and utilities, to choose action a_1 rather than a_2, while the group chooses a_2.[8] The question then is: Can this result be avoided by choosing other aggregation procedures?

In a 1968 discussion paper, Richard Zeckhauser [**2**] answered this question in the negative,[9] though he gave only a heuristic proof of his result. In this paper, the result is stated and proved formally. Moreover, the assumptions involved here are somewhat weaker than the ones used in the 1968 paper.

Our results depend on a few more assumptions than the basic Bayesian model and the principle of Pareto optimality. For one thing, there is the requirement that probabilities and utilities be aggregated separately. One could imagine "Bayesian" group decision procedures in which the group probability assessment depends on the individual utility functions as well as on the individual probability assessments, and similarly for the group utility function. But an essential part of the Bayesian theory for individual decision making is that probabilities and utilities can be defined independently; therefore, separate aggregation is a natural requirement of any generalization to group decision making.[10]

[6] Finiteness is not essential for any of the three sets. If the set of states is infinite, probability measures must be used below instead of probability vectors, and some assumptions must be made concerning existence of expectations.

[7] An alternative setting in which the probabilities of the states depend on the action taken can easily be reformulated to fit the framework of the text.

[8] See example in Raiffa [**1**, Section 8.12].

[9] See also Raiffa [**1**, Sections 8.12–8.13], where this result is described.

[10] We do not claim that this provides an irrefutable argument that separate aggregation must be a part of any Bayesian theory of group decision making under uncertainty. Primarily we are concerned not with arguing whether separability is a reasonable condition, but with investigating its consequences. Indeed, our results can be viewed as arguments against the principle of separate aggregation. The separability assumption in effect rules out one type of log-rolling in the decision process, namely the possibility that one person gives in on the question of probabilities in return for getting more influence on the social utility scheme, or vice versa. It is often the case that restrictions on log-rolling opportunities make it difficult to achieve desirable outcomes. The results of this paper are consistent with this general proposition.

Moreover, we assume that the probability aggregation procedure satisfies the following conditions: (i) If every individual has the same probability assessment, the group's assessment is also equal to this. (ii) No individual is dominant in the sense that this person's probability assessment is always taken as the group's assessment. (Similar conditions can be imposed on the utility aggregation procedure, but they are unnecessary for our impossibility results.)

Condition (i) is necessary for our proofs. (We believe, however, that it can be dropped, provided that a couple of more technical assumptions are strengthened; see note 22.) The inclusion of this condition does not, in our opinion, weaken the significance of the results: A person who insists that the Pareto principle be maintained says, in effect, that nothing outside the individuals' preferences can override a preference unanimously held by them. It then seems inconsistent to allow anything outside the individuals' probability assessment to override their unanimously held beliefs. We claim, therefore, that anybody who accepts the Pareto principle should also accept (i).

Condition (ii) is imposed to rule out certain trivial procedures. (If strong Pareto optimality is required, (ii) can be dropped; see Section 6.) Following the convention of social choice theory, we will call an aggregation procedure "dictatorial" if it violates (ii). This term strongly suggests something undesirable, a connotation which need not be appropriate in our connection: There is nothing undemocratic about the existence of an undisputed expert whose assessment of the world is always accepted by the group.[11] Moreover, even if a "probability dictator" exists, in violation of (ii), it is not clear that any one person can dictate the group's decision.[12] Therefore, (ii) should rather be viewed as a non-degeneracy condition; it guarantees that we are talking about genuine group decision procedures, and not procedures which represent no real generalization from the case of individual decision making.[13]

One might expect probability and utility to play a dual role, so that (i) and (ii) equally well can be expressed in terms of one or the other. In our framework, this is not strictly the case. A condition parallel to (i) for utilities is stronger than (i) and can therefore be substituted for it in the theorems. If we add an extra condition relating to the choice among actions with equal expected utility, the two conditions become equivalent; see Section 7. In order to formulate a "utility condition" which can replace (ii), we have to rule out certain "trivial" cases, namely ones in which one action dominates another, so that uncertainty (the probability

[11] There is a problem here: If one person is an undisputed expert, one should expect everybody to adopt that person's probability assessment when all information, including the expert's credentials and assessment, has been communicated. But we would like to have a procedure which works even if this does not happen, that is, if the expert is not undisputed after all. In such a case there may be something "dictatorial" in letting the expert decide the group's probability assessment. If anything, these considerations strengthen the case for including condition (ii).

[12] The rest of our axioms will, however, imply that in "nontrivial" situations the chosen action is always one to which the "probability dictator" assigns maximal utility. See the last paragraph of Section 8, where the condition of nontriviality is also made precise.

[13] We thank Thomas Schelling for having drawn our attention to the issue discussed in this paragraph.

assessment) plays no role in the decision problem. This issue is discussed in Section 8.

What should a procedure for group decision making under uncertainty look like, given that it cannot satisfy all the conditions mentioned above? We will not discuss this issue in detail, just raise a few points. One possibility would be to abandon the entire Bayesian framework or the principle of separate aggregation of beliefs and values. Two less drastic alternatives might be suggested:

(a) One can give up the attempt to construct a procedure which works for all combinations of individual probability assessments and utilities. In particular, it may be claimed that there is something inherently irrational in the members of a group maintaining different probability assessments when they have had unlimited opportunity of exchanging information and trying to convince each other. And if we only permit different individual utility functions, not different probability assessments, the problems posed by the theorems of this paper disappear.

