Probabilistic Opinion Pooling

Franz Dietrich
Paris School of Economics / CNRS
& University of East Anglia

Christian List
London School of Economics

2 March 2014, final version 13 October 2014

Forthcoming in A. Hajek and C. Hitchcock (eds.), Oxford Handbook of Philosophy and Probability, Oxford University Press

1 Introduction

How can several individuals' opinions on some factual matters be aggregated into unified collective opinions? This question arises in many contexts. A panel of climate experts may have to aggregate the panelists' conflicting opinions into a compromise view, in order to deliver a report to the government. A jury may have to arrive at a collective opinion on the facts of a case, despite disagreements between the jurors, so that the court can reach a verdict. In Bayesian statistics, we may wish to specify some all-things-considered prior probabilities by aggregating the subjective prior probabilities of different statisticians. In meta-statistics, we may wish to aggregate the probability estimates that different statistical studies have produced for the same events. An individual agent may wish to combine his or her own opinions with those of another, so as to resolve any peer disagreements. Finally, in a purely intra-personal case, an agent may seek to reconcile different 'selves' by aggregating their conflicting opinions on the safety of mountaineering, in order to decide whether to undertake a mountain hike and which equipment to buy.

How *should* opinions be aggregated in each of these cases? Perhaps surprisingly, this question has no obvious answer. Of course, if there is unanimity on the facts in question, we can simply take the unanimous opinions as the collective ones. But as soon as there are even mild disagreements, the aggregation problem becomes non-trivial. The aim of this article is to review and assess some salient proposed solutions.

Our focus will be on the aggregation of *probabilistic* opinions, which is often called the problem of *opinion pooling*. For present purposes, the opinions take the form of assignments of probabilities to some events or propositions of interest. Suppose, for instance, our climate panel consists of three experts, who assign the probabilities 0.3, 0.5, and 0.7 to the event that the global temperature will rise by more than one degree Celsius in the next 20 years. One proposal is to compute the linear average of these probabilities, so that the collective probability of the event is $\frac{1}{3}0.3 + \frac{1}{3}0.5 + \frac{1}{3}0.7 = 0.5$. Another proposal is to compute a weighted linear average of the form $w_1 0.3 + w_2 0.5 + w_3 0.7$, where w_1, w_2, and w_3 are non-negative weights whose sum-total is 1. Each expert's weight could reflect his or her competence, so that more competent experts have greater influence on the collective opinions. If expert 1 is deemed more competent than experts 2 and 3, then w_1 may be closer to 1, while w_2 and w_3 may be closer to 0. (In the special case of equal weights, we speak of an *unweighted* average.) A third proposal is to compute a geometric, rather than linear, average of the individuals' probabilities, which could also be weighted or unweighted. Generally, a *pooling function*, defined formally below, is a function from individual to collective probability assignments. Clearly, we can define many different pooling functions; the linear, weighted linear, and geometric functions are just illustrations.

Which pooling function is appropriate depends on the context and the intended status of the collective opinions. At least three questions are relevant:

- Should the collective opinions represent a *compromise* or a *consensus*? In the first case, each individual may keep his or her own personal opinions and adopt the collective opinions only hypothetically when representing the group or acting on behalf of it. In the second case, all individuals are supposed to take on the collective opinions as their own, so that the aggregation process can be viewed as a consensus formation process.
- Should the collective opinions be justified on *epistemic* or *procedural* grounds? In the first case, the pooling function should generate collective opinions that are epistemically well-justified: they should 'reflect the relevant evidence' or 'track the truth', for example. In the second case, the collective opinions should be a fair representation of the individual opinions. The contrast between the two approaches becomes apparent when different individuals have different levels of competence, so that some individuals' opinions are more reliable than others'. The epistemic approach then suggests that the collective opinions should depend primarily on the opinions of the more competent individuals, while the procedural approach might require that all individuals be given equal weight.
- Are the individuals' opinions based only on *shared information* or also on *private information*? This, in turn, may depend on whether the group has deliberated about the subject matter before opinions are aggregated. Group deliberation may influence individual opinions as the individuals learn new information and become aware of new aspects of the issue. It may help remove interpersonal asymmetries in information and awareness.

As we will see, *linear pooling* (the weighted or unweighted linear averaging of probabilities) can be justified on procedural grounds but not on epistemic ones, despite the possibility of giving greater weight to more competent individuals. Epistemic considerations support two other pooling methods: *geometric pooling* (the weighted or unweighted geometric averaging of probabilities), and *multiplicative pooling* (where probabilities are multiplied rather than averaged). The choice between geometric and multiplicative pooling, in turn, depends on whether the individuals' opinions are based on shared information or on private information.

After setting the stage in Sections 2 and 3, we discuss linear pooling in Sections 4 and 5, geometric pooling in Sections 6 and 7, and multiplicative pooling in Sections 8 and 9. We give an axiomatic characterization of each class of pooling functions and assess its plausibility. The characterizations are well-known in the case of linear and geometric pooling, but – to the best of our knowledge – new in the case of multiplicative pooling. In Section 10, we briefly mention some further approaches to opinion pooling: *supra-Bayesian pooling* (a radically Bayesian approach), the *aggregation of imprecise or qualitative probabilities*, the aggregation of binary yes/no opinions, known as *judgment aggregation*, and some generalized kinds of opinion aggregation.

There is a growing interdisciplinary literature on probabilistic opinion pooling; some references are given below (for a classic review, see Genest and Zidek 1986). While a complete review of the literature is beyond the scope of this article, we aim to give a flavour of the variety of possible approaches. We will discuss what we take to be the main arguments for and against the three central approaches we are focusing on: linear, geometric, and multiplicative pooling. As we will argue, these approaches promote different goals and rest on different assumptions.

2 The problem of probabilistic opinion pooling

We consider a group of $n \geq 2$ individuals, labelled $i = 1, ..., n$, who have to assign probabilities to some events.

The agenda. The *agenda* is the set of events under consideration. We define events as sets of possible worlds. Formally, consider a fixed non-empty set Ω of *possible worlds* (sometimes also called *possible states*). We take Ω to be finite for simplicity (but almost everything we say could be generalized to the infinite case). An *event* is a subset A of Ω; it can also be interpreted as a *proposition*. The complement of any event A is denoted $A^c = \Omega \backslash A$ and can be interpreted as its *negation*. For any two events A and B, the intersection $A \cap B$ can be interpreted as their *conjunction*, and the union $A \cup B$ as their *disjunction*. The events Ω (the entire set) and $\varnothing$ (the empty set) represent the *tautology* and the *contradiction*, respectively. All other events are *contingent*. For present purposes, the agenda is simply the set of all possible events, formally the power set of Ω (the set of all subsets of Ω), denoted $2^\Omega = \{A : A \subseteq \Omega\}$.

The simplest non-trivial example is a set of two worlds, $\Omega = \{\omega, \omega'\}$. Here, the agenda contains only two contingent events, namely $\{\omega\}$ and $\{\omega'\}$, e.g., 'rain' and 'no rain'. Obviously, the agenda grows exponentially in the size of Ω.[1]

[1]While we here take the agenda to consist of all possible events $A \subseteq \Omega$ (so that it is always closed under the Boolean operations of conjunction, disjunction, and negation), this classical

A concrete agenda. As an illustration (from Dietrich and List 2013a,b), consider an expert committee that seeks to form collective opinions about climate change. Possible worlds are vectors (j, k, l) of three characteristics, which may each take the value 0 or 1:

- The first characteristic specifies whether greenhouse gas concentrations exceed some critical threshold $(j = 1)$ or not $(j = 0)$.
- The second characteristic specifies whether there is causal law by which greenhouse gas concentrations above the threshold cause Arctic summers to be ice-free $(k = 1)$ or not $(k = 0)$.
- The third characteristic specifies whether Arctic summers are ice-free $(l = 1)$ or not $(l = 0)$.

