Opinion pooling under asymmetric information

Franz Dietrich[1]
3rd July 2004

Abstract. If each member of a group assigns a certain probability to a hypothesis, what probability should the collective as a whole assign? More generally, how should individual probability functions be merged into a single collective one? I investigate this question in case that the individual probability functions are based on different information sets. Under suitable assumptions, I present a simple solution to this aggregation problem, and a more complex solution that can cope with any overlaps between different persons' information sets. The solutions are derived from an axiomatic system that models the individuals as well as the collective as Bayesian rational agents. Two notable features are that the solutions may be parameter-free, and that they incorporate each individual's information although the individuals need not communicate their (perhaps very complex) information, but rather reveal only the resulting probabilities.

Keywords: opinion pooling, probability aggregation, decision theory, social choice theory, Bayesian rationality, Bayesian aggregation, information.

1 Introduction

Suppose that a group is interested whether a given hypothesis H is true. If every individual assigns H a probability of 70%, what probability should the group assign to H? Is it exactly 70%, or perhaps more since different persons have independently confirmed H? The answer, I will show, crucially depends on the informational states of the individuals. If they have *identical* information, the collective has good reasons to adopt people's unanimous 70% probability judgment. In the case of informational *asymmetry*, a possibly much higher or lower collective probability may be appropriate, depending on the nature of the personal and the shared information.

This problem is an instance of the well-established discipline of *probability aggregation* or *opinion pooling*. In general, individual probabilities need of course not coincide, and more than one hypothesis may be of interest. The aim is to merge a profile $\mathrm{Pr}_1, ..., \mathrm{Pr}_n$ of individual probability measures into a single collective probability measure Pr. Two approaches may be distinguished. The first approach (adopted here) is axiomatic and inherently social-choice-theoretic: normative conditions are imposed on the aggregation rule, and the class of rules satisfying these conditions is derived. Already long ago, the two most prominent classes of rules, *linear* and

[1]Center For Junior Research Fellows, University of Konstanz, 78457 Konstanz (Germany). E-mail: franz.dietrich@uni-konstanz.de. Web: www.uni-konstanz.de/ppm/Dietrich.

logarithmic rules, have been characterised axiomatically. If $\Pr, \Pr_1, ..., \Pr_n$ have associated probability density (or mass) functions $f, f_1, ..., f_n$ (with respect to some fixed measure μ), a linear rule defines f as a weighted *arithmetic* average $\sum_{i=1}^{n} w_i f_i$, while a logarithmic rule defines f as (proportional to) a weighted *geometric* average $\Pi_{i=1}^{n} f_i^{w_i}$, where $w_1, ..., w_n \in [0, 1]$ are weights with sum 1. Linear rules have been characterised (under additional technical assumptions) by the *strong setwise function property* (McConway (1981) and Wagner (1982)), the *marginalisation property* (McConway (1981)), and, in a single-profile framework, by the probabilistic analogue of the *weak Pareto principle* (Mongin, 1995, 1998); and logarithmic rules famously satisfy *external Bayesianity* as defined in Section 5 (e.g. McConway (1978), Genest (1984), Genest, McConway and Schervish (1986)). Still an excellent reference for the fundamental issues in axiomatic probability aggregation is Genest and Zidek's (1986) literature review. The second approach (not adopted here) is the "supra Bayesian" approach, formally laid down in Morris' (1974) seminal paper and extended by a growing literature. Here, the collective probabilities are derived as the *posterior* probabilities (held by a real or virtual "supra Bayesian") conditional on the observed individual probability assignments. This approach is clearly very general and elegant, but it faces the problem of having to specify prior probabilities and likelihoods of the individual probability assignments.

In the axiomatic approach, a weakness of linear and logarithmic rules is the partial arbitrariness of the choice of the weights $w_1, ..., w_n$. Intuitively, w_i should reflect i's information (and ability to interpret information). As Genest and Zidek (1986) put it, "expert weights do allow for some discrimination [...], but in vague, somewhat ill defined ways" (p. 120), and "no definite indications can be given concerning the choice or interpretation of the weights" (p. 118).

This paper proposes a new axiomatic framework that explicitly accounts for informational asymmetry. I thereby avoid the problematic determination of information-specific weights. The model differs from earlier axiomatic approaches in that it is inherently Bayesian, a feature shared with the "supra Bayesian" approach. Given that both individuals and the collective are modelled as Bayesian rational agents, the findings are relevant to the theory of *Bayesian aggregation*, which aims to merge individual beliefs/values/preferences satisfying Bayesian rationality conditions (in the sense of Savage 1954 or Jeffrey 1983) into equally rational collective ones; for the *ex ante* approach, e.g. Seidenfeld et al. (1989), Broome (1990), Schervish et al. (1991) and Mongin (1995, 1998); for the *ex post* approach, e.g. Hylland and Zeckhauser (1979), Levi (1990), Hild (1998) and Risse (2001); for an excellent overview/critique, see Risse (2003).

In Section 2, I present the model and derive the aggregation rule that follows from it. Section 3 contains a numerical example. Sections 4 and 5 addresses the choice of the "collective prior", the (only) open parameter in the aggregation rule. In Section 6, I discuss a problematic independence assumption made up to this point, and prove that (under certain conditions) it holds *if and only if* no proper subgroup exclusively shares any information. In Section 7, I allow arbitrary information overlap; I generalise my approach into a technique of recursive probability aggregation by defining an algorithm that gradually builds up the collective probability function so as to incorporate more and more and finally all information.

2 An axiomatic model

Consider a group of persons $i = 1, ..., n$ ($n \geq 2$), faced with a (non-empty) finite or countably infinite set $\mathcal{H}$ of hypotheses H. I call *shared* information/knowledge that knowledge held simultaneously by *all* group members. Each person i may additionally hold *personal information.* If the group is a court jury and $\mathcal{H}$ consists of hypotheses about the defendant's extent of guilt, the shared information might include the charge (read by all jurors), while juror 3's information might additionally include an observed smile on the face of the defendant.

Let Ω be a (non-empty) set of "possible worlds", i.e. possible under the shared information. Formally, assume that $\mathcal{H}$ is a partition of Ω, i.e. that each hypothesis $H \in \mathcal{H}$ is a (non-empty) subset of Ω (the set of worlds in which the hypothesis holds) such that $\cup_{H \in \mathcal{H}} H = \Omega$ and $H \cap H' = \emptyset$ for any distinct $H, H' \in \mathcal{H}$. A simple case is a binary problem $\mathcal{H} = \{H, \Omega \backslash H\}$.

Definition 1 Π *is the set of all functions* $\pi : \mathcal{H} \to (0, 1]$ *such that* $\sum_{H \in \mathcal{H}} \pi(H) = 1$, *called "(positive) probability functions (on* $\mathcal{H}$*)" (whereas a "probability measure" is defined on a* σ*-algebras in* Ω*).*

Each person i is asked to submit both

• a probability function $p_i \in \Pi$, representing the probabilities that i assigns based on (i's interpretation of) the *shared* knowledge,

• a probability function $\pi_i \in \Pi$, representing the probabilities that i assigns based on (i's interpretation of) i's *full* (i.e. shared or personal) knowledge.

The task is to derive, based on the submitted $p_1, \pi_1, ..., p_n, \pi_n$,

• a probability function $\pi \in \Pi$, representing the probabilities that the collective assigns based on the group's *full* (shared or any i' personal) knowledge.

The submission of p_i is new compared to the standard approach. While providing π_i has to be an isolated exercise, not so for p_i: as the functions $p_1, ..., p_n$ are all based on the same (shared) information, the group may deliberate over how to interpret this information and what probabilities rationally follow from it, possibly resulting in an agreement $p_1 = ... = p_n$.

To interpret p_i as a *prior* probability function and π_i as a *posterior* probability function, suppose that, to each person i, there are (without being revealed)

- an event $E_i \subseteq \Omega$, i's *personal evidence*, representing i's personal information;

- a set of events $\mathcal{A}_i$, a σ-algebra in Ω, interpreted as the set of events to which i assigns probabilities, while i may be agnostic regarding events $E \notin \mathcal{A}_i$; by assumption, $E_i \in \mathcal{A}_i$ and $\mathcal{H} \subseteq \mathcal{A}_i$;

- a ("prior") probability measure $P_i : \mathcal{A}_i \to [0, 1]$ representing i's probability assignments based on (i's interpretation of) the *shared* information (hence "prior"); by assumption, $P_i(E_i) > 0$ and $P_i(H) > 0$ for all $H \in \mathcal{H}$.

