Franz Dietrich

Bayesian group belief

Working paper

Bayesian Group Belief

Franz Dietrich[1]
May 2008

Abstract. If a group is modelled as a single Bayesian agent, what should its beliefs be? I propose an axiomatic model that connects group beliefs to beliefs of group members, who are themselves modelled as Bayesian agents and, crucially, may have different information. They may also have different prior beliefs and different domains (σ-algebras) on which they hold beliefs, to account for differences in awareness and conceptualisation. As is shown, group beliefs can incorporate all information spread across individuals without individuals having to communicate their information (which may be complex, hard-to-describe, or not describable in principle due to language restrictions); individuals should instead communicate their prior and posterior beliefs. The group beliefs derived here take a simple multiplicative form if people's information is independent (and a more complex form if information overlaps arbitrarily), which contrast with familiar linear or geometric opinion pooling and the (Pareto) requirement of respecting unanimous beliefs. *JEL classification*: D70, D71

Keywords: Opinion pooling, axiomatic social choice theory, subjective probability

1 Introduction

Suppose a group is interested in whether a given hypothesis H is true. If every individual assigns a probability of 70% to H, what probability should the group as a whole 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% belief, following the popular (probabilistic) Pareto principle (e.g. Mongin (1995, 1998)). Under informational asymmetry, by contrast, a possibly much higher or lower collective probability may be appropriate, and the Pareto principle becomes problematic, or so I argue.

The above question is an instance of the classic *opinion pooling/aggregation* problem, with applications for instance in expert panels. In general, individual probabilities need of course not coincide, and also more than one hypothesis may be of interest.

[1] *Affiliations*: London School of Economics & Maastricht University. This paper is based on my old unpublished working paper 'Opinion Pooling under Asymmetric Information,' *Public Economics 0407002, EconWPA*, 2004. Interesting related results were meanwhile obtained independently by Marcus Pivato in his working paper 'The Discursive Dilemma and Probabilistic Judgement Aggregation,' *MPRA Paper 8412*, University Library of Munich, Germany, 2008.

The goal is to merge a profile of n individual probability measures (on a σ-algebra of events) into a single collective probability measure. The literature has proposed different normative conditions on the aggregation rule, and has derived the class of rules satisfying these conditions. The two most prominent types of rules are *linear* and *geometric* rules. Denoting by $\pi_1, ..., \pi_n, \pi$ the restrictions of the individual resp. collective probability measures to the set $\mathcal{H}$ of atoms of the σ-algebra (where I assume that $\mathcal{H}$ forms a countable partition or the underlying state space, so that it suffices to consider the probabilities of the atoms), a linear rule defines π as being a weighted *arithmetic* average $\sum_{i=1}^{n} w_i \pi_i$, and a geometric rule defines π as being proportional to a weighted *geometric* average $\Pi_{i=1}^{n} \pi_i^{w_i}$, where $w_1, ..., w_n \in [0, 1]$ are fixed weights with sum 1. By contrast, our Bayesian axioms will lead to what I call *multiplicative* rules, which define π as proportional to $g\Pi_{i=1}^{n} \pi_i$, the product of all (unweighted) individual functions π_i with some fixed function g. Linear rules have been characterised (under additional technical assumptions) by the *independence* or *setwise function property* (McConway (1981), Wagner (1982, 1985), Dietrich and List (2007); see also Lehrer and Wagner (1981)), the *marginalisation property* (McConway (1981)), and (in a single-profile context) by the probabilistic analogue of the *Pareto principle* (Mongin, (1995, 1998)); and geometric rules famously satisfy *external Bayesianity* as defined in Section 6 (e.g. McConway (1978), Genest (1984), Genest, McConway and Schervish (1986)). Still an excellent reference for fundamental results on opinion pooling is Genest and Zidek's (1986) literature review.

I claim that the classic approach is problematic if, as in this paper, the goal of opinion pooling is taken to be *information aggregation*, i.e. if collective beliefs aim at incorporating all the information spread asymmetrically over the individuals. The classic approach is more suitable if the goal is not information aggregation: the goal might be not epistemic at all (e.g. fair representation), or it might be epistemic yet with the disagreements between individuals caused not by differences in information but by differences in interpretation of the same shared body of information.

One might at first suspect that classic pooling functions can account for informational asymmetries by putting more weight on the beliefs of well-informed individuals. More concretely, it is often suggested that in a linear and geometric rule (as defined above) the weights w_i of well-informed individuals should be higher. However, 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).

To concretely illustrate the difficulty that classic pooling functions have in aggregating information, consider again the introductory example. Suppose each individual i's subjective probability $\pi_i(H) = .7$ is in fact the result of Bayesian conditionalisation on some private information E_i, where the E_is are independent across individuals. What should the collective belief $\pi(H)$ be? If the individuals all had the same *prior* probability of H, say p_0, all depends on how p_0 compares to .7: if $p_0 < .7$ then

$\pi(H)$ should intuitively exceed .7 because $\pi(H)$ should incorporate all the observations $E_1, ..., E_n$, each one of which alone already suffices to push the probability of H up from p_0 to .7. By a similar argument, if H has a common prior above .7 then intuitively $\pi(H) < .7$, and if H has a common prior of exactly .7 then intuitively $\pi(H) = .7$. If people hold *different* prior beliefs of H, then intuitively $\pi(H)$ should be higher or lower than .7 according to whether 'most' individuals' prior of H is lower resp. higher than .7. These considerations highlight that knowing just the individuals' current (i.e. posterior) opinions $\pi_1, ..., \pi_n$ does not suffice to determine a collective opinion π that efficiently aggregates private information. So our model will have to deviate from standard opinion pooling in that π will not be a function of $\pi_1, ..., \pi_n$ alone. On what else must collective opinions depend? The example lets us suspect that individual *prior* beliefs matter.

The paper confirms this intuition generally, by presenting an axiomatic framework that unlike the classic approach explicitly models the information states of the individuals. The imposed axioms lead (in the common prior case) to a unique formula for the collective probability function; no weights or other parameters are needed to incorporate all individual information into the collective beliefs. For the reason explained above, the collective beliefs depend not just on people's actual (i.e. posterior) beliefs but also their prior beliefs. This increased individual input is necessary and sufficient to efficiently aggregate information, which might come as a surprise. In short, knowing the (complex) content of people's private information is not needed: knowing people's prior-posterior pairs suffices.

As an alternative to our approach, the *supra-Bayesian* approach might also be able to aggregate information efficiently; however, despite conceptual elegance, the approach suffers from some problems, among which practicable infeasibility.[2]

In modelling both individuals and the collective as Bayesian rationals, our findings are also 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, see Risse (2003).

Section 2 presents the axiomatic model and derives the resulting aggregation rule. Section 3 gives a numerical example. Section 4 identifies our pooling formula as a form of multiplicative opinion pooling. Sections 5 and 6 address the case of no

[2] In the supra-Bayesian approach (introduced by Morris' (1974) seminal work and extended in a large literature), collective beliefs are obtained as *posterior* probabilities (held by the real or virtual 'supra-Bayesian') conditional on the observed individual beliefs (treated as random events or evidence). This presupposes knowing (i) prior probabilities, and (ii) the likelihoods with which the individuals make probability assignments. It is not clear where these prior probabilities and likelihoods can come from; reaching a compromise or consensus on them might involve a more complex opinion pooling problem than the original one.

common prior. Section 7 analyses the independent-information assumption made so far. Section 8 generalises the aggregation rule to arbitrary information overlaps.

2 An axiomatic model

Consider a group of persons $i = 1, ..., n$ ($n \geq 2$) who need collective beliefs on certain *hypotheses*, represented as subsets H of a non-empty set Ω of *possible worlds*, i.e. worlds that are possible under the shared information. Throughout I call information (knowledge, an observation etc.) 'shared' if it is held by all group members. Let $\mathcal{H}$ be the set of hypotheses $H \subseteq \Omega$ of interest, where $\mathcal{H}$ forms a finite or countably infinite partition of Ω and $\emptyset \notin \mathcal{H}$. So, the hypotheses are mutually exclusive and exhaustive. A simple but frequent case is a binary problem $\mathcal{H} = \{H, \Omega\backslash H\}$, where H might be the hypothesis that the defendant in a court trial is guilty. In a non-binary case, $\mathcal{H}$ might contain different hypotheses on the defendant's extent of guilt.

