Marginalization and Linear Opinion Pools

K.J. McCONWAY*

Suppose a decision maker has asked a group of experts to assess subjective probability distributions over some space, and he wishes to find a consensus probability distribution as a function of these. The assumption that finding the consensus distribution commutes with marginalization of the distributions implies that the consensus must be found using a linear opinion pool (weighted average), provided the space being considered contains three or more points. Some of the consequences of this result are discussed.

KEY WORDS: Opinion pool; Marginalization; Consensus; Subjective probability; Cauchy's equation.

1. INTRODUCTION

Suppose that a decision maker needs to assess a subjective probability distribution over some space of states of nature for the purpose of performing some decision analysis. One of the most repeated suggestions in the literature as to how the decision maker should go about obtaining this distribution is that he or she should consult a group of other people, who are in some sense experts on the problem being studied, on their opinions about the possible states of nature. The decision maker should then base his or her probability distribution in some way on the experts' opinions as well as on his or her own.

In this paper we assume that all these opinions are expressed as subjective probability distributions over the appropriate space. We assume further that the experts, whom we shall refer to as *assessors,* do not interact directly with one another to decide on a group assessment; they are merely asked individually to assess distributions over the space, and the decision maker then combines these distributions into a *consensus* distribution using a formal mathematical function. We do not consider here whether this is the best approach to performing the aggregation of opinions: useful comments on this point are made in a review paper by Hogarth (1975). (See also Madansky 1978; Winkler 1968.) In fact the word *consensus* is something of a misnomer, particularly if the decision maker is not a member of the group, since the consensus distribution may not be the subject of any type of agreement by the experts.

* K.J. McConway is with the Faculty of Mathematics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom. Much of the work reported here was supported by a Science Research Council Research Studentship, tenable at University College London. The author wishes to thank A.P. Dawid for supervising the work, and for many invaluable suggestions.

2. LINEAR OPINION POOLS

An obvious method of combining the experts' distributions is to use a weighted average. Suppose there are n individual assessors, who have assessed distributions $\Pi_1, \ldots, \Pi_n$ over a space Ω. The decision maker chooses weights $\alpha_1, \ldots, \alpha_n$, nonnegative and summing to 1, and the consensus distribution Π_C over Ω is defined by putting, for any event A in Ω,

$$\Pi_C(A) = \sum_1^n \alpha_i \Pi_i(A).$$

Stone (1961) seems to have been the first to discuss this method in detail; he refers to it as the *opinion pool.* Bacharach (1974) calls this method the *linear opinion pool,* to distinguish it from the *logarithmic opinion pool,* which is essentially a weighted geometric mean.

Most of the subsequent work on linear opinion pools has been concerned with the implementation of the technique in practice, and in particular with the problem of choosing and revising the weights. The idea behind this is that the decision maker should give relatively high weights to assessors who have in some sense a high degree of expertise, either as probability assessors or with respect to the problem being considered. (See, for example, Staël von Holstein 1970, and Winkler 1971.) We note here that the decision maker's own prior subjective distribution can affect the consensus only if the decision maker is considered to be one of the assessors; this distribution is dealt with differently from the other assessors' distributions only in that it may be given a different weight in the opinion pool.

In this paper we investigate a possible theoretical justification for using linear opinion pools; we show that, under mild regularity conditions, if the process of finding the consensus distribution is to commute with any possible marginalization of the distributions involved, then a linear opinion pool must be used.

3. CONSENSUS FINDING AND MARGINALIZATION

We are given a group of n assessors ($2 \leq n < \infty$), and each has assessed a probability distribution over some σ-algebra S over a space Ω of states of nature. We wish to produce a consensus probability distribution over S, which will be a function of S and of the individual assessors' distributions.

© Journal of the American Statistical Association
June 1981, Volume 76, Number 374
Theory and Methods Section

As well as the implicit assumption that the result of the consensus-finding process is a probability distribution, we are also assuming that the assessors agree not only on the space of Ω of states of nature, but also on the σ-algebra S over Ω they are going to use. In practice this seems rather unimportant as, if the assessors agree on Ω, there will usually be some obvious S to use.