Formally, the set of possible worlds is

$$\Omega = \{(1, 1, 1), (1, 0, 1), (1, 0, 0), (0, 1, 1), (0, 1, 0), (0, 0, 1), (0, 0, 0)\}.$$

This is the set of all triples of 0s and 1s with the exception of $(1, 1, 0)$. The latter triple is excluded because it represents an inconsistent combination of characteristics. The expert committee must assign a collective probability to every event $A \subseteq \Omega$.

The opinions. Opinions are represented by *probability functions*. A *probability function* P assigns to each event $A \subseteq \Omega$ a real number $P(A)$ in $[0, 1]$ such that

- the tautology has probability one: $P(\Omega) = 1$; and
- P is additive: $P(A \cup B) = P(A) + P(B)$ whenever two events A and B are mutually inconsistent, i.e., $A \cap B = \varnothing$.

The probability of a singleton event $\{\omega\}$ is often denoted $P(\omega)$ rather than $P(\{\omega\})$. Clearly, the probability of any event A can be written as the sum $P(A) = \sum_{\omega \in A} P(\omega)$. Thus a probability function P is fully determined by the probabilities $P(\omega)$ of the different worlds ω in Ω. Let

- $\mathcal{P}$ be the set of all probability functions P, and
- $\mathcal{P}'$ be the set of all probability functions P which are *regular*, i.e., $P(\omega) > 0$ for all worlds ω.

Opinion pooling. A combination of probability functions across the n individuals, $(P_1, ..., P_n)$, is called an *opinion profile*. A *pooling function* takes opinion profiles as input and produces collective probability functions as output. Formally, it is a function, F, which maps each opinion profile $(P_1, ..., P_n)$ within some domain of admissible profiles to a single probability function $P_{P_1,...,P_n} = F(P_1, ..., P_n)$. The

but demanding assumption is dropped in Dietrich and List (2013a), where the agenda is not required to be closed under Boolean operations.

notation $P_{P_1,...,P_n}$ indicates that the collective probability function depends on the individual probability functions $P_1, ..., P_n$.

Some pooling functions are defined on the domain of all logically possible opinion profiles, others on a more restricted domain, such as the domain of all profiles of *regular* probability functions. In the first case, F is a function from $\mathcal{P}^n$ to $\mathcal{P}$; in the second case, a function from $\mathcal{P}'^n$ to $\mathcal{P}$. (As is standard, for any set S, we write S^n to denote the set of all n-tuples consisting of elements of S.)

The linear example. The best-known example is a *linear* pooling function, which goes back to Stone (1961) or even Laplace.[2] Here, each opinion profile $(P_1, ..., P_n)$ in the domain $\mathcal{P}^n$ is mapped to the collective probability function satisfying

$$P_{P_1,...,P_n}(A) = w_1 P_1(A) + \cdots + w_n P_n(A) \text{ for every event } A \subseteq \Omega,$$

where $w_1, ..., w_n$ are fixed non-negative weights with sum-total 1. The class of linear pooling functions includes a variety of functions, ranging from linear averaging with equal weights, where $w_i = \frac{1}{n}$ for all i, to an 'expert rule' or 'dictatorship', where $w_i = 1$ for one individual and $w_j = 0$ for everyone else. In the latter case:

$$P_{P_1,...,P_n}(A) = P_i(A) \text{ for every event } A \subseteq \Omega.$$

3 The axiomatic method

As should be clear, there is an enormous number of logically possible pooling functions. Many are unattractive. For example, we would not normally want the collective probability of any event to depend *negatively* on the individual probabilities of that event. (An example of a negative dependence would be a case in which the individual probabilities for some event all go up, while the collective probability goes down, with all relevant other things remaining equal.) Similarly, we would not normally want the collective probabilities to depend only on the probabilities assigned by a single 'dictatorial' individual. How can we choose a good pooling function? Here, the *axiomatic method* comes into play. Under this method, we do not choose a particular pooling function directly, say linear pooling, but instead formulate general requirements on a 'good' pooling function – our *axioms* – and then ask which pooling functions, if any, satisfy them. One example is the axiom of unanimity preservation, which requires that if all individuals hold the same opinions, these opinions become the collective ones. This is satisfied by linear pooling functions, but also by many other pooling functions. So, this axiom does not single out a unique pooling function. However,

[2] For other early contributions, see Bacharach (1972) and DeGroot (1974).

if we add another axiom, as discussed below, we can narrow down the class of possible pooling functions to the class of linear pooling functions alone.

The axiomatic method can guide and structure our search for a good pooling function. The difficult question of which pooling function to use is re-cast as the more tractable question of which axioms to impose. This allows us to assess different axioms one by one rather than having to assess a fully specified pooling function in one go.

Generally, once we have specified a set of axioms, we will be faced with one of three possible situations:

(1) *Exactly one* pooling function – or one salient class of pooling functions – satisfies all our axioms, in which case we have successfully completed our search for a pooling function.
(2) *Several* pooling functions – or even several structurally different classes of pooling functions – satisfy all our axioms. This is a case of *underdetermination*, in which we may wish to impose further axioms.
(3) *No* pooling function satisfies all our axioms. This is a case of *overdetermination*, in which we may have to relax at least one axiom.

4 Linear pooling: the eventwise independent approach

Which axioms characterize the class of linear pooling functions? Aczél and Wagner (1980) and McConway (1981) give an elegant answer to this question, identifying two jointly necessary and sufficient axioms: *eventwise independence* and *unanimity preservation*.

The first, *eventwise independence* (or simply *independence*), requires that the collective probability of any event depend solely on the individual probabilities of *that* event.[3] This reflects the democratic idea that the collective opinion on any issue should be determined by individual opinions on that issue. The underlying picture of democracy is a *non-holistic* one, under which the collective opinion on any issue must not be influenced by individual opinions on other issues.

Independence. For each event $A \in X$, there exists a function $D_A : [0,1]^n \to [0,1]$, called the *local pooling criterion* for A, such that

$$P_{P_1,...,P_n}(A) = D_A(P_1(A), ..., P_n(A))$$

for every opinion profile $(P_1, ..., P_n)$ in the domain of the pooling function.

[3] This axiom is also known as the *weak setwise function property* or *weak label neutrality.*

Each local pooling criterion D_A aggregates any combination of probabilities $(x_1, ..., x_n)$ on a specific event into a single collective probability $D_A(x_1, ..., x_n)$. In the case of a linear pooling function, the local pooling criterion for any event A is simply $D_A = D$, with

$$D(x_1, ..., x_n) = w_1 x_1 + \cdots + w_n x_n,$$

where $w_1, w_2, ..., w_n$ are the weights of the n individuals.

The second axiom, *unanimity preservation*, requires that if all individuals hold the same opinions, these opinions become the collective ones:

Unanimity preservation. For every opinion profile $(P_1, ..., P_n)$ in the domain of the pooling function, if all P_i are identical, then $P_{P_1,...,P_n}$ is identical to them.

This axiom seems very compelling, especially from the procedural perspective of making collective probabilities responsive to individual probabilities. Surprisingly, however, the axiom may be problematic from an epistemic perspective (see Section 7), but for now we do not question it.