I call an *opinion* (*on* $\mathcal{H}$) any function $f : \mathcal{H} \to (0, 1]$ with $\sum_{H\in\mathcal{H}} f(H) = 1$ (whereas *probability measures* are, as usual, defined on σ-algebras of events[3]); let Π be the set of all these functions f.

Let each individual i hold an opinion $\pi_i \in \Pi$, and let the collective also hold an opinion $\pi \in \Pi$. So far, this is entirely classical. Classical opinion pooling would proceed by placing conditions on how π depends on $\pi_1, ..., \pi_n$, resulting in a unique relationship (e.g. $\pi = \frac{1}{n}\pi_1 + ... + \frac{1}{n}\pi_n$) or a class of possible relationships (e.g. all linear relationships).

Before stating the axiomatic approach in full generality (that is, before allowing individuals to hold different prior beliefs defined on different domains of events), I sketch the approach in a simple case. Suppose for the moment that any individual i's opinion $\pi_i : \mathcal{H} \to [0, 1]$ is given by

$$\pi_i(H) = P(H|E_i) \text{ for all } H \in \mathcal{H},$$

where P is a common prbability measure defined on the maximal domain $\mathcal{P}(\Omega)$ (so that people hold prior beliefs about *everything*, the *same* ones!), and $E_i \subseteq \Omega$ is individual i's private information with $P(E_i) > 0$, where $E_1, ..., E_n$ are independent conditional on any hypothesis $H \in \mathcal{H}$. We would like the group opinion $\pi : \mathcal{H} \to [0, 1]$ to include all information spread over the group, i.e., to be given by

$$\pi(H) = P(H|E_1 \cap ... \cap E_n) \text{ for all } H \in \mathcal{H} \tag{1}$$

(where one easily checks that (1) is well-defined, i.e., that $P(E_1 \cap ... \cap E_n) > 0$). One approach would be to ask all individuals i to 'tell' their private experience E_i, so that

[3] By countability of $\mathcal{H}$ and σ-additivity of probability measures, any opinion on $\mathcal{H}$ uniquely extends to a probability measure on the σ-algebra $\sigma(\mathcal{H})$ generated by $\mathcal{H}$, and so we lose nothing by considering functions on $\mathcal{H}$ rather than on $\sigma(\mathcal{H})$. By definition, opinions $f \in \Pi$ never assign zero probability to any hypothesis; this is mainly for technical convenience.

the group could simply gather all experiences and calculate the conditional probabilities (1). But this procedure may be unrealistic, as personal experience may be very complex and hard-to-communicate in normal language and limited time. (Another problem which the current preliminary setting assumes away is that person i's experience E_i may be an event of which the other persons have no prior beliefs, or even no awareness or conceptualisation; asymmetries in awareness or conceptualisation might indeed explain *why* different people make different experiences.) Assuming that private evidence cannot (or is not) communicated, can the beliefs in (1) be calculated at all? The following derivation gives a positive answer. Consider a hypothesis $H \in \mathcal{H}$ and the belief $\pi(H)$ as defined by (1). Applying Bayes' rule and then our independence assumption,

$$\begin{aligned}\pi(H) &= \frac{P(H)P(E_1 \cap ... \cap E_n|H)}{\sum_{H' \in \mathcal{H}} P(H')P(E_1 \cap ... \cap E_n|H')} \\ &= \frac{P(H)P(E_1|H) \cdots P(E_n|H)}{\sum_{H' \in \mathcal{H}} P(H')P(E_1|H') \cdots P(E_n|H')}.\end{aligned}$$

In the numerator and the denominator, each factor of type $P(E_i|H)$ can be rewritten according to

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

Substituting this expression, we obtain

$$\begin{aligned}\pi(H) &= \frac{P(H)\frac{\pi_1(H)P(E_1)}{P(H)} \cdots \frac{\pi_n(H)P(E_n)}{P(H)}}{\sum_{H' \in \mathcal{H}} P(H')\frac{\pi_1(H')P(E_1)}{P(H')} \cdots \frac{\pi_n(H')P(E_n)}{P(H')}} \\ &= \frac{\pi_1(H) \cdots \pi_n(H)/P(H)^{n-1}}{\sum_{H' \in \mathcal{H}} \pi_1(H') \cdots \pi_n(H')/P(H')^{n-1}}.\end{aligned}$$

Interestingly, any private information E_i has dropped out altogether, so that the collective opinion π can be calculated solely on the basis the revealed individual opinions $\pi_1, ..., \pi_n$ (and the fixed prior). Put differently, each individual information E_i has been incorporated without disclosing it. In short, denoting by p the prior opinion $P|_{\mathcal{H}}$ (i.e., the restriction of P to the hypotheses of interest), we have shown that

$$\pi \propto \pi_1 \cdots \pi_n / p^{n-1}.$$

Here and throughout, I call functions $f, g : \mathcal{H} \to \mathbf{R}$ *proportional*, written $f \propto g$, if there exists a constant $k \neq 0$ such that $f(H) = kg(H)$ for all $H \in \mathcal{H}$.

After this preliminary analysis, let us start again from the beginning, this time in full generality. Recall that we consider individual opinions $\pi_1, ..., \pi_n \in \mathcal{H}$ and a collective opinion $\pi \in \mathcal{H}$. The further elements introduced in the preliminary analysis (namely, $P, E_1, ..., E_n$) are now re-introduced in their general and official form. For each person i let there be:

- an event $E_i \subseteq \Omega$, i's *personal information*;
- a set of events $\mathcal{A}_i$ ($\supseteq \mathcal{H} \cup \{E_i\}$), a σ-algebra on Ω, representing the domain within which i holds beliefs (whereas on events outside $\mathcal{A}_i$ the individual may lack a belief, or even lack awareness or conceptualisation);
- a ('prior') probability measure $P_i : \mathcal{A}_i \to [0, 1]$ representing i's beliefs based on the *shared* information (hence prior to observing E_i), where $P_i(E_i) > 0$ and $P_i(H) > 0$ for all $H \in \mathcal{H}$.[4]

These model resources allow us to state a standard rationality condition:

Individual Bayesian Rationality (IBR) $\pi_i(H) = P_i(H|E_i)$ for each person i and hypothesis $H \in \mathcal{H}$.[5]

Note in particular that a person i's belief domain $\mathcal{A}_i$ may fail to contain another person j's observation E_j, and this for (at least) two reasons. First, the fact that j but not i observed E_j may be due precisely to j having subjectively conceptualised E_j but i not having done so; juror j in a trial may be the only juror to observe the suspicious smile on the defendant's face *because* the other jurors i do not even know what a suspicious smile would be like. Second, j's information E_i may be so detailed and complex that prior to j observing it belonged not even to j's own belief domain, let alone to i's; that is, it was only while observing E_j that person j extended his prior beliefs to a larger domain $\mathcal{A}_j$ containing E_j.

Following the paradigm of social choice theory, I treat the collective as a separate virtual agent with its own beliefs. While this agent is typically a construction (i.e. there needn't exist any real individual holding these beliefs), the social choice paradigm requires it to be as rational as any real individual.[6] 'Rationality' refers to different things in different contexts (e.g. to transitivity of preferences in Arrowian Arrowian preference aggregation, to von-Neumann-Morgenstern rationality in Harsanyi's Theorem on group preferences over lotteries, to logical consistency in judgment aggregation, and so on). In the present context, it naturally refers to Bayesian rationality. To formulate this, I suppose that there are

- a σ-algebra $\mathcal{A}$ ($\supseteq \mathcal{H} \cup \{E_1, ..., E_n\}$) on Ω, representing the domain within which the collective holds beliefs:

[4]The term "prior" need not have a temporal meaning: the observation of E_i need not come after that of shared information.