Formally, then, let $P(S)$ (depending on Ω) be the set of all probability measures over (Ω, S). Suppose the ith assessor ($i = 1, \ldots, n$) has assessed a distribution $\Pi_i(\,.\,) \in P(S)$, considered as a function $\Pi_i\colon S \to [0, 1]$. The consensus distribution is also in $P(S)$, and depends on S and $\Pi_1, \ldots, \Pi_n$. We denote it by $C_S(\Pi_1, \ldots, \Pi_n)$, so that the consensus probability of $A \in S$ is denoted by $\{C_S(\Pi_1, \ldots, \Pi_n)\}(A)$. C_S is called a *consensus function*; $C_S\colon \{P(S)\}^n \to P(S)$.

We shall be interested in comparing consensus functions for different S's (given a fixed Ω). Hence let us suppose that a consensus function has been defined for each different S so that we have a class of consensus functions $\mathscr{C}$ for Ω. Let Σ (depending on Ω) be the set of all σ-algebras over Ω.

Definition. $\mathscr{C}$ is a *class of consensus functions* if $\mathscr{C} = \{C_S\colon S \in \Sigma\}$, where $C_S\colon \{P(S)\}^n \to P(S)$ for all $S \in \Sigma$.

Note that every element of Σ is a sub-σ-algebra of the power set 2^Ω of Ω; if we do not wish to consider all possible σ-algebras we could instead let Σ be the set of all sub-σ-algebras of some maximal σ-algebra, $S_{\max}$ say, over Ω. The following proofs would still be valid (with obvious slight modifications).

It seems intuitively appealing that one should arrive at the same final consensus distribution whether (a) the assessors' distributions are first combined to form a consensus distribution, and then some marginalization is performed on this, or (b) the individual assessors all perform the marginalization separately, and the resulting individual marginal distributions are combined into a consensus distribution.

We formalize marginalization in the usual way. If Π is a distribution over (Ω, S) and T is any sub-σ-algebra of S, we define the marginal distribution $\Pi^{(T)}$ given by Π over (Ω, T) as the restriction of Π to T, so that $\Pi^{(T)}(B) = \Pi(B)$ for all $B \in T$. Then, to say that any marginalization commutes with finding the consensus, as previously stated, is to insist that $\mathscr{C}$ has the following property.

Marginalization Property (MP). A class $\mathscr{C}$ of consensus functions has the marginalization property if and only if, for all $S \in \Sigma$,

$$\{C_S(\Pi_1, \ldots, \Pi_n)\}^{(T)}(B) = \{C_T(\Pi_1{}^{(T)}, \ldots, \Pi_n{}^{(T)})\}(B) \tag{1}$$

for all sub-σ-algebras T of S, for all $B \in T$, and for all $\Pi_1, \ldots, \Pi_n \in P(S)$.

The left of (1) is the probability of B under the marginalized consensus distribution, and the right is the probability of B under the consensus distribution formed from the individual marginal distributions $\Pi_1{}^{(T)}, \ldots, \Pi_n{}^{(T)}$.

By the definition of $\{C_S(\Pi_1, \ldots, \Pi_n)^{(T)}\}$, the left of (1) is equal to $\{C_S(\Pi_1, \ldots, \Pi_n)\}(B)$, so MP is equivalent to saying that for all $S \in \Sigma$,

$$\{C_S(\Pi_1, \ldots, \Pi_n)\}(B) = \{C_T(\Pi_1{}^{(T)}, \ldots, \Pi_n{}^{(T)})\}(B) \tag{2}$$

for all $T \subseteq S$, $B \in T$, and $\Pi_1, \ldots, \Pi_n \in P(S)$.

In general, given distributions $\Pi_1, \ldots, \Pi_n \in P(S)$, the consensus probability of an event $A \in S$ will depend on S and on the whole of the individual assessors' distributions, and thus on what happens outside the set A. Clearly, however, it would be convenient if this were not true, and the consensus probability of A depended only on the probabilities the individual assessors gave to A, and possibly on the identity of the event A itself. Thus we want the probability of A to be a function only of $(A, \Pi_1(A), \ldots, \Pi_n(A))$. Since we are supposing that $\mathscr{C}$ contains consensus functions for all σ-algebras of Ω, so that A can be any subset of Ω, it is clear that the domain of this function must be the set

$$Q = [(2^\Omega \backslash \{\phi, \Omega\} \times [0, 1]^n] \cup \{(\phi, 0, \ldots, 0), (\Omega, 1, \ldots, 1)\}.$$

We say that a class of consensus functions has the weak setwise function property (WSFP) if it is well behaved in the sense just described.