Recall that a conditional probability $P_i(A|B)$ is defined as $P_i(A \cap B)/P_i(B)$, for all $A, B \in \mathcal{A}_i$ with $P_i(B) \neq 0$; and similarly for any other probability measure.

Individual Bayesian Rationality (IBR). For each person i and hypothesis $H \in \mathcal{H}$,

$$p_i(H) = P_i(H) \text{ and } \pi_i(H) = P_i(H|E_i).$$

p_i and π_i are therefore called i's "prior" and "posterior" probably functions (where "prior" and "posterior" need not have a temporal interpretations as the observation of E_i may precede that of the shared information). By submitting $\pi_i, p_i \in \Pi$, i reveals not i's full P_i, but only the restrictions to $\mathcal{H}$ of P_i and of $P_i(.|E_i)$. Other characteristics of P_i (such as the likelihoods $P_i(E_i|H)$, $H \in \mathcal{H}$) not only remain hidden, but also i need not be fully aware of them. This corresponds to the common "as if" interpretation of theories of rationality: a Bayesian rational agent is assumed update beliefs/hold preferences in a way that is consistent with Bayesian conditionalisation/expected utility maximisation *whether or not* the agent is consciously calculating any posterior probabilities/expected utilities. (A purely "as if" interpretation may of course be questioned.)

I treat the *group* or *collective* as a separate (virtual) agent with its own beliefs. Specifically, I suppose that there are

- a set of events $\mathcal{A}$, a σ-algebra in Ω, interpreted as the set of events to which the collective assigns probabilities; by assumption, $E_1, ..., E_n \in \mathcal{A}$ and $\mathcal{H} \subseteq \mathcal{A}$;

- a ("prior") probability measure $P : \mathcal{A} \to [0, 1]$ representing the collective's probability assignments based on (the collective's interpretation of) the *shared* information (hence "prior"); by assumption, $P(E_1 \cap ... \cap E_n) > 0$ and $P(H) > 0$ for all $H \in \mathcal{H}$.

$\mathcal{A}$ and P are collective counterparts of $\mathcal{A}_i$ and P_i. The counterpart of (IBR) is:

Collective Bayesian Rationality (CBR). For each hypothesis $H \in \mathcal{H}$,

$$\pi(H) = P(H|E_1 \cap ... \cap E_n).$$

So π incorporates both the shared information (contained in P) and all personal evidences $E_1, ..., E_n$ (conditionalised upon). At first, it may seem obscure how one could calculate $\pi(H) = P(H|E_1 \cap ... \cap E_n)$, as P is held by a *virtual* agent (the collective), and the evidences $E_1, ..., E_n$ are not revealed. The key is to connect P to $P_1, ..., P_n$ via assumption (AL) below.

Let $p := P|_{\mathcal{H}}$ be the restriction of P to $\mathcal{H}$ (the pair π, p is the collective counterpart of the pair π_i, p_i). For any hypothesis $H \in \mathcal{H}$, by (CBR) and Bayes' rule

$$\pi(H) = \frac{p(H)P(E_1 \cap ... \cap E_n|H)}{\sum_{H' \in \mathcal{H}} p(H')P(E_1 \cap ... \cap E_n|H')}. \tag{1}$$

I now make an independence assumption to be discussed and relaxed later; the assumption is closely related to Fitelson's (2001) *evidential independence.*

Independence (Ind). For each hypothesis $H \in \mathcal{H}$, the personal evidences $E_1, ..., E_n$ are independent conditional on H, i.e. $P(E_1 \cap ... \cap E_n|H) = P(E_1|H) \cdots P(E_n|H)$.

Applying (Ind) to (1) yields

$$\pi(H) = \frac{p(H)P(E_1|H) \cdots P(E_n|H)}{\sum_{H' \in \mathcal{H}} p(H')P(E_1|H') \cdots P(E_n|H')}. \tag{2}$$

Which values should be used for the collective likelihoods $P(E_i|H)$? I assume:

Acceptance of Likelihoods (AL). For all persons i and hypotheses $H \in \mathcal{H}$, $P(E_i|H) = P_i(E_i|H)$.

This principle requires the collective to take over i's own interpretation of i's evidence E_i as given by i's likelihood assignments $P_i(E_i|H)$, $H \in \mathcal{H}$. (AL) plays an implicit role, as individual likelihoods are not revealed. How can (AL) be justified? Why not take other person's interpretations of E_i also into account by defining $P(E_i|H)$ as some "compromise" of $P_1(E_i|H), ..., P_n(E_i|H)$? First, person i may be the only person to actually possess a likelihood of E_i; i.e., perhaps that for persons $j \neq i$ we have $E_i \notin \mathcal{A}_j$, so that $P_j(E_i|H)$ is not even defined. Indeed, having not observed E_i, j may not have spent any thoughts on E_i, nor on its likelihood. Second, a "likelihood compromise" could only be formed after each person j reveals $P_j(E_i|H)$ (assuming that $E_i \in \mathcal{A}_j$); which in turn supposes that first i communicates the exact nature of E_i to the rest of the group. This is not only at odds with the present approach, but may also be unfeasible in practice: given the possible complexity of E_i and the limitations of language, of time, of i's ability to describe E_i, of j's ($j \neq i$) ability to understand E_i etc., j could probably learn at most some approximation $\tilde{E}_i$ of E_i, and so j could at most provide j's likelihood of $\tilde{E}_i$, which at most approximates j's likelihood of the true E_i ($P_j(\tilde{E}_i) \approx P_j(E_i)$).

By (AL), I may replace each likelihood $P(E_i|H)$ in (2) by $P_i(E_i|H)$. As $P_i(E_i|H)$ is not submitted information, I express it in terms of $\pi_i(H)$ and $p_i(H)$: by (IBR) and Bayes' rule,

$$P_i(E_i|H) = \frac{P_i(H|E_i)P_i(E_i)}{P_i(H)} = \frac{\pi_i(H)}{p_i(H)}P_i(E_i).$$

Substituting this into (2) and then simplifying,

$$\pi(H) = \frac{\frac{\pi_1(H)}{p_1(H)}P_1(E_1) \cdots \frac{\pi_n(H)}{p_n(H)}P_n(E_n)p(H)}{\sum_{H' \in \mathcal{H}} \frac{\pi_1(H')}{p_1(H')}P_n(E_1) \cdots \frac{\pi_n(H')}{p_n(H')}P_n(E_n)p(H')} = \frac{\frac{\pi_1(H)}{p_1(H)} \cdots \frac{\pi_n(H)}{p_n(H)}p(H)}{\sum_{H' \in \mathcal{H}} \frac{\pi_1(H')}{p_1(H')} \cdots \frac{\pi_n(H')}{p_n(H')}p(H')}.$$

This formula simplifies if people have reached an agreement on how to interpret their shared information:

Agreement on the Prior (AP). $p_1 = ... = p_n = p$.

Under (AP), all persons i submit the same prior p_i, and this prior is taken over as the collective prior p (in accordance with the *unanimity principle*). As under (IBR) $p_i = P_i|_{\mathcal{H}}$ and by definition $p = P|_{\mathcal{H}}$, (AP) means that $P_1, ..., P_n, P$ all agree on $\mathcal{H}$ (but perhaps not outside $\mathcal{H}$), i.e. $P_1|_{\mathcal{H}} = ... = P_n|_{\mathcal{H}} = P|_{\mathcal{H}}$.

I now collect in a theorem. Functions $f, g : \mathcal{H} \to \mathbf{R}$ are "proportional", written $f \propto g$, if there exists a constant $k \neq 0$ such that $f(H) = kg(H)$ for all $H \in \mathcal{H}$.