(b) The applicability of the Pareto principle can be questioned. Individual preferences are based on individual probability assessments which may differ from the group's assessment and which, therefore, from the group's point of view may be incorrect. Even if we accept the individuals' right to determine their own utility functions, we need not have the same respect for erroneous assessments of facts.

In this analysis, we assume that individual probability assessments and utilities are revealed honestly. If this is not the case and strategic behavior is possible when people provide these data, new problems arise, as is well known.

2. STATEMENT OF THE PROBLEM

A problem of group decision making under uncertainty is characterized by (i) a set of possible actions,

$$A=\{a_1,\ldots,a_m,\ldots,a_M\};$$

(ii) a set of potential states of the world,

$$S=\{s_1,\ldots,s_n,\ldots,s_N\}; \qquad \text{and}$$

(iii) a group of individuals, designated $1,\ldots,k,\ldots,K$.[14] We assume M, N, $K\geq 2$.

For each person k there is given a vector $p^{(k)}=(p_1^{(k)},\ldots,p_N^{(k)})$, satisfying

(1) $$p_n^{(k)}\geq 0 \qquad \text{for all } n,$$

(2) $$\sum_{n=1}^{N} p_n^{(k)}=1.$$

The vector $p^{(k)}$ represents person k's probability assessments for the states of the

[14] Finiteness of A and S is not required; see note 6 above. Neither is it essential that the set of individuals be finite; see discussion in the Appendix.

world. A vector of length N satisfying (1) and (2) will be called a *probability vector.*

For each k there is also given an $M \times N$ matrix, $u^{(k)}$, where $u^{(k)}_{mn}$ represents person k's utility for the outcome which results if action a_m is applied and state s_n prevails.

We will let $E(a_m | p, u)$ denote the expected utility, according to probability assessment p and utility scheme u, of applying action a_m. That is,

$$E(a_m | p, u) = \sum_{n=1}^{N} p_n u_{mn}$$

We will denote by $B(p, u)$ the set of actions with maximal expected utility according to p and u. That is,

$$a_m \in B(p, u)$$

if and only if

$$E(a_m | p, u) \geq E(a_i | p, u) \qquad \text{for all } i = 1, \ldots, M.$$

A system of probability assessments for all individuals will be called a *probability profile.* Formally, a probability profile is a vector $(p^{(1)}, \ldots, p^{(k)}, \ldots, p^{(K)})$, where each $p^{(k)}$ is a probability vector. Similarly, a vector $(u^{(1)}, \ldots, u^{(k)}, \ldots, u^{(K)})$ of $M \times N$ matrices, giving the utility scheme for each individual, will be called a *utility profile.* The concatenation $(p^{(1)}, \ldots, p^{(K)}; u^{(1)}, \ldots, u^{(K)})$ of a probability profile and a utility profile will be called a *situation.* A situation thus specifies the probability assessments and utility schemes of all individuals.

3. AXIOMS AND PRINCIPAL RESULT

AXIOM 1—Universal Applicability: *There exists a group decision procedure, applicable to all situations* $(p^{(1)}, \ldots, p^{(K)}; u^{(1)}, \ldots, u^{(K)})$. *For any situation, the procedure specifies a non-empty subset of A, called the choice set.*

AXIOM 2—Aggregation of Probability Assessments: *There exists a group probability assessment* $P = (P_1, \ldots, P_N)$, *where* $P_n \geq 0$ *for all n and* $\Sigma_{n=1}^{N} P_n = 1$. *P is a function only of the individual probability assessments; that is, there exists a function f such that* $P = f(p^{(1)}, \ldots, p^{(K)})$.

AXIOM 3—Aggregation of Utility Schemes: *There exists a group utility scheme U which is a function only of the individual utility schemes; that is, there exists a function g such that* [15]

$$U = g(u^{(1)}, \ldots, u^{(K)}).$$

AXIOM 4—Bayesian Group Decision: *The choice set is a function only of P and U as given by Axioms 2 and 3, and it is a subset of* $B(P, U)$.

[15] It may be in better accordance with the spirit of Bayesian decision making to make g a function not of $u^{(1)}, \ldots, u^{(K)}$ as formal mathematical entities, but of the utility functions represented by $u^{(1)}, \ldots, u^{(K)}$. Formally, this amounts to requiring that whenever each $\tilde{u}^{(k)}$ is a positive linear transformation of $u^{(k)}$, then $g(\tilde{u}^{(1)}, \ldots, \tilde{u}^{(K)})$ is a positive linear transformation of $g(u^{(1)}, \ldots, u^{(K)})$. The only reason not to impose this added restriction is that it is unnecessary for our proofs.

AXIOM 5—Pareto Optimality (weak form): *If* $E(a_m|p^{(k)}, u^{(k)}) > E(a_i|p^{(k)}, u^{(k)})$ *for all k, then* a_i *is not an element of the choice set.*

AXIOM 6—Unanimous Beliefs Prevail: *If every individual has the same probability assessment, the group's assessment is identical to this. That is, for all probability vectors p,*

$$f(p, \dots, p) = p.$$

AXIOM 7—Non-dictatorial Group Probability Assessment: *There does not exist any individual k such that*

$$f(p^{(1)}, \dots, p^{(K)}) = p^{(k)}$$

for all probability profiles $(p^{(1)}, \dots, p^{(K)})$.

The choice set for situation $(p^{(1)}, \dots, p^{(K)}; u^{(1)}, \dots, u^{(K)})$ will be denoted $C(f(p^{(1)}, \dots, p^{(K)}), g(u^{(1)}, \dots, u^{(K)})) = C(P, U)$, which is a legitimate notation in view of Axiom 4.