[5]The conditional probability $P_i(H|E_i)$ is well-defined because $E_i, H \in \mathcal{A}_i$ and $P_i(E_i) > 0$. Our assumptions also take care that all other conditional probabilities used in this paper are well-defined.

[6]The collective agent should be rational notably because it forms the basis for collective actions and decisions.

- a (*prior*) probability measure $P : \mathcal{A} \to [0,1]$ meant to represent the collective beliefs based on people's *shared* information (i.e. not on their personal information), where $P(E_1 \cap ... \cap E_n) > 0$ and $P(H) > 0$ for all $H \in \mathcal{H}$.

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

Collective Bayesian Rationality (CBR) $\pi(H) = P(H|E_1 \cap ... \cap E_n)$ for each hypothesis $H \in \mathcal{H}$.

Condition (CBR) requires the collective opinion π to incorporate all information spread over people: the *shared* information (contained in the prior P) and all *personal* information (contained in $E_1, ..., E_n$).

While we have ensured, via (CBR), that collective beliefs 'use' all evidence scattered across individuals, we have done nothing so far to constrain the collective prior probability measure P (which underlies π). Indeed, P may so far be totally disconnected from the individual prior probability measures $P_1, ..., P_n$ (which underlie $\pi_1, ..., \pi_n$). The next condition does something to connect P to $P_1, ..., P_n$. More precisely, it fixes the likelihoods of the individual evidences $E_1, ..., E_n$ (that is, their probabilities conditional on any hypotheses H) by tying these likelihoods to the individuals' own likelihood assignments:

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 information E_i as given by i's likelihood assignments $P_i(E_i|H)$, $H \in \mathcal{H}$. To motivate this condition, let me first explain the context in a little more detail. In statistics, the information that data contain on given hypotheses (as opposed to prior beliefs on these hypotheses) is usually taken to be summarised in the data's likelihood function, which maps any hypothesis to the data's probability given this hypothesis. For instance, the information on humidity contained in a temperature measurement of 20 degrees Celcius is given by the mapping that assigns to each potential humidity level the probability that temperature is 20 degrees Celcius given this humidity level. In our case, the information contained in individual i's evidence E_i is summarised in E_i's likelihood function, mapping any hypothesis H to E_i's probability given H. But how large exactly is E_i's probability given H? For instance, how probable is it that the defendant in a trial has a particular facial expression (E_i) given the hypothesis that he is guilty (H)? The answer may be far from trivial, as one might come up with various different interpretations of the same observation. Condition (AL) requires that the answer that the collective gives matches the answer that the individual *who observed the evidence* gives; that is, $P(E_i|H) = P_i(E_i|H)$. What is the motivation behind identifying $P(E_i|H)$ with $P_i(E_i|H)$? Why not also take other persons' interpretations of E_i into account by defining $P(E_i|H)$ as some compromise of $P_1(E_i|H), ..., P_n(E_i|H)$? First, for reasons explained above, the persons $j \neq i$ may not even hold beliefs on the unobserved event E_i (i.e., $E_i \notin \mathcal{A}_j$), in which case $P_j(E_i|H)$ is simply undefined. Second, assuming that the persons $j \neq i$ do hold such beliefs (i.e., $E_i \in \mathcal{A}_j$), a 'likelihood compromise' could only be formed after each person j reveals $P_j(E_i|H)$; which in turn supposes that first i communicates his informational basis E_i in all detail to the rest of the group. This is not only at odds with the present approach, but may also be infeasible: given the possible complexity of E_i and the limitations of language, time, i's ability to describe E_i, j's ($j \neq i$) ability to understand E_i, and so on, 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 only approximates j's likelihood of the true E_i ($P_j(\tilde{E}_i) \approx P_j(E_i)$).

The next assumption is not a normative condition but rather an assumption on the environment: individuals receive *independent* information. This assumption will be analysed (and relaxed) in later sections. For now, I only mention that it is strong but very common; it is for instance analogous to independent-private-information assumptions often made for Bayesian games, to the independence assumption in the literature on the Condorcet Jury Theorem, to the *Parental Markov Condition* in the theory of Bayesian networks (interpreting the true hypothesis in $\mathcal{H}$ as the parent of each information E_i in a Bayesian network; see Pearl 2000), and to Fitelson's (2001) *confirmational independence* assumption.

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

I denote by $p_1, ..., p_n, p$ the restrictions of the (individual and collective) prior beliefs $P_1, ..., P_n, P$ to the set $\mathcal{H}$ of relevant hypotheses; formally $p_1 := P_1|_{\mathcal{H}}, ..., p_n := P_n|_{\mathcal{H}}, p := P|_{\mathcal{H}}$. So $p_1, ..., p_n, p$ are the prior counterparts of the posterior opinions $\pi_1, ..., \pi_n, \pi$. The pair p_i, π_i represents i's prior and posterior opinions on the relevant hypotheses.

Theorem 1 *If (IBR), (CBR), (AL) and (Ind) hold, the collective opinion π is given by*

$$\pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} p.$$

Proof. Suppose (IBR), (CBR), (AL) and (Ind) hold. We have:

$$\begin{aligned}
\pi(.) &= P(.|E_1 \cap ... \cap E_n) \text{ by (CBR)} \\
&= \frac{P(E_1 \cap ... \cap E_n|.)p(.)}{P(E_1 \cap ... \cap E_n)} \text{ by Bayes' rule} \\
&\propto P(E_1 \cap ... \cap E_n|.)p(.) \\
&= P(E_1|.) \cdots P(E_n|.)p(.) \text{ by (Ind)} \\
&= P_1(E_1|.) \cdots P_n(E_n|.)p(.) \text{ by (AL)} \\
&= \frac{P_1(.|E_1)P_1(E_1)}{p_1(.)} \cdots \frac{P_n(.|E_n)P_n(E_n)}{p_n(.)} p(.) \text{ by Bayes' rule} \\
&\propto \frac{P_1(.|E_1)}{p_1(.)} \cdots \frac{P_n(.|E_n)}{p_n(.)} p(.) \\
&= \frac{\pi_1(.)}{p_1(.)} \cdots \frac{\pi_n(.)}{p_n(.)} p(.) \text{ by (IBR). } \blacksquare
\end{aligned}$$

Two remarks are due.

1. As promised, the collective opinion π is calculated without people having to communicate their arbitrarily complex informational bases E_i or their likelihoods $P(E_i|H)$, $H \in \mathcal{H}$. In practice, all persons i submit their prior-posterior pairs p_i, π_i, and then the collective opinion π is calculated. (The choice of the collective prior p is addressed in Sections 5 and 6.) Compared to standard opinion pooling, we additionally require submission of prior opinions $p_1, ..., p_n$, a complication that enables the incorporation of the individual information $E_1, ..., E_n$ into the collective opinion.

2. Assume a unanimous *posterior* agreement $\pi_1 = ... = \pi_n$ (as in the introduction's example). Then only in special cases does π equal $\pi_1 = ... = \pi_n$, which shows that the unanimity/Pareto principle often required in standard opinion pooling is problematic under informational asymmetries. One such special case is that $\pi_1 = ... = \pi_n = p_1 = ... = p_n = p$, so that none of the personal observations $E_1, ..., E_n$ confirms or disconfirms any hypothesis, i.e., in essence, there is no informational asymmetry.

An important case of Theorem 1 is that where people have managed to agree on how to interpret their shared information, i.e. where they hold a common prior opinion:

Common Prior (CP) $p_1 = ... = p_n = p$ (i.e. the prior probability measures $P_1, ..., P_n, P$ agree on the set $\mathcal{H}$ of relevant hypotheses, though perhaps not elsewhere).

Corollary 1 *If (IBR), (CBR), (AL), (Ind) and (CP) hold, the collective opinion π is given by*

$$\pi \propto \pi_1 \cdots \pi_n / p_1^{n-1}.$$

Let me make two remarks on this corollary.

1. The corollary's formula differs in an important respect from Theorem 1's formula: the parameter p has been eliminated, and so the collective opinion π is fully determined by the individual prior and posterior opinions. By contrast, if (CP) fails, i.e. if the group didn't manage to agree on how to interpret the shared information, Theorem 1's formula does not fully solve the aggregation problem, as we need a way to determine the collective prior p (see Sections 5 and 6).