Theorem 1 *Assume (IBR), (CBR), (Ind) and (AL). Then the collective probability of each hypothesis $H \in \mathcal{H}$ is given by*

$$\pi(H) = \frac{\frac{\pi_1(H)}{p_1(H)} \cdots \frac{\pi_n(H)}{p_n(H)} p(H)}{\sum_{H' \in \mathcal{H}} \frac{\pi_1(H')}{p_1(H')} \cdots \frac{\pi_n(H')}{p_n(H')} p(H')}, \text{ in short } \pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} p.$$

If in addition (AP) holds, then

$$\pi(H) = \frac{\pi_1(H) \cdots \pi_n(H) / p_1(H)^{n-1}}{\sum_{H' \in \mathcal{H}} \pi_1(H') \cdots \pi_n(H') / p_1(H')^{n-1}}, \text{ in short } \pi \propto \pi_1 \cdots \pi_n / p_1^{n-1}.$$

Three remarks:

1. As promised, π was calculated without people having to share their evidences E_i or their likelihoods $P(E_i|H)$, $H \in \mathcal{H}$.

2. If (AP) fails the aggregation problem is not yet fully solved, because the parameter p need still be chosen, a problem addressed in Section 4.

3. Assume a unanimous *posterior* agreement $\pi_1 = ... = \pi_n$. Then only in special cases π equals $\pi_1 = ... = \pi_n$ (suggesting that the *unanimity principle*, often required in standard opinion pooling without $p_1, ..., p_n$, is unjustified in the case of informational asymmetry). One such special case is that $p_i = \pi_i$ for each person i, i.e. the evidences $E_1, ..., E_n$ neither confirm nor disconfirm any hypothesis.

3 A numerical example for a simple case

Consider the simple case of a binary problem $\mathcal{H} = \{H, \Omega \backslash H\}$ with Agreement on the Prior (AP), i.e. $p_1 = ... = p_n = p$. The collective posterior of H is then

$$\pi^* = \frac{\pi_1^* \cdots \pi_n^* / (p^*)^{n-1}}{\pi_1^* \cdots \pi_n^* / (p^*)^{n-1} + (1 - \pi_1^*) \cdots (1 - \pi_n^*) / (1 - p^*)^{n-1}}, \quad (3)$$

where $p^* := p(H)$, $\pi^* := \pi(H)$ and $\pi_i^* := \pi_i(H)$.[2] For the case of group size $n = 2$, Table 1 contains the values of π^* for all possible combinations of values of p^*, π_1^*, π_2^* in the grid $\{.1, .25, .5, .75, .9\}$. Note how drastically π^* depends on the shared prior p^*. By shifting p^* below (above) the π_i^*s, π^* quickly approaches 1 (0); intuitively, if $E_1, ..., E_n$ all point into the same direction, their conjunction points even more into that direction. But if the prior p^* is somewhere in the middle of the π_i^*s, π^* may be

[2] The entries are rounded results if 3 decimal digits are reported, and exact results else.

	π_1^*, π_2^* :	p^* : .1	.25	.5	.75	.9
	.1, .1	.1	.036	.012	.004	.001
	.25, .1	.25	.1	.036	.012	.004
	.25, .25	.5	.25	.1	.036	.012
	.5, .1	.5	.25	.1	.036	.012
	.5, .25	.75	.5	.25	.1	.036
	.5, .5	.9	.75	.5	.25	.1
	.75, .1	.75	.5	.25	.1	.036
π_1^*, π_2^* :	.75, .25	.9	.75	.5	.25	.1
	.75, .50	.964	.9	.75	.5	.25
	.75, .75	.988	.964	.9	.75	.5
	.9, .1	.9	.75	.5	.25	.1
	.9, .25	.964	.9	.75	.5	.25
	.9, .5	.988	.964	.9	.75	.5
	.9, .75	.996	.988	.964	.9	.75
	.9, .9	.999	.996	.988	.964	.9

Table 1: Collective probability $\pi^* = \pi(H)$ in dependence of shared prior $p^* = p(H)$ and individual posteriors $\pi_i^* = \pi_i(H)$ for group size $n = 2$ and $\mathcal{H} = \{H, H^c\}$.

moderate; intuitively, if $E_1, ..., E_n$ point into different directions, their conjunction need not strongly point into any direction. Rewriting (3) as

$$\pi^* = \frac{1}{1 + (1/\pi_1^* - 1) \cdots (1/\pi_n^* - 1)/(1/p^* - 1)^{n-1}}, \tag{4}$$

shows that π^* is a strictly increasing function of $\pi_1^*, ..., \pi_n^*$ for fixed prior p^*, but a strictly decreasing function of p^* for fixed posteriors $\pi_1^*, ..., \pi_n^*$ (where $\pi^* \to 1(0)$ as $p^* \to 0(1)$). How can one make sense of π^* depending negatively on the prior p^*? Does more prior support for H reduces H's posterior probability? No, since more prior support for H increases not only the prior p^* but also each posterior π_i^*. Very intuitively, in (4) the prior information "comes in" n times (through the terms $(1/\pi_i^* - 1)$, which increase π^*) and is "taken out" $n - 1$ times (through the $n - 1$ terms $(1/p^* - 1)$, which reduce π^*), hence is overall accounted for exactly once, as expected.

4 The choice of the collective prior p when there is disagreement over the prior

If the interpretation of the shared information is controversial and hence (AP) fails, the group needs to determine the collective prior p in Theorem 1's formula. At least two strategies are imaginable. First, some (possibly appointed) person, perhaps a group member, may be asked to choose p, either by using his/her own prior, or by taking some inspiration from the submitted priors $p_1, ..., p_n$, or by using statistical estimation techniques if available. This solution has obvious limitations, including a lack of democraticity. An alternative is to replace p by $F(p_1, ..., p_n)$, i.e. to define

$$\pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n), \tag{5}$$

where $F : \Pi^n \to \Pi$ is a standard *opinion pool*. An opinion pool F maps each profile of (individual) functions $(p_1, ..., p_n) \in \Pi^n$ to a (collective) function $F(p_1, ..., p_n) \in \Pi$, the "compromise" between $p_1, ..., p_n$. At first sight, one may wonder what is gained by (5) compared to the standard approach of defining $\pi = F(\pi_1, ..., \pi_n)$ without caring about $p_1, ..., p_n$. Does formula (5) not just shift the original aggregation problem – pooling $\pi_1, ..., \pi_n$ into π – towards an equally complex aggregation problem about priors – pooling $p_1, ..., p_n$ into p? In an important respect, pooling $p_1, ..., p_n$ is simpler than pooling $\pi_1, ..., \pi_n$: pooling $p_1, ..., p_n$ involves no informational asymmetry since each of $p_1, ..., p_n$ is based on the same (shared) information. Hence any disagreement between $p_1, ..., p_n$ is due solely to different *interpretations* of that same body of information. This informational symmetry provides an argument for using *equal weights* in pooling $p_1, ..., p_n$, whereas pooling $\pi_1, ..., \pi_n$ may involve the difficult exercise of assigning information-specific weights to people. Let me explain this more precisely. The literature's two most prominent types of opinion pools $F : \Pi^n \to \Pi$ (see the introduction) are

linear opinion pools	$F(p_1, ..., p_n) = w_1 p_1 + ... + w_n p_n,$
logarithmic opinion pools	$F(p_1, ..., p_n) \propto p_1^{w_1} \cdots p_n^{w_n},$

where $w_1, ..., w_i \in [0, 1]$ are "weights" with $\sum_{i=1}^n w_i = 1$. The former is a (weighted) arithmetic average of $p_1, ..., p_n$, the latter is (proportional to) a (weighted) geometric average of $p_1, ..., p_n$ with factor of proportionality given by $\sum_{H \in \mathcal{H}} F(p_1, ..., p_n)(H) = 1$, i.e.

$$F(p_1, ..., p_n) = p_1^{w_1} \cdots p_n^{w_n} \Bigg/ \sum_{H \in \mathcal{H}} [p_1(H)]^{w_1} \cdots [p_n(H)]^{w_n} .$$

If F is a linear resp. logarithmic opinion pool, formula (5) becomes

$$\pi = \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} (w_1 p_1 + ... + w_n p_n) \tag{6}$$

$$\text{resp. } \pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} p_1^{w_1} \cdots p_n^{w_n} = \frac{\pi_1}{p_1^{1-w_1}} \cdots \frac{\pi_n}{p_n^{1-w_n}}. \tag{7}$$

How should the weights $w_1, ..., w_n$ be chosen? In general, unequal weights may be justified *either* by different information states *or* by different competence, i.e. ability to interpret information. The former reason does not apply here, since $p_1, ..., p_n$ are by definition based on the same (shared) information. If, in addition, differences of competence are either inexistent, or unknown, or not to be taken into account for reasons of procedural fairness, then equal weights $w_1 = ... = w_n = 1/n$ are justified.