Axioms 2 and 3 are vacuous in themselves, but when combined with Axiom 4, they express the principle of separate aggregation of beliefs and values. Axiom 4 requires that the group decision making be "almost Bayesian," insomuch as a chosen action must be one for which expected group utility is maximized. If several actions are tied in expected utility, we allow the group to employ some secondary principle to choose among them. For example, the maximin principle can be used. The tie-breaking mechanism can also involve P; for example, ties can be broken on basis of some fractile of the utility distributions given by the actions. But only information included in P and U can be used; if two situations lead to identical group probability assessments and utility schemes, they must have exactly the same choice set.[16]

THEOREM 1—Principal Result: *Axioms 1–7 are inconsistent.*

4. HEURISTIC PROOF OF THE CASE $M = N = K = 2$

Theorem 1 is proved formally in Section 5 and the Appendix. In this section, an informal argument is given that is more accessible and brings out the intuitive motivation.

We assume that there are two actions, two states, and two individuals. Then a probability assessment p is characterized by a single number p_1 in the interval $[0, 1]$, representing the probability of s_1. An individual's utility assessment for an action can then be represented by two points and a straight line between them, as shown in Figure 1. The distance AB is $E(a_m|p, u)$.

[16] The faithful Bayesian may object to the possibility of $C(P, U)$ being a proper subset of $B(P, U)$. If a person's beliefs and values satisfy the consistency requirements which make the use of subjective probabilities and utilities appropriate, then the person is truly indifferent between two actions with equal expected utility, and there can be no basis for including the one and not the other in the choice set. Our answer to this is that P and U are not really individual probability assessments and utility schemes, therefore, the argument need not apply. Moreover, the axiom does not prohibit $C(P, U) = B(P, U)$; it just opens further possibilities.

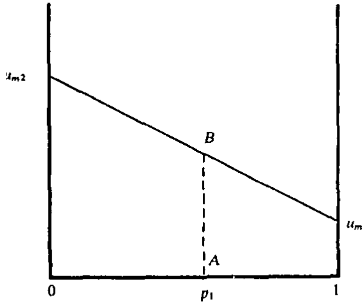

FIGURE 1

In the examples below, both individuals assign utility 0 to action a_2 in both states. Hence a_2 is represented by the axis. Individual k's preference ordering of a_1 and a_2, for a given probability assessment, depends only on whether the line representing $u_{1\cdot}^{(k)}$ lies above or below the axis at the relevant p_1 value.

The most natural non-dictatorial aggregation procedure for probability assessments, is one for which $P = f(p^{(1)}, p^{(2)})$ is a weighted average of $p^{(1)}$ and $p^{(2)}$. This is shown in Figure 2, where two possible utility schemes are also drawn. We see that in the situation $(p^{(1)}, p^{(2)}; u^{(1)}, u^{(2)})$, both individuals strictly prefer a_1 to a_2. Hence a_1 must be chosen, by the Pareto principle. In the situation $(P, P; u^{(1)}, u^{(2)})$, a_2 must be chosen for the same reason. But this is impossible. By

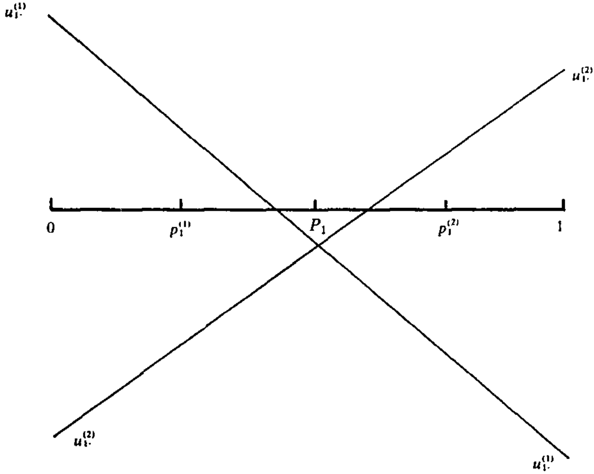

FIGURE 2

assumption and Axiom 6, the group probability assessment is the same for the two situations. The group utility scheme is unchanged because no individual scheme has changed. Then the principle of separate aggregation implies that the group decision must be the same in the two cases, contrary to what we just found.

If P_1 lies outside the interval given by $p_1^{(1)}$ and $p_1^{(2)}$, a similar example can be constructed. Hence, for all $p^{(1)}$ and $p^{(2)}$, $f(p^{(1)}, p^{(2)})$ is equal to either $p^{(1)}$ or $p^{(2)}$.

Only one possibility remains for f being non-dictatorial; perhaps there exist $p^{(1)}$, $p^{(2)}$, $q^{(1)}$, and $q^{(2)}$ such that $p^{(1)} \neq p^{(2)}$, $q^{(1)} \neq q^{(2)}$, $f(p^{(1)}, p^{(2)}) = p^{(1)}$, and $f(q^{(1)}, q^{(2)}) = q^{(2)}$. Let $u^{(1)}$ and $u^{(2)}$ be as shown in Figure 3, and consider the situations $(p^{(1)}, p^{(2)}; u^{(1)}, u^{(2)})$ and $(q^{(2)}, p^{(1)}; u^{(1)}, u^{(2)})$. In the former, a_1 is chosen, in the latter, a_2 must be the group decision. The group utility scheme is the same in the two situations. If the group probability assessment is also equal, a contradiction can be derived as above. Hence $f(q^{(2)}, p^{(1)}) \neq f(p^{(1)}, p^{(2)}) = p^{(1)}$.