Theorem 1. (Aczél and Wagner 1980; McConway 1981) Suppose $|\Omega| > 2$. The linear pooling functions are the only independent and unanimity-preserving pooling functions (with domain $\mathcal{P}^n$).[4]

This result is surprising, because eventwise independence seems, at first, to leave a great degree of freedom in the specification of the local pooling criteria D_A. In conjunction with unanimity preservation, however, independence becomes quite restrictive. First, each local pooling criterion D_A must then be a linear averaging criterion. Second, the local pooling criteria D_A must be the same for all events A. This precludes defining the collective probability for any event A as the weighted average

$$P_{P_1,...,P_n}(A) = D_A(P_1(A), ..., P_n(A)) = w_1^A P_1(A) + \cdots + w_n^A P_n(A), \tag{1}$$

where an individual i may have different weights w_i^A for different events A. One might consider such event-dependent weights plausible, because an individual need not be equally good at estimating the probabilities of different events. Ideally, one

[4] To be precise, Aczél and Wagner (1980) and McConway (1981) use another, logically independent unanimity axiom, called *zero preservation*: if some event is assigned zero probability by each individual, then it is assigned zero probability collectively. As another alternative, one could use the following axiom, which weakens both of these conditions: if some *world* ω is assigned probability 1 by every individual (so that everyone holds the same degenerate probability function), then ω is assigned probability 1 collectively. Other axiomatic characterizations of linear pooling are given by Mongin (1995) and Chambers (2007). See also Lehrer and Wagner (1981), who use linear opinion pooling to build a theory of consensus formation in groups.

might wish to give each individual a greater weight in determining the collective probability for events within his or her area of expertise than for events outside that area. Unfortunately, formula (1) does not guarantee a well-defined collective probability function unless each individual's weight w_i^A is the same for all events A ($\neq \Omega, \varnothing$), as in standard linear pooling. In particular, if weights vary across events, the function defined in (1) can violate additivity.

What can be said in defence of eventwise independence? There are at least two pragmatic arguments for it. First, eventwise independent aggregation is easy to implement, because it permits the subdivision of a complex aggregation problem into multiple simpler ones, each focusing on a single event. Our climate panel can first consider the event that greenhouse gas concentrations exceed some critical threshold and aggregate individual probabilities for that event; then do the same for the second event; and so on. Second, eventwise independent aggregation is invulnerable to agenda manipulation. If the collective opinion about each event depends only on the individual opinions about that event, then an agenda setter who might wish to influence the outcome of the aggregation will not be able to change the collective opinion about any event by adding further events to the agenda or removing others from it. For instance, the agenda setter could not affect the collective probability for the event 'snow' by adding the event 'hail' to the agenda.[5] McConway (1981) proves that eventwise independence is equivalent to the requirement that collective opinions be invariant under changes in the specification of the agenda; see also Genest (1984b).[6]

5 The limitations of eventwise independent aggregation

There are a number of objections to eventwise independence and consequently to linear pooling. First, it is questionable whether eventwise independent aggregation can be justified epistemically. The collective opinions it generates may not adequately incorporate the information on which individual opinions are based. As we will see in Sections 6 to 9, some axioms that capture the idea of 'adequately incorporating information' – namely the axioms of *external Bayesianity* and *individualwise Bayesianity* – typically lead to pooling functions that violate eventwise independence.

[5] A change in the agenda would have to be represented mathematically by a change in the underlying set of worlds Ω. In order to add the event 'hail' to the agenda, each world ω in the original set Ω must be replaced by two worlds, ω_1 and ω_2, interpreted as ω combined with the occurrence of hail and ω combined with its non-occurrence, respectively.

[6] McConway captures this requirement by the so-called *marginalization property*, which requires aggregation to commute with the operation of reducing the relevant algebra (agenda) to a sub-algebra (sub-agenda); this reduction corresponds to the removal of events from the agenda.

Second, eventwise independence becomes implausible when this requirement is applied to 'artificial' composite events, such as conjunctions or disjunctions of intuitively unrelated events. There seems no reason, for example, why the collective probability for the disjunction 'snow *or* wind' should depend only on individual probabilities for that disjunction, rather than on individual probabilities for each disjunct. Except in trivial cases, the agenda will always contain some 'artificial' composite events, since it is closed under Boolean operations (conjunction, disjunction, and negation). (Eventwise independence may become more plausible if we relax this closure requirement on the agenda; see Dietrich and List 2013a.)

Finally, eventwise independence conflicts with the principle of *preserving probabilistic independence*. This requires that any two events that are uncorrelated according to every individual's probability function remain uncorrelated according to the collective probability function. For instance, if each climate expert took the events of high greenhouse gas concentrations and ice-free Arctic summers to be uncorrelated, then these two events should remain uncorrelated according to the collective probabilities. Unfortunately, as shown by Wagner (1984), eventwise independent pooling functions do not preserve probabilistic independence (setting aside degenerate pooling functions such as dictatorial ones).

In fairness, we should mention that the failure to preserve probabilistic independence can be held not just against eventwise independent pooling functions but against a much wider class of pooling functions (Genest and Wagner 1987). This includes all linear, geometric, and multiplicative pooling functions that are non-dictatorial. Further, the preservation of probabilistic independence is itself a normatively questionable requirement. Why, for example, should probabilistic independence judgments be preserved even when they are purely accidental, i.e., not driven by any insight into the causal connections between events? It is more plausible to require that only *structurally relevant* probabilistic independencies be preserved, i.e., those that are due to the structure of causal connections rather than being merely accidental. On the preservation of causally motivated probabilistic independencies, see Bradley, Dietrich, and List (2014).

6 Geometric pooling: the externally Bayesian approach

We now turn to a class of pooling functions based on geometric, rather than linear, averaging. While the linear average of n numbers, such as $x_1, x_2, ..., x_n$, is $\frac{x_1+x_2+...+x_n}{n}$, the geometric average is $\sqrt[n]{x_1 x_2 \cdots x_n} = x_1^{\frac{1}{n}} x_2^{\frac{1}{n}} \cdots x_n^{\frac{1}{n}}$. Just as a linear average can be generalized to take the weighted form $w_1 x_1 + w_2 x_2 + ... + w_n x_n$, so a geometric average can be generalized to take the weighted form $x_1^{w_1} x_2^{w_2} \cdots x_n^{w_n}$, where $w_1, ..., w_n$ are non-negative weights with sum-total 1.

A geometric pooling function determines the collective probabilities in two steps. In the first step, it takes the collective probability of each possible world (rather than event) to be a geometric average of the individuals' probabilities of that world. In the second step, it renormalizes these collective probabilities in such a way that their sum-total becomes 1.

Formally, a pooling function is called *geometric* (or also *logarithmic*) if it maps each opinion profile $(P_1, ..., P_n)$ in the domain $\mathcal{P}'^n$ to the collective probability function satisfying

$$P_{P_1,...,P_n}(\omega) = c[P_1(\omega)]^{w_1} \cdots [P_n(\omega)]^{w_n} \text{ for every world } \omega \text{ in } \Omega,$$

where $w_1, ..., w_n$ are fixed non-negative weights with sum-total 1 and c is a normalization factor, given by

$$c = \frac{1}{\sum_{\omega' \in \Omega} [P_1(\omega')]^{w_1} \cdots [P_n(\omega')]^{w_n}}.$$

The sole point of the normalization factor c is to ensure that the sum-total of the collective probabilities across all worlds in Ω becomes 1. Two technical points are worth noting. First, geometric pooling functions are defined by specifying the collective probabilities of worlds, rather than events, but this is of course sufficient to determine the collective probabilities of all events. Second, to ensure well-definedness, the domain of a geometric pooling function must be $\mathcal{P}'^n$ rather than $\mathcal{P}^n$, admitting only *regular* individual probability functions as input.[7]

As in the case of linear pooling, geometric pooling functions can be weighted or unweighted, ranging from geometric averaging with equal weights, where $w_i = \frac{1}{n}$ for all i, to an 'expert rule' or 'dictatorship', where $w_i = 1$ for one individual and $w_j = 0$ for everyone else, so that $P_{P_1,...,P_n} = P_i$.

How can geometric pooling be justified? First, it is clearly unanimity-preserving. Second, unlike linear pooling, it is not eventwise independent (except in the limiting case of an expert rule or dictatorship). Intuitively, this is because the renormalization of probabilities introduces a holistic element.