Taking equal weights in (6) resp. (7) yields

$$\pi = \frac{1}{n}\frac{\pi_1}{p_1}\cdots\frac{\pi_n}{p_n}(p_1 + ... + p_n) \tag{8}$$

$$\text{resp. } \pi \propto \frac{\pi_1}{p_1^{1-1/n}}\cdots\frac{\pi_n}{p_n^{1-1/n}}, \tag{9}$$

which contains no unknown parameters, hence solves the aggregation problem.

5 External and interpersonal Bayesianity

Let me give two potential arguments in favour of a logarithmic (or, more generally, *externally Bayesian*) opinion pool F in (5), rather than for instance a linear opinion pool. Note first that in (5) π is a function of the vector $(p_1, \pi_1..., p_n, \pi_n) \in (\Pi \times \Pi)^n = \Pi^{2n}$, containing every person's prior and posterior.

Definition 2 *A "generalised opinion pool" ("GOP") or "generalised probability aggregation rule" is a function* $G : \Pi^{2n} \to \Pi$.

Unlike a standard opinion pool $F : \Pi^n \to \Pi$, a GOP G also takes as inputs the p_is, i.e. people's interpretations of the shared information. I have shown that, under certain conditions, a GOP G should take the form (5), i.e.

$$F(p_1, \pi_1, ..., p_n, \pi_n) \propto \frac{\pi_1}{p_1}\cdots\frac{\pi_n}{p_n}F(p_1, ..., p_n), \tag{10}$$

where $F : \Pi^n \to \Pi$ is a standard opinion pool that merges the priors $p_1, ..., p_n$.

From a Bayesian perspective, (at least) two natural conditions may be imposed on a GOP, which I call *external* and *interpersonal Bayesianity.* The former is the analogue of external Bayesianity for standard opinion pools F, whereby it should not matter whether an information comes in before or after pooling, i.e. the pooling operator F should commute with updating. Formally, for every probability function $p \in \Pi$ and every ("likelihood") function $l : \mathcal{H} \to (0, 1]$ I denote by p^l the (updated) probability function $p^l \in \Pi$ defined by

$$p^l(H) := \frac{l(H)p(H)}{\sum_{H' \in \mathcal{H}} l(H')p(H')}, \text{ in short } p^l \propto lp. \tag{11}$$

Here, l is interpreted as the likelihood function $P(E|.)$ for some evidence E, so that p^l is a posterior probability. A standard opinion pool $F : \Pi^n \to \Pi$ is called *externally Bayesian* if

$$F(p_1^l, ..., p_n^l) = F(p_1, ..., p_n)^l$$

for every profile $(p_1, ..., p_n) \in \Pi^n$ and ("likelihood") function $l : \mathcal{H} \to (0, 1]$ (Madansky 1964). In particular, logarithmic opinion pools F are externally Bayesian. To defend the use of an externally Bayesian opinion pool F in (10), let me define an analogous concept for GOPs:

Definition 3 *A GOP G is called "externally Bayesian" if*

$$G(p_1^l, \pi_1^l, ..., p_n^l, \pi_n^l) = G(p_1, \pi_1, ..., p_n, \pi_n)^l$$

for every profile $(p_1, \pi_1, ..., p_n, \pi_n) \in \Pi^{2n}$ and ("likelihood") function $l : \mathcal{H} \to (0, 1]$.

On the left hand side of this equation not only all posteriors are updated (π_i^l), but also all priors (p_i^l), because the incoming information is observed by everybody, hence part of the shared information, hence contained in the priors.

While external Bayesianity requires that it be irrelevant whether pooling happens before or after updating, a different question is whether it matters *who* in the group has observed a given information. *Interpersonal Bayesianity* requires that it be irrelevant whether every or just a single person obtains a given information:

Definition 4 *A GOP G is called "interpersonally Bayesian" if, for each person i,*

$$G(p_1, \pi_1, ..., p_{i-1}, \pi_{i-1}, p_i, \pi_i^l, p_{i+1}, \pi_{i+1}, ..., p_n, \pi_n) = G(p_1^l, \pi_1^l, ..., p_n^l, \pi_n^l)$$

for every profile $(p_1, \pi_1, ..., p_n, \pi_n) \in \Pi^{2n}$ and ("likelihood") function $l : \mathcal{H} \to (0, 1]$.

On the left hand side of this equation, i's prior is not updated (p_i, not p_i^l), because the incoming information, being observed just by person i, is not part of the shared knowledge, hence not reflected in any prior. Interpersonal Bayesianity is based on the idea that the collective probabilities should incorporate all information available *somewhere* in the group, whether it is held by a single or every person. External and interpersonal Bayesianity together imply that, for each person i,

$$G(p_1, \pi_1, ..., p_{i-1}, \pi_{i-1}, p_i, \pi_i^l, p_{i+1}, \pi_{i+1}, ..., p_n, \pi_n) = G(p_1, \pi_1, ..., p_n, \pi_n)^l$$

for every profile $(p_1, \pi_1, ..., p_n, \pi_n) \in \Pi^{2n}$ and ("likelihood") function $l : \mathcal{H} \to (0, 1]$.

It turns out that, *if* a GOP G takes the form (10), *then* external and interpersonal Bayesianity are in fact equivalent, and equivalent to external Bayesianity of F:

Theorem 2 *If a generalised opinion pool G has the form* (10), *where F is some opinion pool, the following conditions are equivalent:*
(i) *G is externally Bayesian;*
(ii) *G is interpersonally Bayesian;*
(iii) *F is externally Bayesian.*

Proof. I show that (i) is equivalent with each of (ii) and (iii). By (10),

$$G(p_1^l, \pi_1^l, ..., p_n^l, \pi_n^l) \propto \frac{\pi_1^l}{p_1^l} \cdots \frac{\pi_n^l}{p_n^l} F(p_1^l, ..., p_n^l),$$

and hence by (11)

$$G(p_1^l, \pi_1^l, ..., p_n^l, \pi_n^l) \propto \frac{l\pi_1}{lp_1} \cdots \frac{l\pi_n}{lp_n} F(p_1^l, ..., p_n^l) = \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1^l, ..., p_n^l). \tag{12}$$

On the other hand, again by (10) and (11),

$$G(p_1, \pi_1, ..., p_n, \pi_n)^l \propto l\frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n) \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n)^l. \tag{13}$$

Relations (12) and (13) together immediately imply that G is externally Bayesian if and only if F is externally Bayesian. Further, again by (10) and (11),

$$\begin{aligned} G(p_1, \pi_1, ..., p_{i-1}, \pi_{i-1}, p_i, \pi_i^l, p_{i+1}, \pi_{i+1}, ..., p_n, \pi_n) &\propto l\frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n) \\ &\propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n)^l. \end{aligned}$$

This together with (12) implies that G is interpersonally Bayesian if and only it F is externally Bayesian. ■

So, if one desires G to be externally or interpersonally Bayesian, one is bound to use an externally Bayesian opinion pool F in (10). The simplest examples of externally Bayesian opinion pools F are logarithmic opinion pools; they lead to a GOP of the form (7), and to the GOP (9) in the case of equal weights. There exist other, more complicated externally Bayesian opinion pools F, characterised in full generality by Genest, McConway, and Schervish (1986, Theorem 2.5); yet logarithmic opinion pools become the only solutions if additional properties are required from F and $|\mathcal{H}| \geq 3$ (see Genest, McConway, and Schervish (1986), Corollary 4.5).

6 When is Independence (Ind) violated?

Let us go back to the foundations of the model. Beside Individual Bayesian Rationality (IBR), the most problematic assumption is Independence (Ind). An important source for failure of (Ind) is what I call "*problematic information overlap*", that is, the existence of information held by more than one but less than every person. I will prove that, under certain conditions, (Ind) holds *if and only if* there is no problematic information overlap.

By a person i's *information set* I mean, informally, the (possibly quite enormous) collection of i's relevant observations/items of information. (Formally, one may define i's information set $\mathcal{I}_i$ as a set of non-empty events $E \subseteq \Omega$, where each $E \in \mathcal{I}_i$ represents an item of information.[3]) In the case of a jury faced with hypotheses about the defendant's guilt, i's information set might include the observations "an insecure smile on the defendant's face", "the defendant's fingerprint near the crime scene", "two contradictory statements by witness x", etc.