Weak Setwise Function Property (WSFP). A class $\mathscr{C}$ of consensus functions has this property if and only if there exists a function $F\colon Q \to [0, 1]$ such that, for all $S \in \Sigma$,

$$\{C_S(\Pi_1, \ldots, \Pi_n)\}(A) = F\{A, \Pi_1(A), \ldots, \Pi_n(A)\} \tag{3}$$

for all $A \in S$ and all $\Pi_1, \ldots, \Pi_n \in P(S)$.

We now prove that WSFP is equivalent to MP.

Theorem 3.1. A class $\mathscr{C}$ of consensus functions has the WSFP if and only if it has the MP.

Proof. (a) WSFP $\Rightarrow$ MP. Follows from the definition of the properties.

(b) MP $\Rightarrow$ WSFP. Suppose $\mathscr{C}$ satisfies MP. Let A be any subset of Ω. We must show that if $S \in \Sigma$ contains A, then the consensus probability $\{C_S(\Pi_1, \ldots, \Pi_n)\}(A)$ of A depends only on A and $\Pi_1(A), \ldots, \Pi_n(A)$.

If $A \neq \phi$ or Ω, the result is trivial, since the range of C_S is $P(S)$.

Now suppose A is a proper subset of Ω. We can define a σ-algebra $\sigma(A) = \{\phi, A, \bar{A}, \Omega\}$ where $\bar{A}$ is the complement of A in Ω. Any $S \in \Sigma$ contains A, has $\sigma(A)$ as a sub-σ-algebra, and so by the MP

$$\{C_S(\Pi_1, \ldots, \Pi_n)\}(A) = \{C_{\sigma(A)}(\Pi_1{}^{(\sigma(A))}, \ldots, \Pi_n{}^{(\sigma(A))})\}(A) \tag{4}$$

using (2). But $\sigma(A)$ is uniquely defined by A, and any distribution over $\sigma(A)$ is uniquely defined by giving the probability of A under the distribution. Thus the right of

(4) is defined by A and by

$$\Pi_i^{(\sigma(A))}(A) = \Pi_i(A)$$

for $i = 1, \ldots, n$, which completes the proof that $\mathscr{C}$ satisfies WSFP.

Now although a class of consensus functions satisfying WSFP is clearly much simpler than one that does not, an important further simplification would occur if we could remove the dependence of the consensus probability of A on A itself, so that it depends only on the probabilities given to A by the individual assessors. We shall thus be interested in classes of consensus functions with the following property.

Strong Setwise Function Property (SSFP). A class $\mathscr{C}$ of consensus functions has the strong setwise function property if and only if there exists a function G: $[0, 1]^n \rightarrow [0, 1]$ such that for all $S \varepsilon \Sigma$

$$\{C_S(\Pi_1, \ldots, \Pi_n)\}(A) = G\{\Pi_1(A), \ldots, \Pi_n(A)\} \quad (5)$$

for all $A \varepsilon S$ and all $\Pi_1, \ldots, \Pi_n \varepsilon P(S)$.

Now it is fairly clear that MP is not equivalent to SSFP, and a counter example to this equivalence can be constructed by essentially taking the consensus distribution to be concentrated on a fixed point in Ω, independent of $\Pi_1, \ldots, \Pi_n$. If we impose, however, another mild condition in addition to MP, we do obtain such an equivalence. If all the assessors consider some particular event to have probability zero, it seems reasonable that the event should also have probability zero under the consensus distribution. A class of consensus functions for which this holds will be characterized as follows.

Zero Probability Property (ZPP). A class $\mathscr{C}$ of consensus functions is said to have the zero probability property if and only if, for all $S \varepsilon \Sigma$, for all $\Pi_1, \ldots, \Pi_n \varepsilon P(S)$, and for all $A \varepsilon S$, $\Pi_1(A) = \ldots = \Pi_n(A) = 0 \Rightarrow \{C_S(\Pi_1, \ldots, \Pi_n)\}(A) = 0$.