By a similar argument, $f(q^{(2)}, p^{(1)}) \neq f(q^{(1)}, q^{(2)}) = q^{(2)}$. But we have just found that $f(q^{(2)}, p^{(1)})$ must be either $q^{(2)}$ or $p^{(1)}$. The argument does not depend on the particular ordering of $p^{(1)}$, $p^{(2)}$, and $q^{(2)}$ in Figure 3; appropriate utility schemes $u_{1\cdot}^{(1)}$ and $u_{1\cdot}^{(2)}$ can always be found. We have assumed that $p^{(1)}, p^{(2)}, q^{(1)}$, and $q^{(2)}$ are all different. But if they are not, we can find $\bar{p}^{(1)}$ and $\bar{p}^{(2)}$ which are different from all of them. Then the argument can be applied twice; first to $p^{(1)}, p^{(2)}, \bar{p}^{(1)}$, and $\bar{p}^{(2)}$, and then to $\bar{p}^{(1)}$, $\bar{p}^{(2)}$, $q^{(1)}$, and $q^{(2)}$. The conclusion is that f must be dictatorial.

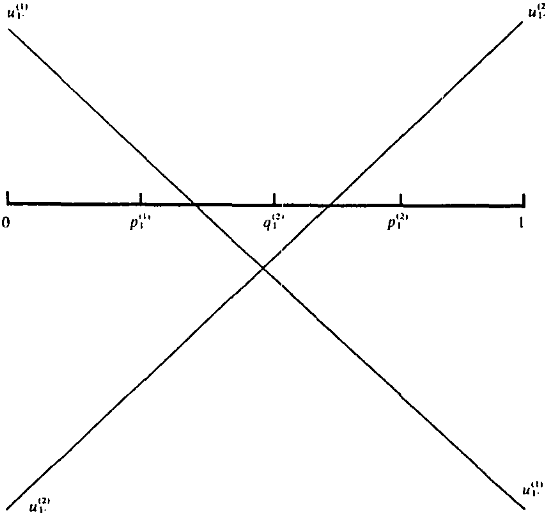

FIGURE 3

In Section 6, we strengthen the Pareto principle in a way which, when $M = K = 2$, implies that a_1 must be chosen whenever one individual prefers a_1 to a_2 and the other is indifferent. Then a contradiction can be derived without using the assumption that f is non-dictatorial (Axiom 7). Choose $p^{(1)}$ and $p^{(2)}$ such that $p^{(1)} \neq p^{(2)}$. As above, $f(p^{(1)}, p^{(2)})$ is either $p^{(1)}$ or $p^{(2)}$; assume the former. Let $u_{1n}^{(1)} = 0$ for $n = 1, 2$, and let $u_{1\cdot}^{(2)}$ be given by Figure 4. Consider the situations $(p^{(1)}, p^{(2)}; u^{(1)}, u^{(2)})$ and $(p^{(1)}, p^{(1)}; u^{(1)}, u^{(2)})$. By assumption and Axiom 6, the group probability assessment is the same; hence the group decision must also be equal. But the strengthened Pareto principle requires that a_1 be chosen in the first and a_2 in the second situation.

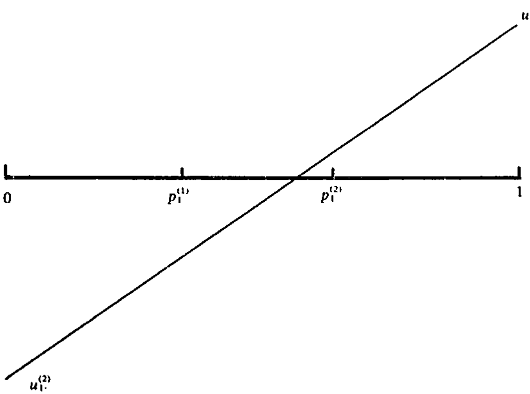

FIGURE 4

5. RIGOROUS PROOF OF THEOREM 1 IN THE GENERAL CASE

The proof is comprised of two lemmas.

LEMMA 1: *Let the probability profiles* $(p^{(1)}, \ldots, p^{(K)})$ *and* $(q^{(1)}, \ldots, q^{(K)})$ *satisfy*

(3) $\qquad p^{(k)} \neq q^{(k)} \qquad$ *for all* $k = 1, \ldots, K$.

Then

(4) $\qquad f(p^{(1)}, \ldots, p^{(K)}) \neq f(q^{(1)}, \ldots, q^{(K)})$

PROOF: Suppose that the lemma is wrong, and let $(p^{(1)}, \ldots, p^{(K)})$ and $(q^{(1)}, \ldots, q^{(K)})$ provide a counter-example. That is, assume that (3) is true and (4) is false. For any k, $p^{(k)}$ and $q^{(k)}$ satisfy (2). Hence they are both non-zero vectors. Equations (2) and (3) imply that they are not parallel; therefore, there exists a vector $x^{(k)} = (x_1^{(k)}, \ldots, x_N^{(k)})$ with

(5) $\qquad x^{(k)} \cdot p^{(k)} < 0,$

(6) $\qquad x^{(k)} \cdot q^{(k)} > 0.$

Define a utility scheme $u^{(k)}$ by

$$(7)\qquad \begin{aligned} u^{(k)}_{1n} &= x^{(k)}_n \qquad \text{for } n = 1, \ldots, N; \text{ and} \\ u^{(k)}_{mn} &= 0 \qquad \text{for } m = 2, \ldots, M \text{ and } n = 1, \ldots, N. \end{aligned}$$

Having done this for all k, we consider the situations

(i) $(p^{(1)}, \ldots, p^{(K)}; u^{(1)}, \ldots, u^{(K)})$,

(ii) $(q^{(1)}, \ldots, q^{(K)}; u^{(1)}, \ldots, u^{(K)})$.