Figure 1 shows information sets, *not* sets of possible worlds $A \subseteq \Omega$. These two concepts are in fact opposed to each other: the larger the information set, the smaller

[3]This definition is in one respect odd (without this causing any problem here): any observation made by *every* person is represented by the same (maximal) event $E = \Omega$ (i.e. $E \in \mathcal{I}_1 \cap ... \cap \mathcal{I}_n$ implies $E = \Omega$), as by assumption *each* $\omega \in \Omega$ is a possible world under the shared information.

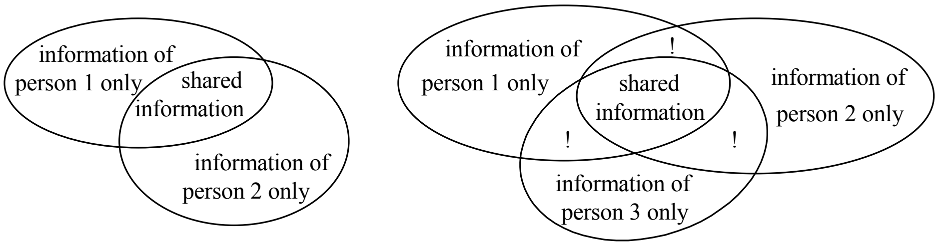

Figure 1: Information sets in a group of $n = 2$ perons (no illigal information overlap), and a group of $n = 3$ persons (problematic information overlap marked with "!")

the corresponding set of worlds (satisfying the information); the *union* of information sets compares to the *intersection* of the sets of worlds. (Formally, to an information set $\mathcal{I}$ corresponds the set of worlds $\cap_{E\in\mathcal{I}} E \subseteq \Omega$, interpreted as Ω if $\mathcal{I} = \emptyset$. Thus i's evidence E_i equals $\cap_{E\in\mathcal{I}_i\setminus(\mathcal{I}_1\cup\ldots\cup\mathcal{I}_n)} E$, the intersection of all of i's not-shared items of information; by footnote 3, this actually equals $\cap_{E\in\mathcal{I}_i} E$.)

Here is the potential problem. Consider any piece of information contained in the information sets of more than one but less than all persons i – something impossible in groups of size $n = 2$ but possible in larger groups, as illustrated by the "!" fields in Figure 1. This information is not part of the shared information, but part of many person i's personal evidences E_i. Such "problematic information overlap" typically creates positive correlations between the E_is in question. As a stylised example, consider a jury of $n = 3$ jurors faced with the hypothesis of guilt of the defendant (H). All jurors have read the charge (shared information), and moreover juror 1 has listened to the first witness report and observed the defendant's nervousness (E_1), juror 2 has listened to the second witness report and observed the defendant's smiles (E_2), and juror 3 has listened both witness reports and had a private chat with the defendant (E_3). Note the problematic information overlap between jurors 1 and 3 and between jurors 2 and 3, which typically causes E_3 to be positively correlated with E_1 and with E_2.

Let me now prove a formal result.

Definition 5 *A subgroup, M, is a non-empty subset of the group $N := \{1, ..., n\}$; it is called a "proper" subgroup if* $1 < |M| < n$.

Now suppose that to each subgroup M there is a non-empty event $E^M \subseteq \Omega$, *M's "exclusively shared evidence"*, representing all information held by *each of and only* the persons in M, where by assumption

- $E_i = \cap_{\{i\}\subseteq M\subseteq N} E^M$ for all persons i (as i has observed those E^M with $i \in M$);[4]
- $E^N = \Omega$ (as *any* world $\omega \in \Omega$ is assumed possible under the shared information);

[4]Why not rather assume that $E_i = \cap_{\{i\}\subseteq M\subsetneq N} E^M$, as E_i should not contain information held by *everybody*? In fact, both assumption are equivalent since by $E^N = \Omega$ an additional intersection with E^N has no effect.

• each E^M belongs to $\mathcal{A}$, the domain of the probability measure P (which holds in particular if $\mathcal{A}$ contains *all* subsets of Ω).

For instance, the "!" fields in Figure 1 represent $E^{\{1,2\}}$, $E^{\{1,3\}}$ and $E^{\{2,3\}}$. (E^M is interpretable as the intersection $\cap_{E \in (\cap_{i \in M} \mathcal{I}_i) \backslash (\cup_{i \notin M} \mathcal{I}_i)} E$ of all items of information E contained in *each* of the information sets $\mathcal{I}_i$, $i \in M$, but in *none* of the information sets $\mathcal{I}_i$, $i \notin M$, where this intersection is Ω if $(\cap_{i \in M} \mathcal{I}_i) \backslash (\cup_{i \notin M} \mathcal{I}_i) = \emptyset$.)

For all subgroups $M \neq N$, "no information exclusively shared by M" means just $E^M = \Omega$. This leads to a simple definition of "no problematic information overlap":

No Problematic Information Overlap (NoPIO). $E^M = \Omega$ for all proper subgroups M.

In Figure 1, (NoPIO) requires each "!" field to be empty, more precisely $E^{\{1,2\}} = E^{\{1,3\}} = E^{\{2,3\}} = \Omega$. Now, I assume:

(Ind*) The events E^M, $\emptyset \neq M \subseteq N$, are ($P$-)independent conditional on each $H \in \mathcal{H}$.

This new independence assumption is less problematic than (Ind) in that the E^Ms are, unlike the E_is, based on non-overlapping information sets. ((Ind*) holds in particular if the items of information in $\mathcal{I}_1 \cup ... \cup \mathcal{I}_n$ are mutually (conditionally) independent.) For simplicity, suppose finally that

$$P(A) > 0 \text{ for every non-empty event } A \in \mathcal{A}. \tag{14}$$

Theorem 3 *Assume (Ind*) and (14). Then*

(a) *(Ind) holds if and only if (NoPIO) holds;*

(b) *more precisely, if $E^M \neq \Omega$ for proper subgroup M, then the personal evidences E_i, $i \in M$, are pairwise positively correlated conditional on at least one $H \in \mathcal{H}$ (i.e. $P(E_i \cap E_j|H) > P(E_i|H)P(E_j|H)$ for any two distinct $i, j \in M$).*

Proof. I prove part (a); the proof includes a proof of part (b).

(i) First, assume (NoPIO). Then we have, for all persons i,

$$E_i = \cap_{\{i\} \subseteq M \subseteq N} E^M = E^{\{i\}} \cap \left[\cap_{\{i\} \subseteq M \subseteq N \& |M| \geq 2} E^M\right] = E^{\{i\}} \cap \Omega = E^{\{i\}}. \tag{15}$$

Conditional on any $H \in \mathcal{H}$, by (Ind*) the events E^M, $\emptyset \neq M \subseteq N$, are independent, hence so are $E^{\{1\}}, ..., E^{\{n\}}$, and hence so are $E_1, ..., E_n$ by (15).

(ii) Now assume (NoPIO) is violated, and let M^* be a proper subgroup with $E^{M^*} \neq \Omega$. I show that the events E_i, $i \in M^*$, are pairwise positively correlated conditional on at least one $H \in \mathcal{H}$, which proves part (b) and also completes the proof of part (a) since $E_1, ..., E_n$ are then not independent conditional on H. Let $i, j \in M^*$ be distinct. By $E^{M^*} \neq \Omega$ and (14) I have $P(E^{M^*}) < 1$. So there exists an $H \in \mathcal{H}$ with $P(E^{M^*}|H) < 1$. Since $E_i = \cap_{\{i\} \subseteq M \subseteq N} E^M$, we have by (Ind*) $P(E_i|H) = \Pi_{\{i\} \subseteq M \subseteq N} P(E^M|H)$. The analogous argument for j yields $P(E_j|H) = \Pi_{\{j\} \subseteq M \subseteq N} P(E^M|H)$. So

$$P(E_i|H)P(E_j|H) = \left[\Pi_{\{i\} \subseteq M \subseteq N} P(E^M|H)\right] \times \left[\Pi_{\{j\} \subseteq M \subseteq N} P(E^M|H)\right]. \tag{16}$$

Further, we have

$$E_i \cap E_j = [\cap_{\{i\}\subseteq M\subseteq N} E^M] \cap [\cap_{\{j\}\subseteq M\subseteq N} E^M] = [\cap_{\{i\}\subseteq M\subseteq N} E^M] \cap [\cap_{\{j\}\subseteq M\subseteq N\backslash\{i\}} E^M].$$

So, by (Ind*),

$$P(E_i \cap E_j) = [\Pi_{\{i\}\subseteq M\subseteq N} P(E^M)] \times [\Pi_{\{j\}\subseteq M\subseteq N\backslash\{i\}} P(E^M)]. \quad (17)$$

The relations (16) and (17) together entail $P(E_i \cap E_j) > P(E_i|H)P(E_j|H)$, because expression (16) equals expression (17) multiplied with the factor $\Pi_{\{i,j\}\subseteq M\subseteq N} P(E^M)$, which is smaller than 1 since it contains the term $P(E^{M^*}|H) < 1$. ■

7 Recursive opinion pooling

I now turn to opinion pooling in the presence of problematic information overlap. By Theorem 3, a violation of (NoPIO) implies a violation of (Ind) (under (Ind*) and (14)), so that we need a new method to calculate the collective probability function π.

