Forthcoming in the *Journal of Philosophical Logic*

PROBABILISTIC OPINION POOLING WITH IMPRECISE PROBABILITIES

RUSH STEWART AND IGNACIO OJEA QUINTANA

Abstract. The question of how the probabilistic opinions of different individuals should be aggregated to form a group opinion is controversial. But one assumption seems to be pretty much common ground: for a group of Bayesians, the representation of group opinion should itself be a unique probability distribution (Madansky, 1964; Lehrer and Wagner, 1981; McConway, 1981; Bordley, 1982; Genest et al., 1986; Genest and Zidek, 1986; Mongin, 1995; Clemen and Winkler, 1999; Dietrich and List, 2014; Herzberg, 2014). We argue that this assumption is not always in order. We show how to extend the canonical mathematical framework for pooling to cover pooling with *imprecise probabilities* (IP) by employing *set-valued* pooling functions and generalizing common pooling axioms accordingly. As a proof of concept, we then show that one IP construction satisfies a number of central pooling axioms that are not jointly satisfied by any of the standard pooling recipes on pain of triviality. Following Levi (1985), we also argue that IP models admit of a much better philosophical motivation as a model of rational consensus.

1. Introduction

The problem of opinion aggregation is "the problem of determining a sensible formula for representing the opinions of a group" (Genest et al., 1986). Representations of group opinion are important in a number of contexts, from scientific advisory panels (on climate change, for example), to joint efforts in scientific inquiry, to decision making in various kinds of groups. In a Bayesian setting, group consensus is particularly important from a theoretical standpoint. The received view is that probabilistic opinions are *subjective* (Ramsey, 1931; Savage, 1954; de Finetti, 1964). Forms of intersubjective agreement have been sought to replace the surrendered notion of objectivity (Genest and Zidek, 1986; Nau, 2002). Probabilistic opinion pooling is one proposal for finding such consensus. It is widely assumed that, for a group of Bayesians, a representation of group opinion should take the form of a (single) probability distribution.The central position of this essay is that, in certain philosophically interesting and important cases, such an assumption is not always appropriate.

At the end of their review article on pooling, Dietrich and List mention other approaches that "redefine the aggregation problem itself" (2014, p. 20). According to them, one such approach is the aggregation of imprecise probabilities.[1] Of the few accounts of aggregating probabilities that

Date: November 4, 2016.

Several people gave us very helpful feedback on the ideas in this paper. Thanks to Robby Finley and Yang Liu for their comments on a presentation given to the Formal Philosophy reading group at Columbia University, and to members of the audiences at the Columbia Graduate Student Workshop, the Probability and Belief Workshop organized by Hans Rott at the University of Regensburg, and a presentation at CUNY organized by Rohit Parikh. We are grateful to Arthur Heller, Michael Nielsen, Teddy Seidenfeld, Reuben Stern, and Mark Swails for their excellent comments on drafts of the paper. We would like to especially thank Isaac Levi for extensive discussion of the content of this paper and comments on drafts and a presentation. Finally, thanks to both the editor and the anonymous referee. The referee provided engaged and thorough feedback that has undoubtedly improved the essay in a number of ways.

[1]Here we use *IP* as a general term, abstracting from the important distinction Isaac Levi makes between what he calls *imprecise* and *indeterminate* probability, or what Walley calls the *Bayesian sensitivity analysis* and *direct* interpretations, respectively. Roughly speaking, according to the first interpretation, while an agent is normatively committed to or descriptively in a state of numerically precise judgments of credal probability, these precise judgments

deal with imprecision, many tend to focus on cases in which the individual opinions are already imprecise. And such accounts do not proceed by generalizing the pooling framework, axioms, etc. (Moral and Del Sagrado, 1998; Nau, 2002). A general account of probabilistic consensus should cover cases in which probabilities are imprecise at the level of the individual (a topic to which we return towards the end of the paper). However, our aim is to call into question the assumption that group opinion should be represented by a single probability distribution when precision holds at the level of the individuals. In this effort, we extend a line of argument that uses limitative results concerning aggregation—results demonstrating the impossibility of jointly satisfying a set of formal pooling criteria for precise aggregation methods—as a springboard into IP (Walley, 1982; Seidenfeld et al., 1989). That is, the limitations of precise pooling motivate IP in the sense that certain IP models *do* satisfy desiderata for "group" opinion that precise models do not.