By assumption, the group probability assessment is the same in these two situations. Obviously, the group utility scheme is also the same. Hence, by Axiom 4, the choice sets in situations (i) and (ii) are equal.

Equations (5) and (7) give

$$(8)\qquad E(a_1 | p^{(k)}, u^{(k)}) < E(a_2 | p^{(k)}, u^{(k)})$$

for all k. Hence, by Axiom 5, a_1 cannot belong to the choice set for (i).

For situation (ii), (6) and (7) give

$$(9)\qquad E(a_1 | q^{(k)}, u^{(k)}) > E(a_m | q^{(k)}, u^{(k)}),$$

for $m \geqslant 2$ and all k. Therefore, none of the actions $a_2, \ldots, a_M$ can belong to the choice set for (ii).

Hence the common choice set for situations (i) and (ii) must be empty, contradicting Axiom 1 and proving Lemma 1.

Lemma 2 is a purely mathematical result; that is, it makes no reference to the axioms. It is proved and discussed in the Appendix.

LEMMA 2: *Let J be a positive integer and let Y be a set with at least J elements. (Y can be finite or infinite.) Suppose that h is a function from the Cartesian product Y^J into Y. Assume*: (i) *For all $y \in Y$, $h(y, \ldots, y) = y$;* (ii) *if $(y_1, \ldots, y_J)$ and $(z_1, \ldots, z_J)$ satisfy $y_j \neq z_j$ for all $j = 1, \ldots, J$, then*

$$h(y_1, \ldots, y_J) \neq h(z_1, \ldots, z_J).$$

Then there exists a j_0 with $1 \leqslant j_0 \leqslant J$ such that

$$h(y_1, \ldots, y_J) = y_{j_0}$$

for all $(y_1, \ldots, y_J) \in Y^J$.

The theorem easily follows from the two lemmas. Let Y in Lemma 2 be the set of probability vectors, which is an infinite set, and set $J = K$ and $h = f$. Part (i) is satisfied by Axiom 6 and (ii) by Lemma 1. The conclusion of Lemma 2 contradicts Axiom 7. The proof of Theorem 1 is complete.

6. THE PARETO PRINCIPLE

Axiom 5 represents the weak Pareto principle; it excludes a_i from the choice set only when there exists some a_m such that $E(a_m | p^{(k)}, u^{(k)})$ is strictly greater than

$E(a_i|p^{(k)}, u^{(k)})$ for all k. A stronger form of the principle, excluding a_i from the choice set under weaker conditions, is given by the following axiom:

AXIOM 5*—Pareto Optimality (strong form): *If* $E(a_m|p^{(k)}, u^{(k)}) \geq E(a_i|p^{(k)}, u^{(k)})$ *for all k, with strict inequality holding for at least one k, then* a_i *is not an element of the choice set.*

If Axiom 5* is assumed, the explicit exclusion of dictatorial group probability assessment (Axiom 7) can be omitted, and the result will still hold. Also, the proof does not use Lemma 2.

THEOREM 2: *Axioms* 1–4, 5*, *and* 6 *are inconsistent.*

PROOF: Find K different probability vectors $p^{(1)}, \ldots, p^{(K)}$. (This is possible since the set of probability vectors is infinite.) Let $P = f(p^{(1)}, \ldots, p^{(K)})$. P can at most be equal to one of the vectors $p^{(1)}, \ldots, p^{(K)}$, and there is no loss of generality in assuming $P \neq p^{(k)}$ for $k \geq 2$.

P and $p^{(k)}$ satisfy (2); therefore, for each $k \geq 2$, we can find a vector $x^{(k)} = (x_1^{(k)}, \ldots, x_N^{(k)})$ satisfying (5) and

(10) $\qquad x^{(k)} \cdot P > 0.$

Let $x^{(1)} = (0, \ldots, 0)$, and define $u^{(k)}$ by (7). Consider the situations (i) and

(iii) $\qquad (P, \ldots, P; u^{(1)}, \ldots, u^{(K)}).$

By Axiom 6 and the definition of P, the group probability assessment is the same in (i) and (iii). It follows that the choice set is the same in the two situations. Equation (8) holds for $k \geq 2$; for $k = 1$, equality holds in (8). Hence, by Axiom 5*, a_1 is not in the choice set. For situation (iii), we get, for $m = 2, \ldots, M$:

$$E(a_1|P, u^{(1)}) = E(a_m|P, u^{(1)})$$

and

$$E(a_1|P, u^{(k)}) > E(a_m|P, u^{(k)}), \qquad \text{for } k \geq 2$$

Axiom 5* then excludes the actions $a_2, \ldots, a_M$ from the choice set; we have a contradiction to Axiom 1, and Theorem 2 is proved.[17]

7. UNANIMOUS BELIEFS OR VALUES

Corresponding to Axiom 6, one can formulate a similar principle for utilities. This can be done in several ways; the following relatively weak axiom will suffice:

[17] Since $p^{(1)}, \ldots, p^{(K)}$ all were chosen different and $x^{(1)}$ was defined equal to $(0, \ldots, 0)$, we have actually used a principle which is weaker than Axiom 5*; we have only applied the axiom when strict inequality holds for at least $K-1$ values of k, and when the one individual for whom strict inequality does not hold is completely unconcerned (that is, has a constant utility function).