Instead of (NoPIO), assume now more generally that *some* subgroups do and others do not exclusively share information. Specifically, let there be a (more or less large) set of subgroups $\mathcal{M}$ with $N \in \mathcal{M}$, interpreted as the set of subgroups that (potentially) have exclusively shared information. Any subgroup $M \notin \mathcal{M}$ should *not* have exclusively shared information:

(NoPIO*). $E^M = \Omega$ for all subgroups $M \notin \mathcal{M}$.

The practical choice of $\mathcal{M}$ depends on how information is distributed across the group. In last section's jury example, the information is distributed across the $n = 3$ jurors in such a way that $\mathcal{M} = \{\{1\},\{2\},\{3\},\{1,3\},\{2,3\},\{1,2,3\}\}$, as $\{1,2\}$ is the only subgroup with no exclusively shared information. Condition (NoPIO*) is equivalent to (NoPIO) if $\mathcal{M} = \{\{1\},...,\{n\},N\}$, but becomes weaker (and more interesting) if $\mathcal{M}$ also contains some proper subgroups, and becomes empty if $\mathcal{M}$ contains *all* subgroups.

Information sharing. Having originally only (NoPIO*), one strategy is to "enforce" (NoPIO) through active information sharing prior to aggregation: all proper subgroups $M \in \mathcal{M}$ communicate their exclusively shared information to the rest of the group. In Figure 1, all information in "!" fields is communicated to the third person, and in the above jury example the subgroups $\{1,3\}$ and $\{2,3\}$ communicate the exact content of the first resp. second witness report to the third juror. Having thus removed any problematic information overlap, (NoPIO) now holds and probability aggregation along the above lines can start.

Recursive pooling. As an alternative, let me now develop a general technique of recursive probability aggregation under (NoPIO*). For simplicity, I assume that each subgroup $M \in \mathcal{M}$ agrees on the interpretation of its shared information, i.e. on the resulting probabilities. So, different probabilities result from different information, never from different interpretations of information. But, rather than making this assumption explicit by a condition analogous to the earlier Agreement on the Prior (AP), the assumption is implicit by not indexing p_M by i (see below), and by using P instead of P_i throughout, thereby implicitly assuming that $P_i(E) = P(E)$ for all $E \in \mathcal{A}_i \cap \mathcal{A}$.

While earlier each person i submitted functions $p_i, \pi_i \in \Pi$, assume now that

• each subgroup $M \in \mathcal{M}$ submits a probability function $p_M \in \Pi$, representing M's probability assignments based on *M's shared information* (shared information need not be "exclusively" shared, i.e. may be known to other persons).

The aim is still to derive, based now on the functions $p_M \in \Pi$, $M \in \mathcal{M}$,

• a (collective) probability function $\pi \in \Pi$ representing the group's entire information.

p_M corresponds to the earlier function π_i if $M = \{i\}$, and to a (unanimously agreed) prior function $p_1 = ... = p_n = p$ if $M = N$. What is new are the functions p_M for *proper* subgroups $M \in \mathcal{M}$. In practice, every non-singleton subgroup $M \in \mathcal{M}$ will have to "sit together", find out about its shared information, and agree on the resulting probability function p_M.

Example of the technique. As noted above, in last section's jury example $\mathcal{M} = \{\{1\},\{2\},\{3\},\{1,3\},\{2,3\},\{1,2,3\}\}$. So, functions $p_{\{1\}}, p_{\{2\}}, p_{\{3\}}, p_{\{1,3\}}, p_{\{2,3\}}$ and $p_{\{1,2,3\}}$ are submitted. The recursion works as follows, where I use a slightly simplified version of the later notation and skip all formal justifications:

• First, combine $p_{\{1,3\}}$ and $p_{\{2,3\}}$ into a function $p_{\{1,3\},\{2,3\}}$ representing the union of $\{1,3\}$'s shared information and $\{1,3\}$'s shared information. One may apply Theorem 1's formula and put $p_{\{1,3\},\{2,3\}} \propto p_{\{1,3\}}p_{\{2,3\}}/p_{\{1,2,3\}}$.

• Next, combine $p_{\{1\}}$ and $p_{\{2\}}$ into a function $p_{\{1\},\{2\}}$ representing the union of $\{1\}$'s and $\{2\}$'s information. One may apply Theorem 1's formula and put $p_{\{1\},\{2\}} \propto p_{\{1\}}p_{\{2\}}/p_{\{1,2\}}$, where $p_{\{1,2\}}$ is defined as $p_{\{1,2,3\}}$ because the subgroup $\{1,2\}$ has no exclusively shared information.

• Finally, combine $p_{\{1\},\{2\}}$ and $p_{\{3\}}$ into the function $\pi = p_{\{1\},\{2\},\{3\}}$ representing the union of $\{1\}$'s, $\{2\}$'s and $\{3\}$'s information. Again, one may apply Theorem 1's formula and put $\pi = p_{\{1\},\{2\},\{3\}} \propto p_{\{1\},\{2\}}p_{\{3\}}/p_{\{1,3\},\{2,3\}}$.

Recall that i's evidence E_i equals $\cap_{\{i\}\subseteq M\subseteq N} E^M$. I now generalise this to subgroups:

Definition 6 *A subgroup M's "shared evidence" is defined as* $E_M := \cap_{M\subseteq M'\subseteq N} E^{M'}$.

Note that $E_{\{i\}} = E_i$. E_M represents all information held by *at least* all $i \in M$ – as opposed E^M, M's *exclusively* shared evidence, which represents all information held by *only* all persons in M. Information shared by M is precisely information exclusively shared by some subgroup $M \subseteq M' \subseteq N$. Also, note that

$$P(E^M), P(E_M) > 0, \text{ for each subgroup } M,$$

as $P(E^M), P(E_M) \geq P(\cap_{\emptyset \neq M' \subseteq N} E^{M'}) = P(E_1 \cap ... \cap E_n) > 0$. The following condition translate Individual Bayesian Rationality (IBR) to subgroups in $\mathcal{M}$:

(IBR*) $p_M(H) = P(H|E_M)$ for every subgroup $M \in \mathcal{M}$ and hypothesis $H \in \mathcal{H}$.