It is straightforward to show that MP and ZPP are independent. We now restrict our attention to spaces Ω which contain at least three distinct points. (If Σ does not contain all σ-algebras on σ, we suppose the maximal σ-algebra $S_{\max}$ in Σ contains at least three nonempty disjoint events.)

Theorem 3.2. If there are at least three distinct points in Ω, then for a class $\mathscr{C}$ of consensus functions, the following are equivalent: (a) $\mathscr{C}$ satisfies MP and ZPP, and (b) $\mathscr{C}$ satisfies SSFP.

Proof.

I *SSFP $\Rightarrow$ MP and ZPP.* Suppose $\mathscr{C}$ satisfies SSFP so that there exists a G as in (5); then $\Pi_1(A) = \ldots = \Pi_n(A) = 0$

$$\Rightarrow \{C_S(\Pi_1, \ldots, \Pi_n)\}(A) = G(0, \ldots, 0)$$

which is zero by consideration of the consensus probability of the empty set. So $\mathscr{C}$ satisfies ZPP. Also SSFP $\Rightarrow$ WSFP $\Rightarrow$ MP by Theorem 3.1.

II *MP and ZPP $\Rightarrow$ SSFP.* Suppose $\mathscr{C}$ satisfies MP and ZPP. By Theorem 3.1, $\mathscr{C}$ satisfies WSFP, so there exists F: $Q \rightarrow [0, 1]$ such that (3) holds. Since $\mathscr{C}$ satisfies ZPP, we also have that

$$F(A, 0, \ldots, 0) = 0 \quad (6)$$

for all $A \subset \Omega$.

Now suppose B, C are proper subsets of Ω with $B \cup C \neq \Omega$, and $B \cap C = \phi$, and $b_1, \ldots, b_n, c_1, \ldots, c_n \varepsilon [0, 1]$ with $b_i + c_i \leq 1$ for all i. We can choose distributions $\Pi_1, \ldots, \Pi_n$ over some σ-algebra containing both B and C, such that $\Pi_i(B) = b_i$ and $\Pi_i(C) = c_i$ for $i = 1, \ldots, n$. Since B and C are disjoint, it follows that $\Pi_i(B \cup C) = b_i + c_i$ and also that

$$\text{Consensus pr}(B) + \text{Consensus pr}(C) = \text{Consensus pr}(B \cup C).$$

It follows that

$$F(B, b_1, \ldots, b_n) + F(C, c_1, \ldots, c_n) = F(B \cup C, b_1 + c_1, \ldots, b_n + c_n). \quad (7)$$

Putting $c_1 = \ldots = c_n = 0$ in (7) and using (6), we obtain

$$F(B, b_1, \ldots, b_n) = F(B \cup C, b_1, \ldots, b_n). \quad (8)$$

Note that (8) also holds, trivially, if $C = \phi$. So we may say equivalently that if $\phi = D \subseteq E \neq \Omega$, and $b_1, \ldots, b_n \varepsilon [0, 1]$,

$$F(D, b_1, \ldots, b_n) = F(E, b_1, \ldots, b_n). \quad (9)$$

We now show that if A_1, A_2 are any subsets of Ω, $F(A, b_1, \ldots, b_n) = F(A_2, b_2, \ldots, b_n)$ whenever both arguments are in Q, so that $F(A, b_1, \ldots, b_n)$ does not depend on A and SSFP holds.

Suppose first A_1, A_2 are both proper.

Case 1. $A_1 \cap A_2 = B \neq \phi$. By (9) $F(A_1, b_1, \ldots, b_n) = F(B, b_1, \ldots, b_n) = F(A_2, b_1, \ldots, b_n)$.

Case 2. $A_1 \cap A_2 = \phi$, $A_1 \cup A_2 \neq \Omega$. By (8) $F(A_1, b_1, \ldots, b_n) = F(A_1 \cup A_2, b_1, \ldots, b_n) = F(A_2, b_1, \ldots, b_n)$.