2. Condition (CP) can in fact be seen as the conjunction of two conditions. The first (descriptive) condition is that $p_1 = ... = p_n$, i.e. all persons i submit the same prior opinion. The second (normative) condition is that the unanimity (or Pareto) principle holds for the *prior* opinions, i.e. if all submit the same prior opinion, this becomes the collective prior opinion. Applying a unanimity condition to prior opinions is far less problematic than doing so for the posterior opinions $\pi_1, ..., \pi_n, \pi$, because prior opinions contain no informational asymmetry.

3 A numerical example for a simple case

Consider the simple case of a binary problem $\mathcal{H} = \{H, \Omega \backslash H\}$ (H and $\Omega \backslash H$ might mean that the defendant in a court trial is guilty resp. innocent, and persons might be jurors). Suppose Common Prior (CP), i.e. $p_1 = ... = p_n = p$. By Theorem 1 (that is, by its corollary), the collective posterior of H is given by

$$\pi^H = \frac{\pi_1^H \cdots \pi_n^H / (p^H)^{n-1}}{\pi_1^H \cdots \pi_n^H / (p^H)^{n-1} + (1 - \pi_1^H) \cdots (1 - \pi_n^H)/(1 - p^H)^{n-1}}, \tag{2}$$

where $p^H := p(H)$, $\pi^H := \pi(H)$ and $\pi_i^H := \pi_i(H)$.[7] For the case of group size $n = 2$, Table 1 contains the values of π^H for all possible combinations of values of p^H, π_1^H, π_2^H in the grid $\{.1, .25, .5, .75, .9\}$. Note how drastically π^H depends on the prior p^H. By shifting p^H below (above) the π_i^Hs, π^H 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^H is somewhere in the middle of the π_i^Hs, π^H may be moderate; intuitively, if $E_1, ..., E_n$ point into different directions, their conjunction need not strongly point into any direction. Rewriting (2) as

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

shows that group belief π^H is a strictly increasing function of individual beliefs $\pi_1^H, ..., \pi_n^H$ for fixed prior p^H, but a strictly decreasing function of p^H for fixed $\pi_1^H, ..., \pi_n^H$ (where $\pi^H \to 1(0)$ as $p^H \to 0(1)$). How can one make sense of the group posterior π^H depending negatively on the prior p^H? Can more prior support for H really reduce H's posterior probability? The answer is that increasing the prior p^H

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

		p^H :				
		.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^H, π_2^H :	.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 $\pi^H = \pi(H)$ in dependence of the common prior $p^H = p(H)$ and the individual posteriors $\pi_i^H = \pi_i(H)$, for a group of size $n = 2$.

while keeping the individual posteriors $\pi_1^H, ..., \pi_n^H$ fixed implicitly reduces the support that each of the n individual observations $E_1, ..., E_n$ give to H; and this ought indeed to reduce the collective posterior π^H of H, because π^H accounts not just for one E_i (whose reduced support for H exactly compensates the increased prior support) but for the entire conjunction $E_1 \cap ... \cap E_n$ (whose reduced support for H overcompensates the increased prior support).

4 Multiplicative opinion pooling

If we treat the priors $p_1, ..., p_n, p$ as fixed parameters, the pooling formula of Theorem 1 depends just on $\pi_1, ..., \pi_n$, hence defines a classic pooling function $F : \Pi^n \to \Pi$. Specifically, this pooling function is given by $\pi \propto g \cdot \pi_1 \cdots \pi_n$ where g is a fixed function on $\mathcal{H}$ defined as $g := p/(p_1 \cdots p_n)$ (and in particular as p^{1-n} under Common Prior (CP)). So, our axioms lead to what one might call a *multiplicative* opinion pool. Formally, a (classic) opinion pool $F : \Pi^n \to \Pi$ is *multiplicative* if it is given by

$$F(\pi_1, ..., \pi_n) \propto g \cdot \pi_1 \cdots \pi_n \text{ for all } \pi_1, ..., \pi_n \in \Pi,$$

for some fixed function $g : \mathcal{H} \to (0, \infty)$.[8] The simplest multiplicative rule is that in which g takes the value 1 everywhere, so that

$$F(\pi_1, ..., \pi_n) \propto \pi_1 \cdots \pi_n \text{ for all } \pi_1, ..., \pi_n \in \Pi.$$

Note how multiplicative opinion pools differ from the more common linear and geometric opinion pools; these arise from different axiomatic systems that do not make information explicit.

In fact, our axioms not only imply that pooling be multiplicative: they *characterise* multiplicative pooling if $\mathcal{H}$ is finite because every multiplicative rule can be obtained from suitable priors $p_1, ..., p_n, p \in \Pi$.[9]

Our axioms always lead to multiplicative pooling, but it is of course not enough in practice to use *any* multiplicative rule: it matters which one is used, as the resulting collective beliefs are highly sensitive to the parameter g resp. to $p_1, ..., p_n, p$. More precisely, the choice of multiplicative rule determines how the *shared* information is represented in collective beliefs, as shared information is what the prior functions $p_1, ..., p_n, p$ reflect. The next section addresses this issue.

5 Choosing the collective prior p when there is no common prior

If the interpretation of the shared information is controversial and hence (CP) fails, the group needs to determine the collective prior p in Theorem 1's formula. At least three strategies are imaginable. First, one might define p as a uniform or maximum-entropy prior if available. Second, someone, not necessarily a group member, may be appointed to choose p, either by drawing on his own prior beliefs, or by taking inspiration from the submitted priors $p_1, ..., p_n$, or by using statistical estimation techniques if available. These two solutions have obvious limitations, including some ad-hoc-ness and a lack of democracy. A third alternative is to replace p by $F(p_1, ..., p_n)$ and thus define the collective opinion by

$$\pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n), \tag{4}$$

where $F : \Pi^n \to \Pi$ is a standard opinion pool. Note that F is used here not to aggregate people's actual (posterior) opinions $\pi_1, ..., \pi_n$ but to aggregate their prior opinions $p_1, ..., p_n$, namely into a 'compromise prior'. At first sight, one may wonder what is gained by formula (4) compared to the standard approach of defining $\pi = F(\pi_1, ..., \pi_n)$ without having to care about priors $p_1, ..., p_n$. Does formula (4) not just

[8] As $F(\pi_1, ..., \pi_n)$ sums to 1, the factor or proportionality is $\left(\sum_{H \in \mathcal{H}} g(H) \cdot \pi_1(H) \cdots \pi_n(H)\right)^{-1}$.

[9] For any multiplicative rule $F : \Pi^n \to \Pi$, say generated by the function g, if for instance $p_1 = ... = p_n = p \propto g^{-1/(n-1)}$ then $g \propto p/(p_1 \cdots p_n)$, and hence the multiplicative rule generated by g coincides with that arising in Theorem 1.

shift the classic 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 $\pi_1, ..., \pi_n$: unlike $\pi_1, ..., \pi_n$, the prior opinions $p_1, ..., p_n$ involve no informational asymmetry since each p_i is based on the same (shared) information.[10] Hence any disagreement between $p_1, ..., p_n$ is due solely to different *interpretations* of that same body of information. This may facilitate the choice of F. For instance, aggregation may be guided by the unanimity/Pareto principle (which is problematic under informational asymmetry, as we have seen). Further, aggregation may place *equal weights* on each or the priors $p_1, ..., p_n$ (whereas pooling $\pi_1, ..., \pi_n$ may involve the difficult and vague exercise of assigning more weight to better informed people). The literature's two most prominent types of opinion pools $F : \Pi^n \to \Pi$ are

$$\begin{array}{ll} \textit{linear opinion pools:} & F(p_1, ..., p_n) = w_1 p_1 + ... + w_n p_n, \\ \textit{geometric opinion pools:} & F(p_1, ..., p_n) \propto p_1^{w_1} \cdots p_n^{w_n}, \end{array}$$

with weights $w_1, ..., w_i \in [0, 1]$ that add up to 1 (where in the geometric pool the factor of proportionality is chosen such that $\sum_{H \in \mathcal{H}} F(p_1, ..., p_n)(H) = 1$). If F is a linear resp. geometric opinion pool, our pooling formula (4) becomes

$$\pi = \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} (w_1 p_1 + ... + w_n p_n) \tag{5}$$

$$\text{resp. } \pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} p_1^{w_1} \cdots p_n^{w_n} = \frac{\pi_1}{p_1^{1-w_1}} \cdots \frac{\pi_n}{p_n^{1-w_n}}. \tag{6}$$

How should the weights $w_1, ..., w_n$ be chosen in practice? 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, so that our pooling formula becomes

$$\pi = \frac{1}{n} \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} (p_1 + ... + p_n) \tag{7}$$

$$\text{resp. } \pi \propto \frac{\pi_1}{p_1^{1-1/n}} \cdots \frac{\pi_n}{p_n^{1-1/n}}, \tag{8}$$

which is parameter-free, hence uniquely solves the aggregation problem.