AXIOM 6′—Unanimous Values Prevail: *For all utility schemes u and all probability vectors p,*[18]

$$B(p, g(u, \ldots, u)) \subseteq B(p, u).$$

Axiom 6′ places restrictions on $g(u^{(1)}, \ldots, u^{(K)})$ only when $u^{(1)}, \ldots, u^{(K)}$ are formally equal, not when they are positive linear transformations of each other and therefore represent the same utility function.[19]

Informally, Axiom 6′ says the following: When $U = g(u, \ldots, u)$, U is not required to be equal to u; U is allowed to discriminate among actions that are treated equally by u. But for any probability vector, U must choose its preferred actions from among those which have maximal expected utility according to u.

Axiom 6′ can replace Axiom 6 in Theorems 1 and 2, as shown by the following lemma:

LEMMA 3: *Assume Axioms* 1–5 *hold. Then Axiom* 6′ *implies Axiom* 6.

PROOF: Suppose Axioms 1–5 and 6′ hold, and assume, in violation of Axiom 6, that p and P are probability vectors with $P = f(p, \ldots, p)$ and $P \neq p$. As before, we can find $x = (x_1, \ldots, x_N)$ such that

(11) $$x \cdot P > 0,$$

(12) $$x \cdot p < 0.$$

Let u be defined as in (7), set $U = g(u, \ldots, u)$, and consider the situation

$$(p, \ldots, p; u, \ldots, u).$$

From (11), $B(P, u) = \{a_1\}$. By Axioms 4 and 6′, $C(P, U) \subseteq B(P, U) \subseteq B(P, u)$. Hence $C(P, U) = \{a_1\}$. This contradicts (12) and Axiom 5, and the lemma is proved.

The converse implication is not true; in the presence of Axioms 1–5, Axiom 6′ is in a trivial way stronger than Axiom 6.[20] If we strengthen Axiom 4 and require $C(P, U) = B(P, U)$, the two become equivalent.

Axiom 6 is necessary for the proofs of Theorems 1 and 2. If it is omitted (and no substitute inserted), the whole idea of separate aggregation can be "subverted." Then it is possible to build into P complete information about $p^{(1)}, \ldots, p^{(K)}$, and

[18] An alternative is to require that $g(u, \ldots, u)$ be equal to u, or to a positive linear transformation of u. This clearly implies Axiom 6′, and in a formal sense it is stronger, as the following example illustrates: Let a_1 dominate a_2 in u, that is, assume $u_{1n} > u_{2n}$ for all n. Then we want a_2 to be dominated in $U = g(u, \ldots, u)$ as well, which is guaranteed by Axiom 6′. But there is no reason to worry about the particular values of U_{2n}.

[19] See note 15.

[20] Let $f(p^{(1)}, \ldots, p^{(K)}) = p^{(1)}$. Define $g(u^{(1)}, \ldots, u^{(K)}) = U$ by $U = u^{(1)}$ except when $u^{(1)}_{1n} > u^{(1)}_{2n}$ for all n, in which case $U_{2n} = u^{(1)}_{1n}$ and $U = u^{(1)}$ otherwise. Finally, let $C(P, U) = \{a_2\}$ if $B(P, U) = \{a_2\}$, otherwise, $C(P, U) = B(P, U) - \{a_2\}$. This system satisfies Axioms 1–6, but not Axiom 6′.

similarly for U and $u^{(1)}, \ldots, u^{(K)}$. C can then retrieve this information and choose an action consistent with Axiom 5 (or 5*).[21]

As argued in the Introduction, Axiom 6 is something one should accept if one accepts the Pareto principle. Its necessity, therefore, is not an important weakening of the results.[22]

8. DICTATORSHIP

Clearly, Axiom 7 is necessary. The straightforward "dictatorship procedure," given by $f(p^{(1)}, \ldots, p^{(K)}) = p^{(1)}$, $g(u^{(1)}, \ldots, u^{(K)}) = u^{(1)}$, and $C(P, U) = B(P, U)$, satisfies Axioms 1–6 (and 6′).

Can Axiom 7, in the same way as Axiom 6, be replaced by a principle which relates to utilities instead of probabilities? The immediate answer is no. A "utility dictator" can be defined in several ways, of which we choose the following: Individual k is said to be a utility dictator if and only if

$$(13) \qquad B(p, g(u^{(1)}, \ldots, u^{(K)})) \subseteq B(p, u^{(k)})$$

for all probability vectors p and all utility profiles $(u^{(1)}, \ldots, u^{(K)})$.

The intuitive meaning of this is as follows: Apply the group probability assessment to individual k's utility scheme and find the actions which maximize expected utility. If k is able to dictate group utility, the chosen action should be found among these. And it is easily seen that (13), together with Axiom 4, implies this.

This definition of utility dictator is, however, a little too strong, and hence an axiom saying that such a dictator does not exist is too weak to be substituted for Axiom 7 in Theorem 1. That is, there exists procedures satisfying Axioms 1–6 in which there is no utility dictator.[23]

EXAMPLE: Let $M = K = 2$, and define f and C by $f(p^{(1)}, p^{(2)}) = p^{(1)}$ and $C(P, U) = B(P, U)$. If one action dominates the other in $u^{(2)}$, that is, if $u_{1n}^{(2)} > u_{2n}^{(2)}$ for all n or the opposite inequality holds for all n, then $g(u^{(1)}, u^{(2)}) = u^{(2)}$;

[21] The point is that the arguments of f and g are a finite number of real numbers. It is possible to code any finite string of real numbers into one real number in an unambiguous way. (The coding and decoding functions will not have such nice properties as continuity, but they will be effectively computable, in the strongest sense this can ever be true for real-valued functions.) Now let $P = (p^*, 1 - p^*, 0, \ldots, 0)$, where p^* codes $p^{(1)}, \ldots, p^{(K)}$, while $U_{mn} = u^*$ for all m and n, where u^* codes $u^{(1)}, \ldots, u^{(K)}$. $B(P, U)$ is then always equal to A, and C can use all information about $p^{(1)}, \ldots, p^{(K)}$ and $u^{(1)}, \ldots, u^{(K)}$ without violating Axiom 4. Therefore, Axioms 1–5 (and 5*) can be satisfied. Axiom 7 also holds.