Case 3. $A_1 \cap A_2 = \phi$, $A_1 \cup A_2 = \Omega$. Since there are at least three points in Ω, there exists a proper subset of at least one of A_1 and A_2. Without loss of generality, suppose A_1 has a proper subset D. Then $F(A_1, b_1, \ldots, b_n) = F(D, b_1, \ldots, b_n)$ by (9). But $A_2 \cap D = \phi$ and D is a subset of A_1, so $A_2 \cup D \neq \Omega$ and we can use the argument of Case 2 to deduce that $F(A_2, b_1, \ldots, b_n) = F(D, b_1, \ldots, b_n)$.

Hence, whenever A_1, A_2 are *proper* subsets of Ω, $F(A_1, b_1, \ldots, b_n) = F(A_2, b_1, \ldots, b_n)$ whenever both arguments are in Q. We complete the proof by noting that (6) implies that, for any proper A, $F(A, 0, \ldots, 0) = 0 = F(\phi, 0, \ldots, 0)$, and by applying (6) to $\bar{A}$ we obtain $F(A, 1, \ldots, 1) = 1 = F(\Omega, 1, \ldots, 1)$.

The above result does not hold in the case where Ω has only two points.

In the proof of Theorem 3.2, we used the fact that $C_S(\Pi_1, \ldots, \Pi_n)$ is a probability distribution. We now exploit elementary probabilistic arguments further to show that if the consensus probability of an event is a function of the individual assessors' probabilities for that event in the SSFP sense, then we can identify the form of the function G in (5). It must be a weighted average. This shows that if a class $\mathscr{C} = \{C_S\}$ of consensus functions has the SSFP, each C_S is a linear pool, and furthermore the weights in the pool do not depend on S. Again, this result holds only if Ω contains at least three points.

Theorem 3.3. If there are at least three distinct points in Ω, then for a class $\mathscr{C}$ of consensus functions the following are equivalent.

I $\mathscr{C}$ satisfies SSFP.

II There exist real numbers $\alpha_1, \ldots, \alpha_n$, nonnegative and summing to 1, such that for all $S \in \Sigma$, all $A \in S$, and all $\Pi_1, \ldots, \Pi_n \in P(S)$

$$\{C_S(\Pi_1, \ldots, \Pi_n)\}(A) = \sum_{i=1}^{n} \alpha_i \Pi_i(A).$$

Proof. (II) $\Rightarrow$ (I) obvious. (I) $\Rightarrow$ (II).

Suppose $\mathscr{C}$ satisfies SSFP. Then there exists a function G as in (5). We require to prove that there exists $\alpha_1, \ldots, \alpha_n$, nonnegative and summing to 1, such that for all $x_1, \ldots, x_n \in [0, 1]$,

$$G(x_1, \ldots, x_n) = \sum_{1}^{n} \alpha_i x_i. \tag{10}$$

Let $a_i, b_i (i = 1, \ldots, n)$ be nonnegative real numbers, with $a_i + b_i \leq 1$ for $i = 1, \ldots, n$.

We can partition Ω into three disjoint proper subsets, say A_1, A_2, and A_3. Let S be the minimal σ-algebra containing these subsets, and define probability distributions $\Pi_1, \ldots, \Pi_n$ on S by putting

$$\Pi_i(A_1) = a_i,$$
$$\Pi_i(A_2) = b_i, \quad \text{and}$$
$$\Pi_i(A_3) = 1 - a_i - b_i \quad \text{for } i = 1, \ldots, n.$$

Then $\Pi_i(A_1 \cup A_2) = a_i + b_i$, and the consensus probability of $A_1 \cup A_2$ is the sum of those of A_1 and A_2, so that

$$\begin{aligned} G(a_1, \ldots, a_n) &+ G(b_1, \ldots, b_n) \\ &= G(a_1 + b_1, \ldots, a_n + b_n). \end{aligned} \tag{11}$$