6 External and internal Bayesianity

I now give an argument in defence of defining F in (4) as a geometric (or more generally, externally Bayesian) opinion pool, hence in defence of our pooling formulae

[10] One might even argue that, while pooling $p_1, ..., p_n$ into p is possible without using extra information (due to the informational symmetry), pooling $\pi_1, ..., \pi_n$ into π is impossible without extra information (such as $p_1, ..., p_n$).

(6) and (8). Note first that in (4) 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 1 *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. As shown above, our axioms imply that a GOP G should take the form (4), i.e. the form

$$G(p_1, \pi_1, ..., p_n, \pi_n) \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n) \tag{9}$$

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

From a Bayesian perspective, two natural conditions may be imposed on a GOP, to be called *external* and *internal Bayesianity.* The former is an analogue of the equally-named classic condition for standard opinion pools F: it should not matter whether information arrives before or after pooling, i.e. pooling should commute with Bayesian updating. Formally, for every opinion $p \in \Pi$ and (*likelihood*) function $l : \mathcal{H} \to (0, 1]$ the (*updated*) opinion $p^l \in \Pi$ is defined by

$$p^l(H) := \frac{l(H)p(H)}{\sum_{H' \in \mathcal{H}} l(H')p(H')}, \text{ in short } p^l \propto lp. \tag{10}$$

Here, l is interpreted as a likelihood function $P(E|.)$ for some observation 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, geometric opinion pools are externally Bayesian. An analogous concept can be defined for GOPs:

Definition 2 *A GOP* $G : \Pi^{2n} \to \Pi$ *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. *Internal Bayesianity* requires that it be irrelevant whether every or just a single person obtains a given information:

Definition 3 *A GOP* $G : \Pi^{2n} \to \Pi$ *is called* internally 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 information, hence not reflected in any prior. Internal 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 internal 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 (9), *then* external and internal Bayesianity are in fact equivalent, and equivalent to external Bayesianity of F:

Theorem 2 *If a generalised opinion pool* $G : \Pi^{2n} \to \Pi$ *has the form* (9) *where* $F : \Pi^n \to \Pi$ *is any opinion pool, the following conditions are equivalent:*
(i) G *is externally Bayesian;*
(ii) G *is internally Bayesian;*
(iii) F *is externally Bayesian.*

So, if one desires G to be externally or internally Bayesian, one is bound to use an externally Bayesian opinion pool F in our pooling formula (9), for instance a geometric opinion pool F, which leads to pooling formula (6), hence to (8) in the equal-weight case. There also exist more complex (non-geometric) externally Bayesian opinion pools F, characterised in full generality by Genest, McConway, and Schervish (1986, Theorem 2.5), but geometric ones become the only solutions if $|\mathcal{H}| \geq 3$ and F has some additional properties (see Genest, McConway, and Schervish (1986), Corollary 4.5).

Proof. I show that (i) is equivalent with each of (ii) and (iii). By (9),

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

and hence by (10)

$$G(p_1^l, \pi_1^l, ..., p_n^l, \pi_n^l) \propto \frac{l\pi_1}{lp_1} \cdots \frac{l\pi_n}{lp_n} F(p_1^l, ..., p_n^l) = \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1^l, ..., p_n^l). \tag{11}$$

On the other hand, again by (9) and (10),

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

Relations (11) and (12) together immediately imply that G is externally Bayesian if and only if F is externally Bayesian. Further, again by (9) and (10),

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

This together with (11) implies that G is internally Bayesian if and only it F is externally Bayesian. ■

7 When is information independent, when not?

Let us go back to Theorem 1's assumption of Independent Information (Ind). This assumption is often a useful idealisation, even in situations where it fails. But what exactly are these real situations where (Ind) fails? An important source for failure is what I call *subgroup information*, that is, information held by more than one but less than all persons. I will prove that, under certain conditions, (Ind) holds *if and only if* there is no subgroup information.

By a person i's *observation set* I mean, informally, the (possibly quite enormous) collection of i's relevant observations/items of information. (Formally, one may define i's observation set as a set $\mathcal{O}_i$ of non-empty *observations* $O \subseteq \Omega$.[11]) In the case of a jury faced with hypotheses about the defendant's guilt, i's observation 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.

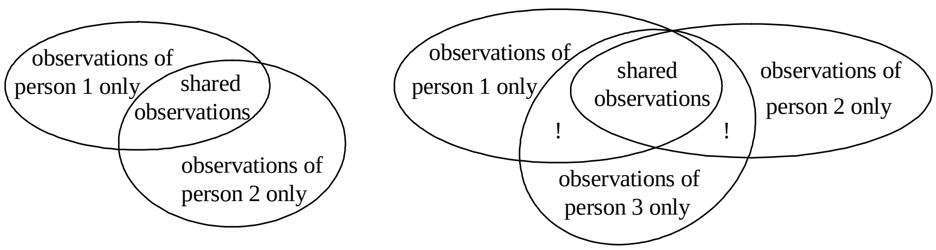

Figure 1: Observation sets in a group of $n = 2$ perons (no subgroup information), and a group of $n = 3$ persons (with subgroup information marked by "!")

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

[11] An observation made by *every* person is represented by the sure event $O = \Omega$, because Ω is interpreted as containing the worlds that are possible under shared information. Formally, $O \in \mathcal{O}_1 \cap ... \cap \mathcal{O}_n$ implies $O = \Omega$.

ler the corresponding set of worlds (in which the observations hold); the *union* of observation sets compares to the *intersection* of the sets of worlds.[12]

Here is the problem. Consider any observation contained in the observation 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 observation is not part of the shared information, but of the personal information E_i of *many* individuals i. Such subgroup information 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 to both witness reports and had a private chat with the defendant (E_3). Note the subgroup information of jurors 1 and 3, and that of jurors 2 and 3, which typically causes E_3 to be positively correlated with E_1 and with E_2. By contrast, individuals 1 and 2 together have no subgroup information. This situation is depicted in Figure 1 on the right.

To formally clarify the relationship between subgroup information and independence violation, some preparation is needed.

Definition 4 *A* subgroup *is a non-empty subset M of the group $N := \{1, ..., n\}$. A subgroup is* proper *if it contains more than one but less than all persons.*

To formalise the notion of subgroup information, suppose that to each subgroup M there is a non-empty event $E^M \subseteq \Omega$, *M's exclusively shared information,* 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 $i \in M$);[13]
- $E^N = \Omega$ (as *any* world $\omega \in \Omega$ is assumed possible under the shared information);
- 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 $E^{\{2,3\}}$.[14]

[12] Formally, to an observation set $\mathcal{O}$ corresponds the set of worlds $\cap_{O \in \mathcal{O}} O \subseteq \Omega$, interpreted as Ω if $\mathcal{O} = \emptyset$. Thus i's information E_i equals $\cap_{O \in \mathcal{O}_i \setminus (\mathcal{O}_1 \cup ... \cup \mathcal{O}_n)} O$, the intersection of all of i's observations except from any shared one; by footnote 11, this actually equals $\cap_{O \in \mathcal{O}_i} O$.

[13] 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.