[22] It is necessary for the example of note 21 that we allow $C(P, U) \neq B(P, U)$. The authors conjecture that Axiom 6 can be dropped from the premise of Theorem 2 if Axiom 4 is strengthened to require $C(P, U) = B(P, U)$. To do the same in Theorem 1, it may be necessary to strengthen Axiom 7 by requiring that there be no k such that $f(p^{(1)}, \ldots, p^{(K)})$ is a function only of $p^{(k)}$. But it is not immediate how this should be proved; and in view of the discussion of Axiom 6 in the Introduction, we do not find it very important to eliminate it from the premise of the theorems.

[23] Alternatives to (13) might be to reverse the inclusion sign, to replace it by equality, or to require that $g(u^{(1)}, \ldots, u^{(K)})$ be equal to (a positive linear transformation of) $u^{(k)}$. The corresponding non-dictatorship axioms will all be consistent with Axioms 1–6 (and remain so even if reasonable extra conditions are added).

otherwise, $g(u^{(1)}, u^{(2)}) = u^{(1)}$. This system satisfies Axioms 1–6, and in no way can one say that one person dictates the group utility function. Neither will any one individual unilaterally determine the group's final action.

The procedure of the previous paragraph is trivial in the sense that person 2 has influence only when one action dominates the other, that is, only when probability assessments play no role in determining person 2's ranking of the actions. If such cases are ruled out, Axioms 1–6 will imply the existence of an individual who dictates the group's decision. To be precise, assume that a procedure is given which satisfies Axioms 1–6. By Theorem 1, a probability dictator exists; let this be person 1. Moreover, suppose that action a_1 is chosen in the situation $(p^{(1)}, \ldots, p^{(K)}; u^{(1)}, \ldots, u^{(K)})$ although person 1 prefers a_2 to a_1. That is, suppose $a_1 \in C(P, U)$ and $E(a_1|p^{(1)}, u^{(1)}) < E(a_2|p^{(1)}, u^{(1)})$. Then there has to be some individual k, where $2 \leq k \leq K$, for whom a_1 weakly dominates a_2, that is, $u_{1n}^{(k)} \geq u_{2n}^{(k)}$ for all n. Therefore, whenever the choice between a_1 and a_2 really is a decision problem under uncertainty for all individuals $2, \ldots, K$, then person 1 dictates this choice. If we also require $C(P, U) = B(P, U)$, (13) will essentially hold, with $k = 1$, under the same nontriviality conditions.[24]

9. RANDOMIZED DECISIONS

Can one avoid the dilemma posed by Theorems 1 and 2 by using randomized decision rules? It is well known that an individual Bayesian decision maker can never gain by introducing randomized decision rules; for any such rule there exists a deterministic one having at least as high expected utility. In the group decision problem, therefore, $B(P, U)$ will always contain a deterministic decision. But $B(P, U)$ can also contain randomized decisions, and it is conceivable that the group can have reasons for choosing one of these.

When randomized decisions are allowed, one must take care to formulate the axioms correctly. It is not possible to let A consist of all deterministic or randomized decisions and apply the axioms as formulated above; since Axiom 1 then will contradict the expected utility principle. It turns out, however, that even when Axiom 1 is appropriately weakened, the results above hold when randomized decisions are permitted.

The proof of this is not difficult. Let there be at least two pure actions, suppose that randomized actions are permitted, and modify Axiom 1 to allow arbitrary assignment of utility values only for pure actions. (Axiom 5 is still supposed to apply to all actions.) The proofs can then be carried through essentially unchanged. In the proof of Lemma 1, the definition (7) of $u^{(k)}$ shall assign value 0 for all deterministic actions except a_1. Equation (8) excludes the deterministic action a_1 from the choice set. Equation (9) will be true for, and will exclude from

[24] The example in note 20 shows that the added condition $C(P, U) = B(P, U)$ is necessary. The precise sense in which (13) holds is the following: Suppose, for some p and $u^{(1)}, \ldots, u^{(K)}$, that $a_1 \in B(p, g(u^{(1)}, \ldots, u^{(K)}))$ but not $a_1 \in B(p, u^{(1)})$. Choose any a_2 with $E(a_1|p, u^{(1)}) < E(a_2|p, u^{(1)})$. (Such an a_2 must exist; for example, any element of $B(p, u^{(1)})$ can be chosen.) Then there exists a person k in whose utility function a_1 dominates a_2, as in the text.

the choice set, any action a_m which includes a positive probability of any pure action but a_1. Lemma 1 follows. Similar modification can be made in the other parts of the proofs.

Harvard University

Manuscript received March, 1978; revision received November, 1978.