However, geometric pooling satisfies another, epistemically motivated axiom, called *external Bayesianity* (proposed by Madansky 1964). This concerns the effects that informational updating has on individual and collective probability functions. Informally, the axiom requires that, if probabilities are to be updated based on some information, it should make no difference whether they are updated *before* aggregation or *after* aggregation. We should arrive at the same collective

[7]Without this restriction, it could happen that, for every world, *some* individual assigns a probability of zero to it, so that the geometric average of individual probabilities is zero for all worlds, a violation of probabilistic coherence. A similar remark applies to the definition of multiplicative pooling in the next section.

probability function irrespective of whether the individuals first update their probability functions and then aggregate them, or whether they first aggregate their probability functions and then update the resulting collective probability function, where the update is based on the same information.

To formalize this, we represent information by a *likelihood function*. This is a function L which assigns, to each world ω in Ω, a positive number $L(\omega)$, interpreted as the degree to which the information supports ω, or more precisely the likelihood that the information is true in world ω. In our climate-panel example, the information that a revolutionary carbon-capture-and-storage technology is in use may be expressed by a likelihood function L that takes lower values at worlds with high greenhouse gas concentrations than at worlds with low greenhouse gas concentrations. This is because the information is more likely to be true at worlds with low greenhouse gas concentrations than at worlds with high ones. (The revolutionary carbon-capture-and-storage technology would remove greenhouse gases from the atmosphere.)

What does it mean to update a probability function based on the likelihood function L? Suppose an agent initially holds the probability function P and now learns the information represented by L. Then the agent should adopt the new probability function P^L satisfying

$$P^L(\omega) := \frac{P(\omega)L(\omega)}{\sum_{\omega' \in \Omega} P(\omega')L(\omega')} \text{ for every world } \omega \text{ in } \Omega. \tag{2}$$

This definition can be motivated in Bayesian terms. For a simple illustration, consider a limiting case of a likelihood function L, where $L(\omega) = 1$ for all worlds ω within some event A and $L(\omega) = 0$ for all worlds ω outside A. Here L simply expresses the information that event A has occurred. (This is a limiting case of a likelihood function because L is not positive for all worlds ω, as required by our definition, but only non-negative.) Formula (2) then reduces to the familiar requirement that the agent's posterior probability function after learning that event A has occurred be equal to his or her prior probability function conditional on A. In the Appendix, we discuss the notion of a likelihood function and the Bayesian motivation for formula (2) in more detail.

The axiom of external Bayesianity can now be stated as follows:

External Bayesianity. For every opinion profile $(P_1, ..., P_n)$ in the domain of the pooling function and every likelihood function L (where the updated profile $(P_1^L, ..., P_n^L)$ remains in the domain), pooling and updating are commutative, i.e., $P_{P_1^L, ..., P_n^L} = P^L_{P_1, ..., P_n}$.

Theorem 2. (e.g. Genest 1984a) The geometric pooling functions are externally Bayesian and unanimity-preserving.

Let us briefly explain why a geometric pooling function (say with weights $w_1, ..., w_n$) is externally Bayesian. Without loss of generality, we can view any probability function as a function from the set Ω of worlds into $[0,1]$, rather than as a function from the set 2^Ω of events into $[0,1]$. Consider any opinion profile $(P_1, ..., P_n)$ (in the domain $\mathcal{P}'^n$) and any likelihood function L. To show that $P_{P_1^L,...,P_n^L} = P^L_{P_1,...,P_n}$, we observe that each side of this equation is proportional to the function $[P_1]^{w_1} \cdots [P_n]^{w_n} L$. (Since we are dealing with probability functions, proportionality then implies identity.) First, note that $P^L_{P_1,...,P_n}$ is proportional to this function by definition. Second, note that $P_{P_1^L,...,P_n^L}$ is proportional to the product of functions $[P_1^L]^{w_1} \cdots [P_n^L]^{w_n}$, also by definition. But, since each function P_i^L is proportional to the product $P_i L$, the product $[P_1^L]^{w_1} \cdots [P_n^L]^{w_n}$ is, in turn, proportional to the function

$$[P_1 L]^{w_1} \cdots [P_n L]^{w_n} = [P_1]^{w_1} \cdots [P_n]^{w_n} L^{w_1 + \cdots + w_n} = [P_1]^{w_1} \cdots [P_n]^{w_n} L,$$

as required.

Why is external Bayesianity a plausible requirement? If it is violated, the time at which an informational update occurs can influence the collective opinions. It will then matter whether the informational update takes place before or after individual opinions are aggregated. This would open the door to manipulation of the collective opinions by someone who strategically discloses a relevant piece of information at the right time. Of course, someone acting in this way need not have bad intentions; he or she might simply wish to 'improve' the collective opinions. Nonetheless, the need to decide whether $P_{P_1^L,...,P_n^L}$ or $P^L_{P_1,...,P_n}$ is a 'better' collective probability function raises all sorts of complications, which we can avoid if external Bayesianity is satisfied. See Raiffa (1968, pp. 221-226) for some examples of strategic information retention when external Bayesianity is violated.

Geometric pooling functions are not the only externally Bayesian and unanimity-preserving pooling functions. The two axioms are also compatible with a generalized form of geometric pooling, in which the weights $w_1, ..., w_n$ may depend on the opinion profile $(P_1, ..., P_n)$ in a systematic way.[8] Genest, McConway, and Schervish (1986) characterize all pooling functions satisfying the conditions of Theorem 2, or just external Bayesianity. Once some additional axioms are imposed, over and above those in Theorem 2, geometric pooling becomes unique (Genest 1984a; Genest, McConway, and Schervish 1986). However, the additional axioms are technical and arguably not independently compelling. So, we still lack a fully compelling axiomatic characterization of geometric pooling. For a further

[8] Let us write $w_i^{P_1,...,P_n}$ for individual i's weight when the profile is $(P_1, ..., P_n)$. In the profile-dependent specification of weights, all one needs to ensure is that, for all i, $w_i^{P_1,...,P_n} = w_i^{P'_1,...,P'_n}$ whenever the profile $(P'_1, ..., P'_n)$ is 'accessible via update' from the profile $(P_1, ..., P_n)$ in the sense that there is a likelihood function L such that $P_i^L = P'_i$ for every i. Accessibility via updates defines an equivalence relation between profiles in $\mathcal{P}'^n$. Since there are many equivalence classes (provided $|\Omega| > 1$), there are many generalized geometric pooling functions.

discussion and comparison of linear and geometric pooling, see Genest and Zidek (1986).

7 From symmetric to asymmetric information

Although we have justified geometric pooling in epistemic terms – by invoking the axiom of external Bayesianity – there are conditions under which geometric pooling is not epistemically justified. These conditions motivate another approach to opinion pooling, called *multiplicative pooling* (Dietrich 2010), which we introduce in the next section. To identify those conditions, we must consider not just the probability functions P_1, P_2, ..., P_n that are to be pooled, but their *informational bases*: the information that the individuals have used to arrive at them.

Let us contrast two diametrically opposed cases, setting aside any intermediate cases for simplicity. (We comment briefly on intermediate cases in Section 10.)

Case 1: informational symmetry. The individuals' probability functions P_1, ..., P_n are based on exactly the same information. Any differences in these probability functions stem at most from different ways of interpreting that shared information.

Case 2: informational asymmetry. The individuals' probability functions P_1, ..., P_n are based on different information, and there is no overlap between different individuals' information, apart from some fixed background information held by everyone. Each individual i's probability function P_i is derived from some prior probability function by conditionalizing on i's private information. That is, $P_i = p_i^{L_i}$, where p_i is i's prior probability function and L_i is the likelihood function representing i's private information. For simplicity, we assume a shared prior probability function $p_i = p$ for every individual i, which reflects the individuals' shared background information.

Case 1 might occur if there is group deliberation and exchange of information prior to the pooling of opinions. Case 2 might occur in the absence of such group deliberation or exchange of information. We will now show that the axioms by which we have justified geometric pooling – unanimity preservation and external Bayesianity – are plausible in Case 1, but not in Case 2.