Now repeatedly using (11), we have for any $x_1, \ldots, x_n \in [0, 1]$

$$\begin{aligned} G(x_1, \ldots, x_n) &= G(x_1, 0, \ldots, 0) + G(0, x_2, \ldots, x_n) \\ &= G(x_1, 0, \ldots, 0) \\ &\quad + G(0, x_2, 0, \ldots, 0) + \ldots \\ &\quad + G(\underbrace{0, \ldots, 0}_{i-1}), x_i, 0, \ldots, 0) + \ldots \\ &\quad + G(0, \ldots, 0, x_n), \end{aligned}$$

or, putting $G_i(x) = G(\underbrace{0, \ldots, 0}_{i-1}, x_i, 0, \ldots, 0)$ for $i = 1, \ldots, n$, we have $G(x_1, \ldots, x_n) = \sum_{1}^{n} G_i(x_i)$. Also, (11) implies that $G_i(a + b) = G_i(a) + G_i(b)$ for $a, b \geq 0$, $a + b \in [0, 1]$, so that each G_i satisfies Cauchy's functional equation. Now the fact that G represents a probability implies that for all $x_1, \ldots, x_n \in [0, 1]$, $G(x_1, \ldots, x_n) \geq 0$, so that

$$G_i(x) = G(0, \ldots, 0, x, 0, \ldots, 0) \geq 0 \tag{12}$$

for $x \in [0, 1]$ and all i. Hence (Aczél 1966, Sec. 2.1.1.) $G_i(x) = \alpha_i x$ where α_i is a real constant. So (10) holds. Further, (12) implies that $\alpha_i \geq 0$ for all i.

It remains only to show that the α_i's sum to 1. This follows from putting $x_1 = \ldots = x_n = 1$ in (10) and noting that $G(1, \ldots, 1) = 1$. (Consider the consensus probability of Ω.)

Again, this theorem is not valid if Ω contains only two points.

From Theorem 3.2 and 3.3, we have, immediately, the following corollary.

Corollary 3.4. If there are at least three distinct points in Ω, then for a class $\mathscr{C}$ of consensus functions the following are equivalent.

I $\mathscr{C}$ satisfies MP and ZPP.

II Each $C \in \mathscr{C}$ is a linear opinion pool with weights independent of S.

Here we have a strong justification for using a linear opinion pool. If we want our consensus-finding process to commute with marginalization in the sense of MP, and we are further willing to impose the mild ZPP, then we must use a linear opinion pool. The restriction to spaces Ω with three or more points is hardly relevant to this argument, because if Ω has only two points, no nontrivial marginalization can be performed.

4. DISCUSSION

Important difficulties with the arguments here can arise from the fact that the consensus distribution is being formed as a function only of the individual assessors' distributions and the σ-algebra over which they are defined. Nothing has been said about how the consensus distribution should depend on the assessors' prior knowledge or probability assessment skills. We assumed that a consensus distribution was defined for all n tuples of assessors' distributions over any sub-σ-algebra of a σ-algebra $S_{\max}$ over the space Ω of states of nature. This assumption can be applied to several different situations.

First, one can assume that an appropriate class of consensus functions is to be found for a fixed probability assessment situation, with a fixed set of assessors who each have a fixed amount of prior information. Here it is clear that, for consensus finding to commute with marginalization, a linear opinion pool should be used. Prior information enters only through the values chosen for the weights. The justification for the marginalization assumption will be discussed.

Second, one could search for a method of combining distributions which applies to a fixed assessment task with *any* set of n assessors. Here, although the theorems still work, the assumption that the consensus functions do not depend on prior information is clearly unrealistic.

Third, one might require a method of combining distributions for a fixed set of assessors, but an arbitrary assessment situation. If we assume each assessor has a fixed amount of prior information, no theoretical problem arises. We can merely take Ω to be the set of all possible states of nature we might wish to consider, so that any σ-algebra over which distributions may be assessed is a σ-algebra over Ω. Then, if we impose the marginalization and zero probability properties (MP and ZPP) on our class of consensus functions, we must in every assessment situation use the same linear opinion pool. This implies that, for a fixed set of assessors with fixed prior information, we should not allow the weights in a linear opinion pool to depend on the assessors' relative amount of knowledge about a particular assessment problem.

If the assessors are experts in differing fields and hence have disparate prior knowledge, and we wish to consider a wide range of assessment situations, this may well be unreasonable. This amounts to saying that MP or ZPP is not reasonable here, and in fact MP seems intuitively unreasonable if applied to a marginalization from a multivariate distribution over a wide range of variables to a marginal distribution for a variable on which only one assessor has specialized knowledge.