APPENDIX

Proof and Discussion of Lemma 2

LEMMA 2: *Let J be a positive integer and let Y be a set with at least J elements. (Y can be finite or infinite.) Suppose that h is a function from the Cartesian product Y^J into Y. Assume:* (i) *For all $y \in Y$, $h(y, \ldots, y) = y$;* (ii) *if $(y_1, \ldots, y_J)$ and $(z_1, \ldots, z_J)$ satisfy $y_j \neq z_j$ for all $j = 1, \ldots, J$, then $h(y_1, \ldots, y_J) \neq h(z_1, \ldots, z_J)$. Then there exists a j_0 with $1 \leq j_0 \leq J$ such that*

$$h(y_1, \ldots, y_J) = y_{j_0}$$

for all $(y_1, \ldots, y_J) \in Y^J$.

This is a purely mathematical result; that is, it makes no reference to and does not depend on the axioms listed in Section 3. It looks like a general combinatorial result, and may therefore be of interest in connections other than ours. We do not know any reference to other applications or earlier proofs.

In the presence of (ii), (i) is equivalent to: (i') For all $(y_1, \ldots, y_J) \in Y^J$, there exists a j with $1 \leq j \leq J$ and

$$h(y_1, \ldots, y_J) = y_j$$

It is obvious that (i') implies (i). To prove the converse, suppose that (i) and (ii) hold, and assume that (i') is false. Then there exist $y_1, \ldots, y_J$ and z such that $z = h(y_1, \ldots, y_J)$ and $z \neq y_j$ for all j. Hence, by (ii), $h(z, \ldots, z) \neq h(y_1, \ldots, y_J)$. But $h(z, \ldots, z) = z$ by (i) and we have a contradiction; thus (i') is proved.

When (i') is substituted for (i), Lemma 2 can be interpreted as a uniformization result. The premise (i') says that for each $(y_1, \ldots, y_J)$ there exists a j such that $h(y_1, \ldots, y_J) = y_j$, but j may depend on $(y_1, \ldots, y_J)$. The conclusion of the lemma is that j can be chosen uniformly, that is, independently of $(y_1, \ldots, y_J)$.

PROOF OF LEMMA 2: The case $J = 1$ is trivial. Hence we can assume $J \geq 2$. We can also use (i').

Choose J different elements $y_1, \ldots, y_J$ from Y; this is possible by assumption. Elements $y_1, \ldots, y_J$ are kept fixed throughout the proof. By (i'), $h(y_1, \ldots, y_J)$ is equal to one of the elements y_j. There is no real loss of generality in assuming

(A1) $\qquad h(y_1, \ldots, y_J) = y_1.$

When we have made this assumption, we will of course get $j_0 = 1$ in the conclusion of the lemma.

First, consider elements of Y^J in which the $J-1$ last coordinates are equal. We shall prove that for all $y, z \in Y$,

(A2) $\qquad h(y, z, \ldots, z) = y.$

For $y = z$, this follows from (i). Otherwise, $h(y, z, \ldots, z)$ is equal to either y or z by (i'); hence we need only prove that it is different from z. We consider three cases:

(a) $y \neq y_1$; $z = y_1$. Then $(y_1, \ldots, y_J)$ and $(y, z, \ldots, z)$ are different on each coordinate by the choice of $y_1, \ldots, y_J$. (ii) and (A1) give the desired conclusion.

(b) $y = y_1$, $z \neq y_1$. By case (a), $h(z, y, \ldots, y) = z$. The vectors $(y, z, \ldots, z)$ and $(z, y, \ldots, y)$ are different on each coordinate, and (A2) again follows by (ii).

(c) y_1, y, and z are all different. Assume, contrary to (A2), that $h(y, z, \ldots, z) = z$. The vectors $(y, z, \ldots, z)$ and $(z, y_1, \ldots, y_1)$ are different on each coordinate, and (ii) gives $h(z, y_1, \ldots, y_1) \neq z$. This contradicts case (a); hence (A2) must hold here also.

If the lemma is wrong, there must exist $(z_1, \ldots, z_J) \in Y^J$ with

(A3) $$h(z_1, \ldots, z_J) = y \neq z_1.$$

Since Y has at least J elements, there exists a $z \in Y$ which is different from all of $z_2, \ldots, z_J$. Now the vectors $(y, z, \ldots, z)$ and $(z_1, \ldots, z_J)$ are different on all coordinates. Then (A2) and (A3) contradict (ii). This completes the proof of Lemma 2.

The lemma is also true if J is infinite, provided that the cardinality of Y is greater than or equal to that of J. The set Y^J, on which h is defined, must be thought of as the set of functions from J into Y. (If J is countable, Y^J is the set of infinite sequences of elements from Y.) The proof of (A2) applies without change. The last paragraph of the proof must, however, be modified somewhat. (When J and Y are of equal, infinite cardinality, it is possible that $z_2, z_3, \ldots$ exhaust all of Y.)

The lemma may be wrong if Y has fewer elements than J. For example, assume that $J > 2$ and Y has two elements, and define $h(y_1, \ldots)$ to be the element of Y which occurs two or three times among y_1, y_2, y_3. This is well-defined and satisfies (i) and (ii), but the conclusion of the lemma does not hold.

On the other hand, the condition that Y has at least J elements is not the weakest possible premise. For example, the lemma is true if $J = 4$ and Y has three elements. We have not made any attempt to find the weakest possible condition on the cardinality of Y, as it depends on J.

REFERENCES

[1] RAIFFA, HOWARD: *Decision Analysis.* Reading, Massachusetts: Addison-Wesley, 1968.

[2] ZECKHAUSER, RICHARD: "Group Decision and Allocation," Discussion Paper #51, Harvard Institute of Economic Research, Cambridge, Massachusetts, 1968.