[14] E^M is interpretable as the intersection $\cap_{O \in (\cap_{i \in M} \mathcal{O}_i) \setminus (\cup_{i \notin M} \mathcal{O}_i)} O$ of all observations O contained in *each* of the observation sets $\mathcal{O}_i$, $i \in M$, but in *none* of the observation sets $\mathcal{O}_i$, $i \notin M$, where this intersection is Ω if $(\cap_{i \in M} \mathcal{O}_i) \setminus (\cup_{i \notin M} \mathcal{O}_i) = \emptyset$.

What we have to exclude is that a proper subgroup M exclusively shares information; in other words, E^M must be the no-information event Ω:

No Subgroup Information (NoSI) All proper subgroups M have $E^M = \Omega$ (i.e. do not exclusively share any information).

This condition is empty if there are just $n = 2$ individuals, it requires $E^{\{1,2\}} = E^{\{1,3\}} = E^{\{2,3\}} = \Omega$ if $n = 3$, and it requires the '!' fields in Figure 1 to be empty. Finally, consider the following independence assumption:

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

(Ind*) is a more generally acceptable condition than (Ind) in that the E^Ms, unlike the E_is, are based on non-overlapping observation sets. Indeed, a subgroup M's exclusively shared information E^M, by the very meaning of 'exclusively', represents different observations than any other subgroup's exclusively shared information.[15] For simplicity, suppose finally that

$$P(A) > 0 \text{ for every non-empty event } A \in \mathcal{A}. \tag{13}$$

Theorem 3 *Assume (Ind*) and (13). Then:*

(a) *Independent Information (Ind) is equivalent to No Subgroup Information (NoSI);*

(b) *specifically, if $E^M \neq \Omega$ for proper subgroup M, then conditional on some $H \in \mathcal{H}$ the personal observations E_i, $i \in M$, are pairwise positively correlated (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 (NoSI). Then we have, for all persons i,

$$E_i = \cap_{\{i\}\subseteq M\subseteq N} E^M = E^{\{i\}} \cap \left[\cap_{\{i\}\subseteq M\subseteq N \& |M|\geq 2} E^M\right] = E^{\{i\}} \cap \Omega = E^{\{i\}}. \tag{14}$$

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 (14).

(ii) Now assume (NoSI) 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 (13) 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*)

[15] (Ind*) holds if the observations in $\mathcal{O}_1 \cup ... \cup \mathcal{O}_n$ are mutually (conditionally) independent.

$P(E_i|H) = \Pi_{\{i\}\subseteq M\subseteq N} P(E^M|H)$. The analogous argument for j yields $P(E_j|H) = \Pi_{\{j\}\subseteq M\subseteq N} P(E^M|H)$. So

$$P(E_i|H)P(E_j|H) = \left[\Pi_{\{i\}\subseteq M\subseteq N} P(E^M|H)\right] \times \left[\Pi_{\{j\}\subseteq M\subseteq N} P(E^M|H)\right]. \quad (15)$$

Further, we have

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

So, by (Ind*),

$$P(E_i \cap E_j) = [\Pi_{\{i\}\subseteq M\subseteq N} P(E^M)] \times [\Pi_{\{j\}\subseteq M\subseteq N\setminus\{i\}} P(E^M)]. \quad (16)$$

The relations (15) and (16) together entail $P(E_i \cap E_j) > P(E_i|H)P(E_j|H)$, because expression (15) equals expression (16) multiplied with the factor $\Pi_{\{i,j\}\subseteq M\subseteq N} P(E^M)$, which is smaller than 1 since it contains the term $P(E^{M^*}|H) < 1$. ■

8 Opinion pooling in the presence of subgroup information

One may always try to 'remove' subgroup information through active information sharing prior to aggregation: all proper subgroups with exclusively shared information communicate this information to the rest of the group. In Figure 1, the observations in each '!' field are 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 subgroup information, (NoSI) and hence (in view of Theorem 3) Independent Information (Ind) hold, so that opinion pooling can proceed along the lines of Sections 2-5.

But suppose now that such information sharing is not feasible, e.g. due to the complexity of subgroup information. Then (NoSI) fails, and hence (Ind) fails, so that we need to modify our pooling formula. It is at first not obvious whether and how one can generalise Theorem 1 to arbitrary information overlaps, i.e. whether and how collective opinions can incorporate all information spread around the group. The generalisation is possible, as will be seen. Roughly speaking, we have to replace Theorem 1's axioms of Individual Bayesian Rationality (IBR) and Independent Information (Ind) by corresponding axioms based on subgroups rather than individuals. Theorem 1's two other axioms, Acceptance of Likelihoods (AL) and Common Prior (CP), will not anymore appear explicitly, but are build implicitly into the model, as explained in a moment. The adapted axioms will again lead to a unique collective opinion π, calculated in a somewhat more complicated way than in Theorem 1.

First, let me state the model ingredients. On the informational side, Theorem 1's model contained individual information $E_1, ..., E_n$; the present model moreover contains each subgroup M's exclusively shared information E^M, as introduced in the last section. Recall that in Theorem 1's model (in its common prior version) people provide individual opinions $\pi_1, ..., \pi_n$ and a common prior opinion p based on the group's shared information; so, technically, the model contained the opinions $\pi_1, ..., \pi_n, p$ reflecting the shared information of the *improper* subgroups $\{1\}, ..., \{n\}, N$, respectively. Our new model adds to this the opinions reflecting the shared information of *proper* subgroups $M \subseteq N$. More precisely, it suffices here to consider subgroups with exclusively share information: let $\mathcal{M}$ be a set of subgroups $M \subseteq N$ containing at least the (proper or improper) subgroups M with exclusively shared information, i.e. with $E^M \neq \Omega$; and let $N \in \mathcal{M}$ without loss of generality.[16] Each subgroup M in $\mathcal{M}$ submits an opinion $p_M \in \Pi$, 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 too; see Definition 5 below). Theorem 1's model (in the common prior version) is the special case that $\mathcal{M} = \{\{1\}, ..., \{n\}, N\}$ ($= \{M : M$ is an improper subgroup$\}$) with $p_{\{1\}} = \pi_1, ..., p_{\{n\}} = \pi_n, p_N = p$. In the last section's jury example with $n = 3$ individuals, we may put $\mathcal{M} = \{\{1\}, \{2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$ because $\{1,2\}$ has no exclusively shared information.

In practice, every non-singleton subgroup $M \in \mathcal{M}$ will have to 'sit together', find out about its shared information, and come up with a resulting opinion p_M. As mentioned, this amounts to a common prior assumption: the present model allows difference in opinion to come only from difference in information. But, rather than making this assumption explicit by a condition analogous to the earlier Common Prior (CP), the assumption is implicit by not indexing p_M by individuals i, and by using P instead of P_i throughout, thereby implicitly assuming that $P_i(A) = P(A)$ for all $A \in \mathcal{A}_i \cap \mathcal{A}$.[17]

The technique to calculate the (collective) opinion $\pi \in \Pi$ from the subgroup opinions p_M, $M \in \mathcal{M}$, will be recursive. Let me first illustrate it using the last section's jury example. Here, $n = 3$ and $\mathcal{M} = \{\{1\}, \{2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$. 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, merge $p_{\{1,3\}}$ and $p_{\{2,3\}}$ into a function $p_{\{1,3\},\{2,3\}}$ that combines $\{1,3\}$'s shared information and $\{1,3\}$'s shared information. One may apply Theorem 1's formula: $p_{\{1,3\},\{2,3\}} \propto p_{\{1,3\}} p_{\{2,3\}} / p_{\{1,2,3\}}$.
- Next, merge $p_{\{1\}}$ and $p_{\{2\}}$ into a function $p_{\{1\},\{2\}}$ that combines $\{1\}$'s and $\{2\}$'s information. One may apply Theorem 1's formula: $p_{\{1\},\{2\}} \propto p_{\{1\}} p_{\{2\}} / p_{\{1,2\}}$,

[16] One may always define $\mathcal{M}$ as containing *all* subgroups, but in practice this maximal choice adds unnecessary steps to the recusive pooling procedure introduced below. The minimal choice is $\mathcal{M} = \{M : \emptyset \neq M \subsetneq N$ and $E^M \neq \Omega\} \cup \{N\}$.