The aim is still that the collective probability function satisfies (CBR), i.e. that

$$\pi(H) = P(H|E_1 \cap ... \cap E_n) \text{ for each hypothesis } H \in \mathcal{H},$$

a condition that may be rewritten in several equivalent ways since (by Definition 6)

$$E_1 \cap ... \cap E_n = E_{\{1\}} \cap ... \cap E_{\{n\}} = \cap_{\emptyset \neq M \subseteq N} E^M = \cap_{\emptyset \neq M \subseteq N} E_M.$$

To calculate π satisfying (CBR), I now introduce "abstract individuals":

Definition 7 *An "abstract individual", A, is a non-empty set of subgroups M.*

I interpret an abstract individual A as a hypothetical agent whose information contains the shared information of *each* subgroup $M \in A$. For instance, $A = \{\{1,3\},\{2,3\}\}$ "knows" $\{1,3\}$'s shared information *and* $\{2,3\}$'s shared information. A's evidence is thus given by $\cap_{M \in A} E_M$. I will calculate for each abstract individual A a function $p_A \in \Pi$ reflecting precisely A's information $\cap_{M \in A} E_M$, i.e. such that

$$p_A(H) = P(H| \cap_{M \in A} E_M) \text{ for each } H \in \mathcal{H}. \tag{18}$$

Definition 8 *The "order" of an abstract individual A is $order(A) := \min\{|M| : M \in A\}$, the size of a smallest subgroup in A.*

I calculate p_A by (backward) recursion over $order(A)$: p_A is calculated first for $order(A) = n$, then for $order(A) = n-1$, ..., then for $order(A) = 1$. This finally yields π, since by (CBR) and (18) $\pi = P(.|E_{\{1\}} \cap ... \cap E_{\{n\}}) = p_A$ where A is the abstract individual $\{\{1\},\{2\},...,\{n\}\}$ of order 1. The induction step will consist essentially in calculating, from the functions p_A and p_{A^*} of two abstract individuals A and A^* the function $p_{A \cup A^*}$ of the abstract individual $A \cup A^*$ whose information is the combined information of A and A^*. To derive $p_{A \cup A^*}$ from p_A and p_{A^*}, I generalise the formula of Theorem 1 to (two) *abstract* individuals. In that formula, the notion of shared information is crucial. What information do A and A^* share? They share precisely the information of the abstract individual

$$A \wedge A^* := \{M \cup M^* : M \in A \text{ and } M^* \in A^*\}.$$

For A and A^* share that information shared *both* by a subgroup $M \in A$ *and* by a subgroup $M^* \in A^*$, i.e. shared by subgroup $M \cup M^*$. So, when combining p_A and p_{A^*}, $A \wedge A^*$'s function $p_{A \wedge A^*}$ plays the role of the "shared prior" p in Theorem 1. Specifically:

Lemma 1 *Assume (Ind*). For any two abstract individuals A and A^*, if each of $p_A, p_{A^*}, p_{A\wedge A^*} \in \Pi$ satisfies (18) and $p_{A\cup A^*} \in \Pi$ is given by*

$$p_{A\cup A^*}(H) = \frac{p_A(H)p_{A^*}(H)/p_{A\wedge A^*}(H)}{\sum_{H'\in\mathcal{H}} p_A(H')p_{A^*}(H')/p_{A\wedge A^*}(H')}, \text{ in short } p_{A\cup A^*} \propto p_A p_{A^*}/p_{A\wedge A^*},$$

then $p_{A\cup A^*}$ also satisfies (18).

Proof. Let $p_A, p_{A^*}, p_{A\wedge A^*} \in \Pi$ each satisfy (18) and $p_{A\cup A^*} \in \Pi$ be given by $p_{A\cup A^*} \propto p_A p_{A^*}/p_{A\wedge A^*}$. For all abstract individuals B,

$$\overline{B} := \{M \subseteq N : M' \subseteq M \text{ for some } M' \in B\},$$

the set of supergroups of subgroups in B. By (18), $p_{A\wedge A^*} = P(.|\cap_{M\in A\wedge A^*} E_M)$, where by Definition 6

$$\cap_{M\in A\wedge A^*} E_M = \cap_{M\in A\wedge A^*} \cap_{M\subseteq M'\subseteq N} E^{M'} = \cap_{M\in\overline{A\wedge A^*}} E^M.$$

So,

$$p_{A\wedge A^*} = P(.|\widetilde{E}) \text{ with } \widetilde{E} := \cap_{M\in\overline{A\wedge A^*}} E^M. \tag{19}$$

Analogously, by (18), $p_A = P(.| \cap_{M\in A} E_M)$, where by Definition 6

$$\cap_{M\in A} E_M = \cap_{M\in A} \cap_{M\subseteq M'\subseteq N} E^{M'} = \cap_{M\in\overline{A}} E^M = E \cap \widetilde{E}$$

with $E := \cap_{M\in\overline{A}\setminus\overline{A\wedge A^*}} E^M$. So $p_A = P(.|E\cap\widetilde{E})$, and hence by Bayes' rule

$$p_A \propto P(.|\widetilde{E})P(E|. \cap \tilde{E}). \tag{20}$$

By an analogous argument for A^*, we have

$$p_{A^*} \propto P(.|\widetilde{E})P(E^*|. \cap \tilde{E}), \tag{21}$$

where $E^* := \cap_{M\in\overline{A^*}\setminus\overline{A\wedge A^*}} E^M$. Since $p_{A\cup A^*} \propto p_A p_{A^*}/p_{A\wedge A^*}$, we have by (19), (20) and (21)

$$\begin{aligned} p_{A\cup A^*} &\propto \left[P(.|\widetilde{E})P(E|. \cap \tilde{E})\right]\left[P(.|\widetilde{E})P(E^*|. \cap \tilde{E})\right]/P(.|\widetilde{E}) \\ &= P(.|\widetilde{E})P(E|. \cap \tilde{E})P(E^*|. \cap \tilde{E}). \end{aligned} \tag{22}$$

(Ind*) implies that, for each $H \in \mathcal{H}$, the events $E, E^*, \widetilde{E}$ are independent given H, and hence E, E^* are independent given $H \cap \widetilde{E}$. So

$$P(E|. \cap \tilde{E})P(E^*|. \cap \tilde{E}) = P(E \cap E^*|. \cap \tilde{E}).$$

Substituting this into (22) and then applying Bayes' rule, I obtain

$$p_{A\cup A^*} \propto P(.|\widetilde{E})P(E\cap E^*|. \cap \tilde{E}) \propto P(.|E\cap E^* \cap \widetilde{E}).$$

But the latter equals $P(.| \cap_{M\in A\cup A^*} E_M)$, as by Definition 6

$$\cap_{M\in A\cup A^*} E_M = \cap_{M\in A\cup A^*} \cap_{M\subseteq M'\subseteq N} E^M = \cap_{M\in\overline{A\cup A^*}} E^M = E\cap E^* \cap \widetilde{E}. \blacksquare$$

Algorithm. Define $p_A \in \Pi$ for abstract individuals A by backward recursion on $order(A)$:

- Let $order(A) = n$. Then $A = \{N\}$. Define $p_A := p_N$.
- Let $order(A) = k < n$ and assume $p_{A'}$ is already defined for $order(A') > k$.

  Case 1: $|A| = 1$. Then $A = \{M\}$. If $M \in \mathcal{M}$, define $p_A = p_M$. If $M \notin \mathcal{M}$, consider the abstract individual $A' := \{M \cup \{i\} : i \notin M\}$ containing all subgroups with exactly one person added to M (*interpretation*: A and A' have the same information by $M \notin \mathcal{M}$); define $p_A := p_{A'}$, where $p_{A'}$ is already defined by $order(A') = k + 1$.

  Case 2: $|A| > 1$. Define p_A by recursion on $|\{M \in A : |M| = k\}|$, the number of subgroups in A of size k:

  - Let $|\{M \in A : |M| = k\}| = 1$. Then $A = \{M\} \cup A^*$, where $|M| = k$ and $order(A^*) > k$. Define

    $$p_A \propto p_{\{M\}} p_{A^*} / p_{\{M\} \wedge A^*},$$

    where $p_{\{M\}}$ is already defined in case 1, and p_{A^*} and $p_{\{M\} \wedge A^*}$ are already defined by $order(A^*) > k$ and $order(\{M\} \wedge A^*) > k$.
  - Let $|\{M \in A : |M| = k\}| = l > 1$ and assume p_{A^*} is already defined for $|\{M \in A^* : |M| = k\}| < l$ (and $order(A^*) = k$). Then $A = \{M\} \cup A^*$ with $|M| = k$ and $|\{M^* \in A^* : |M^*| = k\}| = l - 1$. Define

    $$p_A \propto p_{\{M\}} p_{A^*} / p_{\{M\} \wedge A^*},$$

    where $p_{\{M\}}$ is already defined in case 1, p_{A^*} is already defined by $|\{M^* \in A^* : |M^*| = k\}| = l - 1$, and $p_{\{M\} \wedge A^*}$ is already defined by $order(\{M\} \wedge A^*) > k$.

Now let me prove that this algorithm yields the desired result.