After presenting the basic mathematical framework for probabilistic opinion pooling, we review some of the central limitative results (Section 2). One contribution of the present essay is generalizing the pooling framework, framing pooling with imprecise probabilities in the mathematical language common in research on probability aggregation with precise probabilities (Sections 4 and 5). The particular IP model that we primarily focus on in this paper, as a proof of concept, is presented in Section 4 (in a sense that will be made clear and precise, our case for considering IP models of pooling does *not* rise and fall with this particular format). Even in cases in which individual probabilities are precise, demanding that the output of an aggregation method be a single probability function is overly restrictive. As we show, representations of group opinion in terms of sets of probability functions have some very nice features. On the one hand, IP allows for a plausible philosophical account of rational consensus (Section 3). On the other hand, the construction we study satisfies a number of the central pooling axioms that are not jointly satisfied by any of the standard, precise pooling recipes on pain of triviality (Sections 5 and 6). We close by considering some potential objections (Section 7).

2. Pooling

A general framework for aggregating the probabilistic opinions of a group to form a collective opinion is that of *pooling*. Formally, a pooling method for a group of n individuals is a function

$$F\colon \mathbb{P}^n \to \mathbb{P}$$

mapping profiles of probability functions for the n agents (or simply the n distributions under consideration), $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, to *single* probability functions intended to represent group opinion, $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. The probabilities are assigned to events, which we represent as subsets of a sample space, Ω. We assume that Ω is countable. The *agenda*, or the set of events under consideration, is assumed to be an algebra $\mathscr{A}$ of events over Ω, that is, a set of subsets of Ω closed under complementation and finite unions (in the general case, closure under countable unions yields a σ-algebra).[2] A function $\boldsymbol{p} : \Omega \to [0, 1]$ is a *probability mass function* (pmf) *iff* $\sum_{\omega \in \Omega} \boldsymbol{p}(\omega) = 1$. Abusing notation, we can define a probability *measure*, $\boldsymbol{p}$, on general events for a given pmf by

may not be precisely elicited or introspected. On the second interpretation, imprecision is a feature of the credal state itself and is not attributable to imperfect elicitation or introspection. It is possible, of course, for a credal state to be imprecise in both senses, that is, an indeterminate credal state could be incompletely elicited.

[2]For completeness, we include the probability axioms. A *probability function* is a mapping $\boldsymbol{p} : \mathscr{A} \to \mathbb{R}$ that satisfies the following conditions:

(i) $\boldsymbol{p}(A) \geqslant 0$ for any $A \in \mathscr{A}$;
(ii) $\boldsymbol{p}(\Omega) = 1$;
(iii) $\boldsymbol{p}(A \cup B) = \boldsymbol{p}(A) + \boldsymbol{p}(B)$ for any $A, B \in \mathscr{A}$ such that $A \cap B = \varnothing$.

If, in addition, $\mathscr{A}$ is a σ-algebra and $\boldsymbol{p}$ satisfies the following condition, $\boldsymbol{p}$ is called *countably additive*:

(iv) If $\{A_n\}_{n=1}^{\infty} \subseteq \mathscr{A}$ is a collection of pairwise disjoint events, then $\boldsymbol{p}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \boldsymbol{p}(A_n)$.

In this paper, we assume countable additivity for convenience, not because we take it to be rationally mandatory.

$\boldsymbol{p}(E) = \sum_{\omega \in E} \boldsymbol{p}(\omega)$. Pooling can be formulated in terms of pmfs, and we will appeal to pmfs in discussing geometric pooling functions and the external Bayesianity constraint below.

Various interpretations of pooling are proposed in the literature. Wagner, for example, offers the following (2009, pp. 336-337):

(1) A rough summary of the current probabilities of the n individuals;
(2) a "compromise" adopted by the individuals for the purpose of group decision making;
(3) a rational consensus to which the individuals revise their probabilities after discussion;
(4) the opinion a decision maker external to the group adopts upon being informed of the n expert opinions in the group;
(5) the opinion an individual in the group adopts upon being informed of the $n-1$ opinions of his "epistemic peers" in the group.

These five interpretations do not exhaust the possibilities. Our target interpretation is rational consensus, adopted either *for the sake of the argument* (a compromise) in order to perform some task in group inference or decision making (2) or genuinely by individual group members (3, 5). However, the account we consider could also be used by a decision maker external to the group.