More insidious is the case where the same group of assessors is asked to assess different distributions over the same σ-algebra, say unconditionally and then conditionally on some event. There is, of course, no reason why the two sets of distributions should be combined using the same consensus function. If they are, Dalkey (1972, 1975) has shown that serious difficulties arise. He assumed, essentially, our Strong Setwise Function Property (SSFP) and tried to find a consensus function that would be used to combine the assessed probabilities of A and of B given A, where A and B are two events. He obtained an impossibility theorem that showed that the only consensus function satisfying these assumptions and the probability axioms is the "dictatorship function," which takes a single assessor's probabilities as the consensus.

Whether this result causes any difficulty in our framework depends on the interpretation of the domain of the consensus functions. If one assumes that a particular class of consensus functions applies only to distributions conditional on a fixed amount of knowledge, then one is not necessarily justified in using the same function for the assessment of $\mathrm{pr}(A)$ and $\mathrm{pr}(B \mid A)$, and all Dalkey's theorem tells us is that we must not use the same G if we wish to avoid dictatorship. In fact if we impose MP and ZPP to a space Ω including both A and B, we should use linear opinion pools with the same weights to find the consensus probabilities of (A and B) and of A, and it can then be argued that we should define the consensus probability of B given A to be the ratio of these two consensus probabilities, provided they are not both zero. In fact, if one does this it is easily shown that the standard rule for weight-revision in linear opinion pools after conditioning (Roberts 1965, Raiffa 1968, Ch. 8) applies here. That is, if we use a linear opinion pool with weights $\alpha_1, \ldots, \alpha_n$ for finding the unconditional consensus probabilities, then we should use a linear opinion pool with weights $\beta_1, \ldots, \beta_n$ to find the consensus probability of B given A, where

$$\beta_i = \alpha_i \Pi_i(A)/\{\sum_i \alpha_i \Pi_i(A)\}$$

and $\Pi_i(A)$ is the ith assessor's probability for A. However, if one does for some reason require a consensus-finding method that works for any set of probability assessments conditional on anything, Dalkey's theorem applies. Dalkey's advice is that one should treat the consensus probability as a kind of pseudoprobability for which the multiplication rule does not apply. Probably a better approach is to weaken Dalkey's assumption, which is similar to SSFP, that the consensus probability of an event is a function only of the individual assessor's probabilities for that event.

[*Received April 1980. Revised November 1980.*]

REFERENCES

ACZÉL, J. (1966), *Lectures on Functional Equations and Their Applications*, New York: Academic Press.

BACHARACH, M. (1974), "Bayesian Dialogues," unpublished manuscript, Christ Church, Oxford.

DALKEY, N.C. (1972), "An Impossibility Theorem for Group Probability Functions," Santa Monica, Calif.: RAND Corporation, P-4862.

——— (1975), "Toward a Theory of Group Estimation," in *The Delphi Method: Techniques and Applications*, eds. H.A. Linstone and M. Turoff, Reading, Mass.: Addison-Wesley.

HOGARTH, R.M. (1977), "Methods for Aggregating Opinions," in *Decision Making and Change in Human Affairs*, eds. H. Jungermann and G. DeZeeuw, Dordrecht: D. Reidel.

MADANSKY, A. (1978), "Externally Bayesian Groups" (rev.), unpublished paper, University of Chicago, Graduate School of Business.

RAIFFA, H. (1968), "Decision Analysis: Introductory Lectures on Choices Under Uncertainty," Reading, Mass.: Addison-Wesley.

ROBERTS, H.V. (1965), "Probabilistic Prediction," *Journal of the American Statistical Association*, 60, 50–62.

STAËL VON HOLSTEIN, C.-A.S. (1970), *Assessment and Evaluation of Subjective Probability Distributions*, Stockholm: Economics Research Institute at the Stockholm School of Economics.

STONE, M. (1961), "The Opinion Pool," *Annals of Mathematical Statistics*, 32, 1339–1342.

WINKLER, R.L. (1968), "The Consensus of Subjective Probability Distributions," *Management Sciences*, 15, B61–B75.

——— (1971), "Probabilistic Prediction: Some Experimental Results," *Journal of the American Statistical Association*, 66, 675–685.