Consider unanimity preservation. In Case 1, this axiom is compelling. If all individuals arrive at the same probability function $P_1 = ... = P_n$ based on shared information, there is no reason why this probability function should not also become the collective one. After all, in the present case, the individuals not only *have* the same information, as assumed in Case 1, but also *interpret* it in the same way; otherwise, we would not have $P_1 = ... = P_n$.

In Case 2, by contrast, unanimity preservation is not compelling. If all individuals arrive at the same probability function P_i based on different private information, the collective probability function ought to incorporate that dispersed information. Thus it should incorporate the individuals' likelihood functions $L_1, ..., L_n$, and this may, in turn, require a collective probability function distinct from $P_1 = ... = P_n$.[9] Suppose, for example, that all the experts on our climate panel assign the same high probability of 0.9 to the event that greenhouse gas concentrations exceed the critical threshold. Plausibly, if each expert has some private information that supports assigning a high probability to some event, compared to a much lower prior, then the totality of private information supports the assignment of an even higher probability to it. Thus the collective probability should not be the same as the individual ones, but amplified, above 0.9. Similarly, if all experts, prompted by their own independent evidence, assign the same low probability of 0.1 to some event, then the collective probability should be even lower. Here, the group knows more than each individual member.

Next consider external Bayesianity. In Case 1, where all individuals have the same information, this requirement is well motivated, as should be clear from our discussion in the last section. By contrast, in Case 2, where different individuals have different and non-overlapping private information, external Bayesianity loses its force. Recall that we justified the requirement that $P_{P_1^L,...,P_n^L} = P_{P_1,...,P_n}^L$ by interpreting L as representing information that is received by *all* individuals. In Case 2, however, individuals have only private information (apart from some shared but fixed background information, which cannot include the non-fixed information represented by L).[10] Here, updating all probability functions P_i would mean updating them *on the basis of different private information*. So, the updated profile $(P_1^L, ..., P_n^L)$ would have to be interpreted as expressing the individuals' opinions after incorporating *different* items of private information that happen to be represented by the same likelihood function L for each individual. This interpretation makes it implausible to require that $P_{P_1^L,...,P_n^L}$ and $P_{P_1,...,P_n}^L$ be the same. From the group's perspective, there is not just one item of information to take into account, but n separate such items. While each item of information by itself corresponds to the likelihood function L, the group's information as a whole corresponds to the product of n such functions, namely L^n. In the next section, we introduce an axiom that replaces external Bayesianity in Case 2.

[9] One may want to obtain the collective probability function $P_{P_1,...,P_n}$ by updating some prior probability function p in light of all n likelihood functions. Then $P_{P_1,...,P_n}$ equals $(...((p^{L_1})^{L_2})...)^{L_n}$, which in turn equals $p^{L_1 L_2 \cdots L_n}$, the probability function obtained by updating p in light of the likelihood function defined as the product $L_1 L_2 \cdots L_n$. This is, in effect, what multiplicative pooling does, as should become clear in the next section.

[10] The information represented by L is non-fixed, since it is present in one opinion profile, $(P_1^L, ..., P_n^L)$, and absent in another, $(P_1, ..., P_n)$.

8 Multiplicative pooling: the individualwise Bayesian approach

We now consider a class of pooling functions that are appropriate in Case 2, where the probability functions P_1, P_2, ..., P_n are based on different private information and there is at most some fixed background information held by all individuals. This is the class of multiplicative pooling functions (proposed by Dietrich 2010), which are based on multiplying, rather than averaging, probabilities.

A multiplicative pooling function, like a geometric one, determines the collective probabilities in two steps. In the first step, it takes the collective probability of each possible world to be the product of the individuals' probabilities of that world, calibrated by multiplication with some exogenously fixed probability (whose significance we discuss in Section 9). This differs from the first step of geometric pooling, where the geometric average of the individuals' probabilities is taken. In the second step, multiplicative pooling renormalizes the collective probabilities such that their sum-total becomes 1; this matches the second step of geometric pooling.

Formally, a pooling function is called *multiplicative* if it maps each opinion profile $(P_1, ..., P_n)$ in the domain $\mathcal{P}'^n$ to the collective probability function satisfying

$$P_{P_1,...,P_n}(\omega) = cP_0(\omega)P_1(\omega)\cdots P_n(\omega) \text{ for every world } \omega \text{ in } \Omega,$$

where P_0 is some fixed probability function, called the *calibrating function*, and c is a normalization factor, given by

$$c = \frac{1}{\sum_{\omega' \in \Omega} P_0(\omega')P_1(\omega')\cdots P_n(\omega')}.$$

As before, the point of the normalization factor c is to ensure that the sum-total of the collective probabilities across all worlds in Ω is 1. To see that multiplicative pooling can be justified in Case 2, we now introduce a new axiom that is plausible in that case – *individualwise Bayesianity* – and show that it is necessary and sufficient for multiplicative pooling. (The present characterization of multiplicative pooling is distinct from the one given in Dietrich 2010.)

The axiom says that it should make no difference whether some information is received by a single individual before opinions are pooled or by the group as a whole afterwards. More specifically, we should arrive at the same collective probability function irrespective of whether a single individual first updates his or her own probability function based on some private information and the probability functions are then aggregated, or whether the probability functions are first aggregated and then updated – now at the collective level – given the same information.

Individualwise Bayesianity. For every opinion profile $(P_1, ..., P_n)$ in the domain of the pooling function, every individual i, and every likelihood function L (where the profile $(P_1, ..., P_i^L, ..., P_n)$ remains in the domain), we have $P_{P_1,...,P_i^L,...,P_n} = P^L_{P_1,...,P_n}$.

Just as external Bayesianity was plausible in Case 1, where all individuals' probability functions are based on the same information, so individualwise Bayesianity is plausible in Case 2, where different individuals' probability functions are based on different private information. The argument for individualwise Bayesianity mirrors that for external Bayesianity: any violation of the axiom implies that it makes a difference whether someone acquires private information *before* opinions are pooled or acquires the information and shares it with the group *afterwards*. This would again generate opportunities for manipulation by third parties able to control the acquisition of information.

Theorem 3. The multiplicative pooling functions are the only individualwise Bayesian pooling functions (with domain $\mathcal{P}'^n$).

This (new) result has an intuitive proof, which we now give.

Proof: Let us again view any probability function as a function from the set Ω of worlds into $[0, 1]$, rather than as a function from the set 2^Ω of events into $[0, 1]$. As noted earlier, this is no loss of generality. We first prove that multiplicative pooling functions satisfy individualwise Bayesianity. Consider a multiplicative pooling function, for some exogenously fixed probability function P_0, which serves as the calibrating function. Note that, for any opinion profile $(P_1, ..., P_n)$,

- the function $P_{P_1,...,P_i^L,...,P_n}$ is by definition proportional to the product $P_0 P_1 \cdots (P_i L) \cdots P_n$, and
- the function $P^L_{P_1,...,P_n}$ is by definition proportional to the product $(P_0 P_1 \cdots P_n)L$.

These two products are obviously the same, so individualwise Bayesianity is satisfied. Conversely, we prove that no pooling functions other than multiplicative ones satisfy the axiom. Consider any pooling function with domain $\mathcal{P}'^n$ that satisfies individualwise Bayesianity. Let P^* be the uniform probability function, which assigns the same probability to every world in Ω. We show that our pooling function is multiplicative with calibrating function $P_0 = P_{P^*,...,P^*}$. Consider any opinion profile $(P_1, ..., P_n)$ (in $\mathcal{P}'^n$). The argument proceeds in n steps. It could be re-stated more formally as an inductive proof.