[17] By using P rather than $P_1, ..., P_n$ I implicitly make a common prior assumption that is global, i.e. is not like (CP) restricted to the set $\mathcal{H}$ of relevant hypotheses. I thereby implicitly also assume (AL).

where $p_{\{1,2\}}$ is defined as $p_{\{1,2,3\}}$ because the subgroup $\{1,2\}$ has no exclusively shared information.

- Finally, merge $p_{\{1\},\{2\}}$ and $p_{\{3\}}$ into the function $\pi = p_{\{1\},\{2\},\{3\}}$ that combines $\{1\}$'s, $\{2\}$'s and $\{3\}$'s information. Again, one may apply Theorem 1's formula: $\pi = p_{\{1\},\{2\},\{3\}} \propto p_{\{1\},\{2\}} p_{\{3\}} / p_{\{1,3\},\{2,3\}}$.

Now I come to the formal treatment. Recall that i's information E_i equals $\cap_{\{i\} \subseteq M \subseteq N} E^M$, i.e. i knows precisely the conjunction of what the subgroups containing i exclusively share. This generalises as follows to:

Definition 5 *A subgroup M's* shared information *is defined as $E_M := \cap_{M \subseteq M' \subseteq N} E^{M'}$ (the conjunction of all information exclusively shared by some supergroup of M).*

E_M represents what is known to *at least* all members of M – as opposed to M's *exclusively* shared information E^M, known *exactly* all members of M. Taking the case of a singleton subgroup $M = \{i\}$, the event $E_{\{i\}}$ coincides with E_i. Also, note that

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

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

Subgroup Bayesian Rationality (SBR) $p_M(H) = P(H|E_M)$ for every subgroup $M \in \mathcal{M}$ and hypothesis $H \in \mathcal{H}$.

As in Theorem 1, we would like the collective opinion to satisfy Collective Bayesian Rationality (CBR); that is, we require 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 5)

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

As a technical tool to construct collective opinion π satisfying (CBR), I need to introduce opinions of *abstract* individuals.

Definition 6 *An* abstract individual *is a non-empty set A of subgroups M; its* order *is order$(A) := \min\{|M| : M \in A\}$, the size of a smallest subgroup in A.*

The opinions $p_{\{1,3\},\{2,3\}}, p_{\{1\},\{2\}}, ...$ defined in the example above are in fact the opinions of the abstract individuals $\{\{1,3\},\{2,3\}\}, \{\{1\},\{2\}\}, ...$ More generally, I interpret an abstract individual A as a hypothetical agent who knows the shared information of any subgroup $M \in A$ (and no more). For instance, $A = \{\{1,3\},\{2,3\}\}$

knows $\{1,3\}$'s shared information *and* $\{2,3\}$'s shared information. A's information 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

$$p_A(H) = P(H|\cap_{M\in A} E_M) \text{ for each } H \in \mathcal{H}. \tag{17}$$

Specifically, 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 (17) $\pi = P(.|E_{\{1\}} \cap ... \cap E_{\{n\}}) = p_A$ where A is the abstract individual $\{\{1\},\{2\},...,\{n\}\}$ of order 1. In the recursive construction, the main steps are to calculate from opinions p_A and p_{A^*} of abstract individuals A and A^* the opinion $p_{A\cup A^*}$ of the abstract individual $A \cup A^*$ whose information combines the 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. To do so, the notion of shared information is crucial. What information do A and A^* share? They share precisely the information held by the abstract individual

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

The reason is: the information A and A^* share is precisely the information that A knows *and* A^* knows, i.e. that some subgroup in A shares *and* some subgroup in A^* shares, i.e. that some union $M \cup M^*$ with $M \in A$ and $M^* \in A^*$ shares. So, when combining opinions p_A and p_{A^*}, $A \vee A^*$'s opinion $p_{A\vee A^*}$ plays the role of the common prior p in Theorem 1. More precisely, the crucial result on how to combine opinions of abstract individuals states as follows (and is proved later):

Lemma 1 *Assume (Ind*). Consider abstract individuals B and C, form the abstract individuals $B \vee C$ and $B \cup C$, and let $p_B, p_C, p_{B\vee C}, p_{B\cup C}$ be four opinions in Π. If $p_B, p_C, p_{B\vee C}$ are all given by (17), then the function $p_B p_C / p_{B\vee C}$ is proportional to an opinion in Π (equivalently, has a finite sum $\sum_{H\in\mathcal{H}} p_B(H)p_C(H)/p_{B\vee C}(H)$), and if moreover $p_{B\cup C}$ is this opinion (i.e. $p_{B\cup C} \propto p_B p_C / p_{B\vee C}$) then $p_{B\cup C}$ is given by (17).*

The formula in Lemma 1 guides us in assigning beliefs to abstract individuals. The assignment is recursive, with another nested recursion in 'Case 2':

Definition 7 *Define the opinions $p_A \in \Pi$ of abstract individual A by the following 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 $M \notin \mathcal{M}$) and define $p_A := p_{A'}$ (where $p_{A'}$ is already defined by order$(A') = k+1$).*

Case 2*: $|A| > 1$. Define p_A by another 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 by $p_A \propto p_{\{M\}} p_{A^*}/p_{\{M\} \vee A^*}$ (where $p_{\{M\}}$ is already defined in case 1, and p_{A^*} and $p_{\{M\} \vee A^*}$ are already defined by order$(A^*) > k$ and order$(\{M\} \vee A^*) > k$).*
- *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 by $p_A \propto p_{\{M\}} p_{A^*}/p_{\{M\} \vee A^*}$ (where $p_{\{M\}}$ is already defined in case 1, p_{A^*} is already defined by $|\{M^* \in A^* : |M^*| = k\}| = l - 1$, and $p_{\{M\} \vee A^*}$ is already defined by order$(\{M\} \vee A^*) > k$).*

The existence and uniqueness of the above-defined opinions p_A follows from the recursion theorem.[18] On the last recursion step we reach the opinions p_A of abstract individuals of order 1, hence in particular the opinion of $A = \{\{1\}, ..., \{n\}\}$, and this is the desired opinion that incorporates the group's full information:

Theorem 4 *If (SBR), (CBR) and (Ind*) hold, the collective opinion π is given by $\pi = p_{\{\{1\},...,\{n\}\}}$, the opinion of the abstract individual $\{\{1\}, ..., \{n\}\}$.*

I first prove Lemma 1 and then Theorem 4.

Proof of Lemma 1. Assume (Ind*). Let B, C be abstract individuals, and $p_B, p_C, p_{B \vee C}, p_{B \cup C} \in \Pi$. Suppose $p_B, p_C, p_{B \vee C}$ satisfy (17). For all abstract individuals A, put

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

the set of supergroups of subgroups in A. By (17), $p_{B \vee C} = P(.| \cap_{M \in B \vee C} E_M)$, where by Definition 5

$$\cap_{M \in B \vee C} E_M = \cap_{M \in B \vee C} \cap_{M \subseteq M' \subseteq N} E^{M'} = \cap_{M \in \overline{B \vee C}} E^M.$$

[18] A technical detail is left implicit in Definition 7: in each bullet point of Case 2, I have defined p_A as the member of Π that is proportional to the function a certain function f $(= p_{\{M\}} p_{A^*}/p_{\{M\} \vee A^*})$, but this is only meaningful if there *exists* a $g \in \Pi$ with $g \propto f$ (i.e. if f has a finite sum $\sum_{H \in \mathcal{H}} f(H) < \infty$ so that f can be normalised to a function in Π). Existence does indeed holds under Theorem 4's axioms (see the proof of Theorem 4, which draws on Lemma 1), but strictly speaking this fact should not be anticipated in the recursive definition. This is why Definition 7 strictly speaking needs the following extension. Fix an arbitrary belief $\sigma \in \Pi$, and add to Cases 1 and 2 the clause that p_A is defined as σ if the previous prescription does not apply (i.e. if there is non-existence, as just discussed). The added clause can then be shown to never apply (under Theorem 4's axioms).