Theorem 4 *Assume (IBR*), (Ind*) and (NoPIO*), and define the functions $p_A \in \Pi$ (for abstract individuals A) by the above algorithm. Then the functions p_A satisfy (18). So, assuming (CBR), the collective probability function π equals $p_{\{\{1\},\ldots,\{n\}\}}$.*

Proof. Assume (IBR*), (Ind*) and (NoPIO*). Denote by $\mathcal{A}$ the set of abstract individuals A. Define p_A, $A \in \mathcal{A}$, by the above algorithm. I prove that (18) (i.e. $p_A = P(.| \cap_{M \in A} E_M)$) holds for all $A \in \mathcal{A}$, using (backward) induction on $order(A)$.

- If $order(A) = n$, then $A = \{N\}$, and by definition $p_A = p_N$. So by (IBR*) $p_A = P(.|E_N) = P(.| \cap_{M \in A} E_M)$.
- Now let $order(A) = k < n$, and assume (18) holds for all $A' \in \mathcal{A}$ with $order(A') > k$. I have to show that $p_A = P(.| \cap_{M \in A} E_M)$.

  Case 1: $|A| = 1$. Then $A = \{M\}$ with $|M| = k$. If $M \in \mathcal{M}$, then by definition $p_A = p_M$, so by (IBR*) $p_A = P(.|E_M) = P(.| \cap_{M' \in A} E_{M'})$. Now assume

$M \notin \mathcal{M}$. Then by definition $p_A = p_{A'}$ with $A' := \{M \cup \{i\} : i \notin M\}$. Since $order(A') = k+1$, the induction hypothesis yields $p_{A'} = P(.| \cap_{M' \in A'} E_{M'})$, hence $p_A = P(.| \cap_{M' \in A'} E_{M'})$. So I have to show that $\cap_{M' \in A'} E_{M'} = E_M$. By Definition 6,

$$E_M = \cap_{M \subseteq M' \subseteq N} E^{M'} = E^M \cap \left\{ \cap_{M' \in A'} \left[\cap_{M' \subseteq M'' \subseteq N} E^{M''} \right] \right\}.$$

In this, $E^M = \Omega$ by (NoPIO*) and $\cap_{M' \subseteq M'' \subseteq N} E^{M''} = E_{M'}$ by Definition 6. So $E_M = \cap_{M' \in A'} E_{M'}$, as desired.

Case 2: $|A| > 1$. I show $p_A = P(.| \cap_{M \in A} E_M)$ by induction on the number $|\{M \in A : |M| = k\}|$ of subgroups in A of size k.

- Let $|\{M \in A : |M| = k\}| = 1$. Then by definition $p_A \propto p_{\{M\}} p_{A^*} / p_{\{M\} \wedge A^*}$, where $A = \{M\} \cup A^*$ with $|M| = k$ and $order(A^*) > k$. Now, $p_{\{M\}}$ satisfies (18) by "case 1", and p_{A^*} and $p_{\{M\} \wedge A^*}$ satisfy (18) by $order(A^*) > k$ and $order(\{M\} \wedge A^*) > k$ (and the k-induction hypothesis). So, by Lemma 1, p_A satisfies (18).
- Let $|\{M \in A : |M| = k\}| = l > 1$, and assume A^* satisfies (18) whenever $|\{M \in A^* : |M| = k\}| < l$ (and $order(A^*) = k$). By definition, $p_A \propto p_{\{M\}} p_{A^*} / p_{\{M\} \wedge A^*}$, where $A = \{M\} \cup A^*$ with $|M| = k$ and $|\{M^* \in A^* : |M^*| = k\}| = l - 1$. In this, $p_{\{M\}}$ satisfies (18) by "case 1", p_{A^*} satisfies (18) by $|\{M^* \in A^* : |M^*| = k\}| = l - 1$ (and the l-induction hypothesis), and $p_{\{M\} \wedge A^*}$ satisfies (18) by $order(\{M\} \wedge A^*) > k$ (and the k-induction hypothesis). So, by Lemma 1, p_A satisfies (18). ■

8 Conclusion

The above model addresses probability aggregation in the presence of informational asymmetry. In an attempt to free opinion pooling from open parameters, the informational asymmetry was accounted for *not* by using an aggregation rule that assigns information-specific weights to persons, *but* by asking people to agree on a shared ("prior") probability function p based on the shared information resp. to submit individual functions p_i if there is disagreement over how to interpret the shared information. Specifically, while the standard approach defines the collective probability function π as $F(\pi_1, ..., \pi_n)$ for some opinion pool F, I proposed defining π (in the absence of *problematic information overlap*) by $\pi \propto \pi_1 \cdots \pi_n / p^{n-1}$ (which is parameter-free), resp. by $\pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n)$ if there is disagreement over the prior, where F is a standard opinion pool that involves no information-specific parameters (weights); if F furthermore involves no *competence*-specific parameters (say, because competence levels are unknown or should be ignored for reasons of procedural fairness), then F and hence π may again be parameter-free. In the case of a *problematic information overlap*, information pooling was achieved by a somewhat more complicated algorithmic approach (see Theorem 4).

Overall, compared to earlier approaches, this approach is more demanding on the side of the individuals (due to the additional agreement on p resp. submission of $p_1, ..., p_n$) but typically less demanding on the side of choosing the aggregation rule (as fewer or no parameters need be chosen). The reduced arbitrariness in the procedure choice increases not only practical feasibility but also group autonomy.

Still, the practical use of the model rests on assumptions. It is a *normative* question whether to accept Collective Bayesian Rationality (CBR) and Acceptance of Likelihoods (AL); they may be accepted on epistemic grounds. By contrast, it is a *factual* matter whether Individual Bayesian Rationality (IBR)/(IBR*) and Independence (Ind)/(Ind*) hold. (Ind) is threatened by the possibility of a *problematic information overlap* (see Theorem 3), while (Ind*) is not. The technique of recursive opinion pooling presented in Section 7 can cope with any type of information overlap, but it assumes that each subgroup in $\mathcal{M}$ agrees on how to interpret its shared information. Dropping this assumption would have gone beyond the scope of this paper – but it clearly is an interesting question for future research. The algorithm can possibly be generalised in more than one way.

9 References

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

Fitelson, B. (2001) A Bayesian account of independent evidence with applications, *Philosophy of Science* 68 (Proceedings). S123 - S140

Genest, C. (1984) A characterization theorem for externally Bayesian groups, *Ann. Statist.* 12, p. 1100-1105

Genest, C. and J. V. Zidek (1986) Combining probability distributions: a critique and an annotated bibliography, *Statist. Sci.* 1, p. 114-148

Genest, C., K. J. McConway and M. J. Schervish (1986) Characterization of externally Bayesian pooling operators, *Ann. Statist.* 14, 487-501

Hild, M. (1998) The instability of ex post aggregation, *Typescript*

Hylland, A. and R. Zeckhauser (1979) The impossibility of group decision making with separate aggregation of beliefs and values, *Econometrica* 47, p. 1321-36

Jeffrey, R. (1983) (first published 1965) *The logic of decision*, Chicago: Chicago University Press

Levi, I. (1990) Pareto-unanimity and consensus, *Journal of Philosophy* 87

McConway (1978) The combination of experts' opinions in probability assessments: some theoretical considerations, *Ph.D. thesis*, University College London

McConway (1981) Marginalization and linear opinion pools, *Jour. Amer. Statist. Assoc.* 76, p. 410-414

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

Mongin, P. (1998) The paradox of the Bayesian experts and state-dependent utility theory, *Journal of Mathematical Economics* 29, p. 331-61

Morris, P. A. (1974) Decision analysis expert use, *Management Science* 20, p. 1233-41

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

Risse, M. (2001) Instability of ex post aggregation in the Bolker/Jeffrey framework and related instability phenomena, *Erkenntnis* 55, p. 239-269

Risse, M. (2003) Bayesian group agents and two modes of aggregation, *Synthese*, forthcoming

Savage, L. (1954) *The foundations of statistics*, New York: Wiley

Schervish, M., T. Seidenfeld and J. Kadane (1991) Shared preferences and state-dependent utilities, *Management Science* 37, p. 1575-89

Seidenfeld, T., J. Kadane and M. Schervish (1989) On the shared preferences of two Bayesian decision makers, *Journal of Philosophy* 86, p. 221-44

Wagner, C. G. (1982) Allocation, Lehrer models, and the consensus of probabilities, *Theory and Decision* 14, p. 207-220