- *Step 1*: First, consider the likelihood function $L := P_1$. The function $P_{P_1,P^*,...,P^*}$ is equal to $P_{(P^*)^L,P^*,...,P^*}$. By individualwise Bayesianity, this is equal to $P^L_{P^*,...,P^*}$, which is in turn proportional to $P_{P^*,...,P^*} L = P_0 P_1$, by the definitions of P_0 and L.

- *Step 2*: Now, consider the likelihood function $L := P_2$. The function $P_{P_1,P_2,P^*,...,P^*}$ is equal to $P_{P_1,(P^*)^L,P^*,...,P^*}$. By individualwise Bayesianity, this is equal to $P^L_{P_1,P^*,...,P^*}$, which is in turn proportional to $P_{P_1,P^*,...,P^*}L$, i.e., to $P_0P_1P_2$, by Step 1 and the definition of L.

...

- *Step n*: Finally, consider the likelihood function $L := P_n$. The function $P_{P_1,...,P_n}$ is equal to $P_{P_1,...,P_{n-1},(P^*)^L}$. By individualwise Bayesianity, this is equal to $P^L_{P_1,...,P_{n-1},P^*}$, which is in turn proportional to $P_{P_1,...,P_{n-1},P^*}L$, i.e., to $P_0P_1 \cdots P_n$, by Step $n-1$ and the definition of L. ■

9 How to calibrate a multiplicative pooling function

Recall that the definition of a multiplicative pooling function involves a calibrating probability function P_0. The collective probability of each possible world is not merely the renormalized product of the individuals' probabilities of that world, but it is multiplied further by the probability that P_0 assigns to the world. How should we choose that calibrating probability function?

It is simplest to take P_0 to be the uniform probability function, which assigns the same probability to every world in Ω. In this case, we obtain the *simple multiplicative pooling function*, which maps each opinion profile $(P_1, ..., P_n)$ in $\mathcal{P}'^n$ to the collective probability function satisfying

$$P_{P_1,...,P_n}(\omega) = cP_1(\omega) \cdots P_n(\omega) \text{ for every world } \omega \text{ in } \Omega,$$

for a suitable normalization factor c.

The simple multiplicative pooling function is the only multiplicative pooling function that satisfies an additional axiom, which we call *indifference preservation*. It is a weak version of the unanimity-preservation axiom, which applies only in the special case in which every individual's probability function is the uniform one.

Indifference preservation. If every probability function in the opinion profile $(P_1, ..., P_n)$ is the uniform probability function, then the collective probability function $P_{P_1,...,P_n}$ is also the uniform one (assuming the profile is in the domain of the pooling function).

Corollary of Theorem 3. The simple multiplicative pooling function is the only individualwise Bayesian and indifference-preserving pooling function (with domain $\mathcal{P}'^n$).

When is indifference preservation plausible? We suggest that it is plausible if the individuals have no shared background information at all; all their information is private. Recall that we can view each individual i's probability function P_i as being derived from a shared prior probability function p by conditionalizing on i's private information L_i. If the individuals have no shared background information, it is plausible to take p to be the uniform prior, following the *principle of insufficient reason* (though, of course, that principle raises some well-known philosophical issues, which we cannot discuss here). Any deviations from the uniform probability function on the part of some individual – i.e., in some function P_i – must then plausibly be due to some private information. But now consider the opinion profile $(P_1, ..., P_n)$ in which *every* P_i is the uniform probability function. For the individuals to arrive at this opinion profile, there must be a complete lack of private information, in addition to the lack of collectively shared background information. (If some individuals had relevant private information, some P_i would arguably have to be distinct from the uniform probability function.[11]) In such a situation of no information – private or shared – it seems plausible to require the collective probability function to be uniform. So, indifference preservation is plausible here.

By contrast, if the individuals have some shared background information, indifference preservation is questionable. The individuals' prior probability functions will not normally be uniform in this case, so any uniformity in an individual's posterior probability function P_i points towards the presence of some private information which has led the individual to update his or her probabilities from the non-uniform prior ones to uniform posterior ones. The collective probability function should therefore incorporate both the group's shared background information and the individuals' private information. There is no reason to expect that incorporating all this information will generally lead to the uniform probability function. Consequently, indifference preservation is not plausible here.

How should we choose the calibrating probability function P_0 when we cannot assume indifference preservation? Our answer to this question follows Dietrich (2010). Again, consider Case 2, where different individuals have different private information and there is at most some fixed background information that is collectively shared. Let p be every individual's prior probability function, assuming a shared prior (which may reflect the shared background information).

If none of the individuals holds any additional private information, then each individual i's probability function is simply $P_i = p$, and it is reasonable to require the group to have the same probability function p, because no further information

[11] Alternatively, it is possible for an individual to have multiple pieces of private information that perfectly cancel each other out, so that, on balance, his or her probability function remains uniform. Strictly speaking, to justify indifference preservation in such a case, we must assume that different individuals' private information is uncorrelated (i.e., mutually independent). We briefly discuss the issue of correlated private information in Section 10.

is available to the group. Formally, $P_{p,...,p} = p$.[12] By the definition of multiplicative pooling, the collective probability function $P_{p,...,p}$ is proportional to the product $P_0 p^n$ (where probability functions are viewed as functions defined on the set of worlds Ω). So, p, which is equal to $P_{p,...,p}$, must be proportional to $P_0 p^n$, which implies that P_0 must be proportional to $1/p^{n-1}$. Formally,

$$P_0(\omega) = \frac{c}{[p(\omega)]^{n-1}} \text{ for every world } \omega \text{ in } \Omega,$$

where c is a normalization factor to ensure that P_0 is a probability function.

This shows that the choice of P_0 is not free, but constrained by the individuals' prior probabilities. In particular, the probability assignments made by P_0 must depend strongly negatively on the individuals' prior probabilities. This idea can be generalized to the case in which different individuals have different priors, as shown in the Appendix.

10 Concluding remarks

We have discussed three classes of opinion pooling functions – linear, geometric, and multiplicative – and have shown that they satisfy different axioms and are justifiable under different conditions. Linear pooling may be justified on procedural grounds, but not on epistemic grounds. Geometric and multiplicative pooling may be justified on epistemic grounds, but which of the two is appropriate depends not just on the opinion profiles to be aggregated but also on the information on which they are based. Geometric pooling can be justified if all individuals' opinions are based on the same information (Case 1), while multiplicative pooling can be justified if every individual's opinions are based solely on private information, except for some shared background information held by everyone (Case 2).

There are, of course, many intermediate cases between Case 1 and Case 2, in which the opinion pooling problem becomes more complicated. First, there are cases in which an opinion profile is based on some information that is neither shared by everyone, nor held by a single individual alone, but shared by a proper subset of the individuals. In such cases, neither geometric nor multiplicative pooling is justified but a more complicated pooling function – involving a recursive construction – is needed (see Dietrich 2010).

Second, there are cases in which there are *correlations* between different individuals' private information – a possibility implicitly assumed away in our discussion so far. If different individuals' private information is correlated, the axiom of individualwise Bayesianity loses its force. To see this, note that the combined

[12] To ensure that the opinion profile $(p, ..., p)$ is in the domain of the pooling function, we assume that p belongs to $\mathcal{P}'$, i.e., is a regular probability function.

evidential strength of two pieces of *correlated* private information, represented by the likelihood functions L_1 and L_2, is not their product L_1L_2. So, it is not plausible to demand that $P_{P_1,\ldots,P_i^{L_1},\ldots,P_j^{L_2},\ldots,P_n} = P_{P_1,\ldots,P_n}^{L_1L_2}$, as individualwise Bayesianity (applied twice) would require. (On the subject of dependencies between different individuals' opinions, see Dietrich and List 2004 and Dietrich and Spiekermann 2013.)