So,

$$p_{B\vee C} = P(.|E) \text{ with } E := \cap_{M\in\overline{B\vee C}} E^M. \tag{18}$$

Analogously, by (17), $p_B = P(.|\cap_{M\in B} E_M)$, where by Definition 5

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

with $E_B := \cap_{M\in\overline{B}\backslash\overline{B\vee C}} E^M$. So $p_B = P(.|E_B \cap E)$, and hence by Bayes' rule

$$p_B \propto P(.|E)P(E_B|.\cap E). \tag{19}$$

By an analogous argument for C, we have

$$p_C \propto P(.|E)P(E_C|.\cap E), \tag{20}$$

where $E_C := \cap_{M\in\overline{C}\backslash\overline{B\vee C}} E^M$. By (18), (19) and (20) we have

$$\begin{aligned} p_B p_C / p_{B\vee C} &\propto [P(.|E)P(E_B|.\cap E)]\,[P(.|E)P(E_C|.\cap E)]\,/P(.|E) \\ &= P(.|E)P(E_B|.\cap E)P(E_C|.\cap E). \end{aligned} \tag{21}$$

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

$$P(E_B|.\cap E)P(E_C|.\cap E) = P(E_B \cap E_C|.\cap E).$$

Substituting this into (21) and then applying Bayes' rule, we obtain

$$p_B p_C / p_{B\vee C} \propto P(.|E)P(E_B \cap E_C|.\cap E) \propto P(.|E_B \cap E_C \cap E) \in \Pi.$$

Now suppose $p_{B\cup C} = P(.|E_B \cap E_C \cap E)$. We may rewrite $E_B \cap E_C \cap E$ as

$$\cap_{M\in\overline{B\cup C}} E^M = \cap_{M\in B\cup C} \cap_{M\subseteq M'\subseteq N} E^M = \cap_{M\in B\cup C} E_M,$$

and hence $p_{B\cup C}$ equals $P(.|\cap_{M\in B\cup C} E_M)$, i.e. satisfies (17). ■

Proof of Theorem 4. Assume (SBR) and (Ind*). By backward induction on the order of A I show that each abstract individual A has opinion p_A satisfying (17). This in particular implies that $\{\{1\},...,\{n\}\}$ has opinion

$$p_{\{\{1\},...,\{n\}\}}(H) = P(H|E_1 \cap ... \cap E_n) \text{ for each } H \in \mathcal{H},$$

so that under (CBR) we have $\pi = p_{\{\{1\},...,\{n\}\}}$, as desired.

Denote by $\mathbf{A}$ the set of abstract individuals A. The recursion proceeds as follows.

- If $order(A) = n$, then $A = \{N\}$, and by definition $p_A = p_N$. So by (SBR) $p_A = P(.|E_N) = P(.|\cap_{M\in A} E_M)$, as desired.

- Now let $order(A) = k < n$, and assume (17) holds for all $A' \in \mathbf{A}$ 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 (SBR) $p_A = P(.|E_M) = P(.| \cap_{M' \in A} E_{M'})$, as desired. 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 5,

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

  In this, $E^M = \Omega$ (by $M \notin \mathcal{M}$) and $\cap_{M' \subseteq M'' \subseteq N} E^{M''} = E_{M'}$ (by Definition 5). 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 $A = \{M\} \cup A^*$ with $|M| = k$ and $order(A^*) > k$. Then p_A was defined as the function in Π proportional to $p_{\{M\}} p_{A^*} / p_{\{M\} \vee A^*}$; let me show that (i) such a function does indeed exists (see footnote 18 on potential inexistence) and (ii) satisfies (17), as desired. Now, $p_{\{M\}}$ satisfies (17) by Case 1, and p_{A^*} and $p_{\{M\} \vee A^*}$ satisfy (17) by $order(A^*) > k$ and $order(\{M\} \vee A^*) > k$ (and the k-induction hypothesis). So, by Lemma 1, the function $p_{\{M\}} p_{A^*} / p_{\{M\} \vee A^*}$ is proportional to a function in Π, so that p_A is well-defined. Also by Lemma 1, this function p_A satisfies (17), as desired.
  - Let $|\{M \in A : |M| = k\}| = l > 1$, and assume A^* satisfies (17) whenever $|\{M \in A^* : |M| = k\}| < l$ (and $order(A^*) = k$). By definition, $p_A \propto p_{\{M\}} p_{A^*} / p_{\{M\} \vee A^*}$, where $A = \{M\} \cup A^*$ with $|M| = k$ and $|\{M^* \in A^* : |M^*| = k\}| = l - 1$. Again, we have to show that p_A is well-defined (i.e. Π indeed contains a function proportional to $p_{\{M\}} p_{A^*} / p_{\{M\} \vee A^*}$) and satisfies (17). $p_{\{M\}}$ satisfies (17) by Case 1, p_{A^*} satisfies (17) by $|\{M^* \in A^* : |M^*| = k\}| = l - 1$ (and the l-induction hypothesis), and $p_{\{M\} \vee A^*}$ satisfies (17) by $order(\{M\} \vee A^*) > k$ (and the k-induction hypothesis). So, by Lemma 1, p_A is well-defined and satisfies (17). ■

9 Conclusion

The above model interprets opinion pooling as information pooling: collective opinions should build in the group's entire information, be it shared or personal. According to the pooling formulae I obtained, collective opinions should account for informational asymmetries *not* by taking a standard weighted (linear or geometric) average of the individual opinions with higher weight assigned to better informed individuals *but* by incorporating people's prior opinions in addition to their actual (i.e. posterior) opinions. In practice, people have either to agree on a common prior opinion p, i.e. on how to interpret the shared information, or they have to submit their possibly diverging prior opinions $p_1, ..., p_n$. Based on simple axioms, Theorem 1 shows how to aggregate the (prior and posterior) opinions into a collective opinion. The formula defines a *multiplicative* opinion pool: the collective opinion π is proportional to the product of the individual opinions $\pi_1, ..., \pi_n$ and a function g (that depends on prior opinions).

More precisely, Theorem 1 suggests that, based on individual opinions $\pi_1, ..., \pi_n$, the collective opinion π should be defined by $\pi \propto \pi_1 \cdots \pi_n / p^{n-1}$ if people agree on a common prior p, and by $\pi \propto \frac{\pi_1}{p_1} \cdots \frac{\pi_n}{p_n} F(p_1, ..., p_n)$ if people have arbitrary priors $p_1, ..., p_n$, where F is a standard opinion pool. I have suggested that F should be anonymous (i.e. symmetric in its arguments) because the prior opinions it pools are based on the same (shared) information, giving no individual an informational superiority. More specifically, I have suggested to define F as unweighted geometric pooling, because this generates appealing properties shown in Theorem 2. This choice of F gives collective opinion the form

$$\pi \propto \frac{\pi_1}{p_1^{1-1/n}} \cdots \frac{\pi_n}{p_n^{1-1/n}}.$$

A crucial axiom underlying this formula is that personal information is independent. By Theorem 3, independence is threatened by the possibility of subgroup information, i.e. of information held by more than one but less than all individuals. Theorem 4 therefore generalises the aggregation rule to arbitrary information distributions (allowing for subgroup information). The generalisation is unique, but assumes that each subgroup with subgroup information agrees on how to interpret this information, a kind of common prior assumption. Dropping this assumption would have gone beyond the scope of this paper, but it might be an interesting route for future research.

10 References

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

Dietrich, F. (2004) Opinion Pooling under Asymmetric Information, working paper, *Public Economics 0407002, EconWPA*

Dietrich, F., and C. List (2007) Opinion pooling on general agendas, working paper, *METEOR Research Memorandum 038*, Maastricht University

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

Lehrer, K. and C. Wagner (1981) *Rational Consensus in Science and Society*, Dordrecht: Reidel

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

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

McConway, K. (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

Pearl, J. (2000) *Causality: Models, Reasoning and Inference*, Cambridge: Cambridge University Press

Pivato, M. (2008) The Discursive Dilemma and Probabilistic Judgement Aggregation, *MPRA Paper 8412*, University Library of Munich, Germany

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

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