In sum, it should be clear that there is no one-size-fits-all approach to probabilistic opinion pooling. We wish to conclude by mentioning some other approaches that we have not discussed. One such approach is *supra-Bayesian opinion pooling* (introduced by Morris 1974), a radically Bayesian approach. Here, the collective probability of each possible world is defined as the *posterior* probability of that world (held by a hypothetical Bayesian observer), conditional on learning what the opinion profile is. Opinion pooling then becomes a complex form of Bayesian updating. This presupposes a very rich probability model, which specifies not just the prior probability of each possible world, but also the probability of obtaining each possible opinion profile conditional on each possible world. In practice, it is unclear where such a rich model could come from, and how a group could agree on it. Nevertheless, from a radically Bayesian perspective, supra-Bayesian pooling is a very natural approach – or even the rationally required one.

There are also a number of approaches that not merely lead to different opinion pooling functions but redefine the aggregation problem itself. Here, the opinions to be aggregated are no longer given by probability functions, but by other formal objects. Two examples are the aggregation of *imprecise probabilities* (e.g., Moral and Sagrado 1998) and the aggregation of *ordinal probabilities*, which are expressed by probability orders (using the binary relation 'at least as probable as') rather than probability functions (e.g., Weymark 1997). Similarly, one could in principle use the tools of formal aggregation theory to study the aggregation of *ranking functions* (as discussed, e.g., by Spohn 2012).

In recent years, there has been much work on the aggregation of binary opinions, where a group seeks to assign the values 'true'/'false' or 'yes'/'no' to a set of propositions, based on the individuals' assignments – a problem now known as *judgment aggregation* (e.g., List and Pettit 2002; Dietrich 2007; Dietrich and List 2007; Nehring and Puppe 2010; Dokow and Holzman 2010; for a recent review, see List 2012). Truth-value assignments, especially in classical propositional logic, can be viewed as degenerate probability assignments (restricted to the values 0 and 1). Interestingly, the analogues of the axioms characterizing linear averaging in probabilistic opinion pooling typically lead to dictatorial aggregation in judgment-aggregation problems (for discussion, see Dietrich and List 2010).

Pauly and van Hees (2006) consider judgment-aggregation problems in many-valued (as distinct from two-valued) logics and show that some of the dictatorship results familiar from the two-valued case continue to hold in the many-valued case (for further results, see Duddy and Piggins 2013). Relatedly, Bradley and Wagner (2012) discuss the aggregation of probability functions that take values within a finite grid, such as the grid $\{k/10 : k = 0, 1, ..., 10\}$. They show that this aggregation problem is also susceptible to dictatorship results akin to those in judgment aggregation. Under certain conditions, the only eventwise independent and unanimity-preserving aggregation functions are the dictatorial ones.

The list of examples could be continued. For a unified framework that subsumes several aggregation problems under the umbrella of *attitude aggregation*, see Dietrich and List (2010). In an attitude-aggregation problem, each individual i holds an *attitude function* A_i, which assigns to each proposition on some agenda a value in some set V of admissible values, which could take a variety of forms. We must further specify some criteria determining when an attitude function counts as *consistent* or *formally rational*, and when not. The task, then, is to map each profile $(A_1, ..., A_n)$ of individual attitude functions in some domain to a collective attitude function. It should be evident that probabilistic opinion pooling, two-valued and many-valued judgment aggregation, and finite-grid probability aggregation can all be viewed as special cases of such attitude-aggregation problems, for different specifications of (i) the value set V and (ii) the consistency or rationality criteria. (An extension of this line of research, using an algebraic framework, can be found in Herzberg forthcoming.)

Finally, much of the literature on opinion pooling is inspired, at least in part, by Arrow's pioneering work in social choice theory (Arrow 1951/1963). Social choice theory, in the most general terms, addresses the aggregation of potentially conflicting individual inputs into collective outputs (for a survey, see List 2013). Much of the work in this area, following Arrow, focuses on the aggregation of preferences, welfare, or interests. The theory of opinion pooling can be seen as an epistemically oriented counterpart of Arrovian social choice theory.[13]

11 References

Aczél, J., and C. Wagner (1980) "A characterization of weighted arithmetic means", *SIAM Journal on Algebraic and Discrete Methods* 1(3): 259-260.

Aczél, J., C. T. Ng, and C. Wagner (1984) "Aggregation Theorems for Allocation Problems", *SIAM Journal on Algebraic and Discrete Methods* 5(1): 1-8.

Arrow, K. J. (1951/1963) *Social Choice and Individual Values*, New York: Wiley.

Bacharach, M. (1972) "Scientific disagreement", unpublished manuscript.

Bradley, R., F. Dietrich, and C. List (2014) "Aggregating Causal Judgments", *Philosophy of Science* 81(4): 491-515.

Bradley, R., and C. Wagner (2012) "Realistic opinion aggregation: Lehrer-

[13] We are very grateful to John Cusbert and Alan Hájek for helpful written comments on this article.

Wagner with a finite set of opinion values", *Episteme* 9(2): 91-99.
Chambers, C. (2007) "An ordinal characterization of the linear opinion pool", *Economic Theory* 33(3): 457-474.
DeGroot, M. H. (1974) "Reaching a Consensus", *Journal of the American Statistical Association* 69(345): 118-121.
Dietrich, F. (2007) "A generalised model of judgment aggregation", *Social Choice and Welfare* 28(4): 529-565.
Dietrich, F. (2010) "Bayesian group belief", *Social Choice and Welfare* 35(4): 595-626.
Dietrich, F., and C. List (2004) "A model of jury decisions where all jurors have the same evidence", *Synthese* 142(2): 175-202.
Dietrich, F., and C. List (2007) "Arrow's theorem in judgment aggregation", *Social Choice and Welfare* 29(1): 19-33.
Dietrich, F., and C. List (2010) ""The aggregation of propositional attitudes: towards a general theory", *Oxford Studies in Epistemology* 3: 215-234 [with "Corrigendum" on the authors' webpages].
Dietrich, F., and C. List (2013a) "Opinion pooling generalized – Part one: general agendas", working paper, London School of Economics.
Dietrich, F., and C. List (2013b) "Opinion pooling generalized – Part two: the premise-based approach", working paper, London School of Economics.
Dietrich, F., and K. Spiekermann (2013) "Independent Opinions? On the Causal Foundations of Belief Formation and Jury Theorems", *Mind* 122(487): 655-685.
Dokow, E., and R. Holzman (2010) "Aggregation of binary evaluations", *Journal of Economic Theory* 145(2): 495-511.
Duddy, C., and A. Piggins (2013) "Many-valued judgment aggregation: Characterizing the possibility/impossibility boundary", *Journal of Economic Theory* 148(2): 793-805.
Genest, C. (1984a) "A characterization theorem for externally Bayesian groups", *Annals of Statistics* 12(3): 1100-1105.
Genest, C. (1984b) "Pooling operators with the marginalization property", *Canadian Journal of Statistics* 12(2): 153-163.
Genest, C., K. J. McConway, and M. J. Schervish (1986) "Characterization of externally Bayesian pooling operators", *Annals of Statistics* 14(2): 487-501.
Genest, C., and J. V. Zidek (1986) "Combining Probability Distributions: A Critique and Annotated Bibliography", *Statistical Science* 1(1): 114-135.
Genest, C., and C. Wagner (1987) "Further Evidence against Independence Preservation in Expert Judgement Synthesis", *Aequationes Mathematicae* 32(1): 74-86.
Herzberg, F. (forthcoming) "Universal algebra for general aggregation theory: Many-valued propositional-attitude aggregators as MV-homomorphisms", *Journal of Logic and Computation*.
Lehrer, K., and C. Wagner (1981) *Rational Consensus in Science and Society*,

Dordrecht: Reidel.

List, C., and P. Pettit (2002) "Aggregating sets of judgments: an impossibility result", *Economics and Philosophy* 18(1): 89-110.

List, C. (2012) "The theory of judgment aggregation: An introductory review", *Synthese* 187(1): 179-207.

List, C. (2013) "Social Choice Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), <http://plato.stanford.edu/archives/win2013/entries/social-choice/>.

Madansky, A. (1964) "Externally Bayesian groups", Technical Report RM-4141-PR, RAND Corporation.

McConway, K. J. (1981) "Marginalization and Linear Opinion Pools", *Journal of the American Statistical Association* 76(374): 410-414.

Mongin, P. (1995) "Consistent Bayesian aggregation", *Journal of Economic Theory* 66(2): 313-351.

Moral, S., and J. Sagrado (1998) "Aggregation of imprecise probabilities", in Bouchon-Meunier, B. (ed.), *Aggregation and Fusion of Imperfect Information*, Heidelberg: Physica-Verlag, pp. 162–188.

Morris, P. A. (1974) "Decision analysis expert use", *Management Science* 20(9): 1233-1241.

Nehring, K., and C. Puppe (2010) "Abstract Arrovian Aggregation", *Journal of Economic Theory* 145(2): 467-494.

Pauly, M., and M. van Hees (2006) "Logical Constraints on Judgment Aggregation", *Journal of Philosophical Logic* 35: 569-585.

Raiffa, H. (1968) *Decision Analysis: Introductory Lectures on Decision Analysis*, Reading, Mass.: Addison-Wesley.

Spohn, W. (2012) *The Laws of Belief: Ranking Theory and Its Philosophical Applications*, Oxford: Oxford University Press.

Stone, M. (1961) "The Opinion Pool", *Annals of Mathematical Statistics* 32(4): 1339-1342.

Wagner, C. (1982) "Allocation, Lehrer Models, and the Consensus of Probabilities", *Theory and Decision* 14(2): 207-220.

Wagner, C. (1984) "Aggregating subjective probabilities: some limitative theorems", *Notre Dame Journal of Formal Logic* 25(3): 233-240.

Wagner, C. (1985) "On the Formal Properties of Weighted Averaging as a Method of Aggregation", *Synthese* 62(1): 97-108.

Weymark, J. (1997) "Aggregating Ordinal Probabilities on Finite Sets", *Journal of Economic Theory* 75(2): 407-432.

12 Appendix

12.1 Likelihood functions and their Bayesian interpretation

According to our definition, a *likelihood function* L assigns, to each world ω in Ω, a positive number $L(\omega)$, interpreted as the likelihood that the information is true in world ω. This notion of a likelihood function is slightly non-standard, because in statistics a likelihood function is usually associated with some information (data) that is explicitly representable in the relevant model. In our climate-panel example, by contrast, the information that a revolutionary carbon-capture-and-storage technology is in use cannot be represented by any event $A \subseteq \Omega$. To relate our notion of a likelihood function to the more standard one, we need to make the following construction.

Let us 'split' each world $\omega = (j, k, l)$ in Ω into two more refined worlds: $\omega^+ = (j, k, l, 1)$ and $\omega^- = (j, k, l, 0)$, in which the fourth characteristic specifies whether or not a revolutionary carbon-capture-and-storage technology is in use. The refined set of worlds, Ω', now consists of all such 'four-dimensional' worlds, formally, $\Omega' = \Omega \times \{0, 1\}$. The information that a revolutionary carbon-capture-and-storage technology is in use can then be represented as an event relative to the refined set of worlds Ω', namely the event consisting of all refined worlds whose fourth characteristic is 1; call this event E.

Under this construction, the non-standard likelihood function L on Ω corresponding to this information becomes a standard likelihood function relative to our refined set Ω'. Formally, for any unrefined world $\omega \in \Omega$,

$$L(\omega) = \Pr(E|\omega) = \frac{\Pr(\omega^+)}{\Pr(\omega)},$$

where Pr is a probability function for the refined set of worlds Ω', and any unrefined world ω in Ω is re-interpreted as the event $\{\omega^+, \omega^-\} \subseteq \Omega'$.

One can think of Pr as a refinement of a probability function for the original set Ω. Of course, different individuals i may hold different probability functions P_i on Ω, and so they may hold different refined probability functions $\Pr_i$ on Ω'. Nonetheless, the likelihood function $L(\omega) = \Pr_i(E|\omega)$ is supposed to be the same for all individuals i, as we focus on *objective* (or at least *intersubjective*) information, which has an uncontroversial interpretation in terms of its evidential support for worlds in Ω.

For present purposes, the individuals may disagree about prior probabilities, but not about the evidential value of the incoming information. A paradigmatic example of objective information is given by the case in which worlds in Ω correspond to rival statistical hypotheses (e.g., possible probabilities of 'heads' for a given coin) and the information consists of statistical data (e.g., a sequence of coin tosses).

Finally, we show that our rule for updating a probability function P based on the likelihood function L – formula (2) in the main text – is an instance of ordinary Bayesian updating, applied to the refined model. Note that $P^L(\omega)$, the probability assigned to ω after learning the information represented by L, can be interpreted as $\Pr(\omega|E)$, where E is the event in Ω' that corresponds to the information. By Bayes's theorem,

$$\Pr(\omega|E) = \frac{\Pr(\omega)\Pr(E|\omega)}{\sum_{\omega' \in \Omega} \Pr(\omega')\Pr(E|\omega')},$$

which reduces to

$$P^L(\omega) = \frac{P(\omega)L(\omega)}{\sum_{\omega' \in \Omega} P(\omega')L(\omega')},$$

as in formula (2).

12.2 How to calibrate a mutiplicative pooling function when there is no shared prior

For each individual i, let p_i denote i's prior probability function, and let p denote the prior probability function that the group as a whole will use, without asking – for the moment – where p comes from. Plausibly, in the absence of any private information, when each individual i's probability function is simply $P_i = p_i$, the group should stick to its own prior probability function p. Formally, $P_{p_1,...,p_n} = p$.[14] By the definition of multiplicative pooling, $P_{p_1,...,p_n}$ is proportional to the product $P_0 p_1 \cdots p_n$ (where probability functions are again viewed as functions defined on the set of worlds Ω). So, p, which is equal to $P_{p_1,...,p_n}$, must be proportional to $P_0 p_1 \cdots p_n$, which implies that P_0 must be proportional to $p/(p_1 \cdots p_n)$. Formally,

$$P_0(\omega) = \frac{cp(\omega)}{p_1(\omega) \cdots p_n(\omega)} \text{ for every world } \omega \text{ in } \Omega, \quad (3)$$

where c is an appropriate normalization factor.

This expression still leaves open how to specify p, the group's prior probability function. Plausibly, it should reflect the individual prior probability functions $p_1, ..., p_n$. Since the individuals' prior probabilities are not based on any informational asymmetry – they are, by assumption, based on the same background information – their aggregation is an instance of Case 1. Hence, geometric pooling is a reasonable candidate for determining p on the basis of $p_1, ..., p_n$. If we further

[14] We assume that each p_i belongs to $\mathcal{P}'$, i.e., is a regular probability function. So the profile $(p_1, ..., p_n)$ is in the domain of the pooling function.

wish to treat the individuals equally – perhaps because we equally trust their abilities to interpret the shared background information correctly – we might use *unweighted* geometric pooling, i.e., take p to be proportional to $p_1^{1/n} \cdots p_n^{1/n}$. As a result, expression (3) reduces to the following general formula:

$$P_0(\omega) = \frac{c[p_1(\omega)]^{1/n} \cdots [p_n(\omega)]^{1/n}}{p_1(\omega) \cdots p_n(\omega)} = \frac{c}{[p_1(\omega) \cdots p_n(\omega)]^{1-1/n}} \text{ for every world } \omega \text{ in } \Omega,$$

where c is an appropriate normalization factor.

We have now arrived at a unique solution to our opinion pooling problem, having specified a multiplicative pooling function without any free parameters. However, the present solution is quite informationally demanding. In particular, it requires knowledge of the individuals' prior probabilities. For more details, see Dietrich (2010).