2.1. **Criteria for Pooling Functions.** What properties should a pooling function have? We review some of the most popular properties discussed in this connection. It is important to consider, for each property, the extent to which it is normatively compelling for a particular interpretation and use of pooling functions. Surveys of the material presented here include Simon French's (1985), Genest and Zidek's (1986), and Dietrich and List's (2014).

McConway (1981) and Lerher and Wagner (1981) introduce a convenient property of pooling functions called the *strong setwise function property* and *strong label neutrality* by the respective authors. The property has it that the individual probabilities for an event—and not the entire distributions of each individual—are all that is required to determine the collective probability of that event.

Strong Setwise Function Property. There exists a function $G : [0, 1]^n \to [0, 1]$ such that, for every event A, $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = G(\boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A))$.

What case can be made for the strong setwise function property (SSFP) as a pooling *norm*? SSFP can be seen as a probabilistic analogue of the independence of irrelevant alternatives constraint in the social choice literature. Consider two profiles $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ and $(\boldsymbol{p}'_1, ..., \boldsymbol{p}'_n)$. Suppose that, for some event A, $\boldsymbol{p}_i(A) = \boldsymbol{p}'_i(A)$ for $i = 1, ..., n$, but the two profiles differ on some other event (so for pooling probabilities for A, "irrelevant" parts of the probability functions differ). It can happen that $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) \neq F(\boldsymbol{p}'_1, ..., \boldsymbol{p}'_n)(A)$ despite the fact that $(\boldsymbol{p}_1(A), ...\boldsymbol{p}_n(A)) = (\boldsymbol{p}'_1(A), ..., \boldsymbol{p}'_n(A))$. That is, the "consensus" probabilities for A differ for the two profiles despite no change in individual opinions concerning A. So, the pooled probability for A is not a function merely of the individual probabilities for A. For such an F, no function G exists because such a function would have to map one profile of values in $[0, 1]^n$ to two distinct outputs in $[0, 1]$. Admittedly, such a case for the normative status of SSFP is incomplete.

Many of the axioms proposed in the literature on pooling require that some property of the individual probability functions be preserved under pooling. When the algebra contains at least three events, one such preservation property follows immediately from SSFP, as McConway observes (1981, Theorem 3.2).

Zero Preservation Property. For any event A, if $\boldsymbol{p}_i(A) = 0$ for $i = 1, ..., n$, then $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = 0$.

As Genest and Zidek remark, the zero preservation property (ZPP) is one in a class of constraints requiring that the pool preserves initial shared agreements. The normative status of this sort of preservation axiom has been called into question in the literature (Genest and Wagner, 1987). Of course, ZPP is forced upon those endorsing SSFP. For conceptions of *consensus* on which common ground is sought, that is, a non-question begging position of agreement, ZPP is more compelling. We return to *consensus as shared agreement* or *common ground* below in Section 3.

McConway's Theorem 3.2 shows more. Taken together, the *marginalization property* (MP) and the zero preservation property (ZPP) are equivalent to SSFP. McConway's formal setup differs somewhat from the one presented here. He is concerned with classes of pooling functions that take into account all $\sigma-$algebras on Ω. We, however, are considering pooling functions for a fixed algebra (which seems to be the more common approach). The formal properties of concern to McConway must be modified accordingly. A pooling function satisfies MP if marginalization and pooling commute. We adopt the modification of MP proposed by Genest and Zidek (1986, p. 118). Let $\mathscr{A}'$ be a subalgebra of $\mathscr{A}$.[3] Suppose that $\boldsymbol{p}$ is a distribution over $(\Omega, \mathscr{A})$. The *marginal* distribution $\boldsymbol{p}\restriction_{\mathscr{A}'}$ given by $\boldsymbol{p}$ over $(\Omega, \mathscr{A}')$ is the restriction of $\boldsymbol{p}$ to $\mathscr{A}'$ such that $\boldsymbol{p}(A) = \boldsymbol{p}\restriction_{\mathscr{A}'}(A)$ for all $A \in \mathscr{A}'$. $[\boldsymbol{p}\restriction_{\mathscr{A}'}]$ is a Carathéodory extension of $\boldsymbol{p}\restriction_{\mathscr{A}'}$ to $\mathscr{A}$.

Marginalization Property. Let $\mathscr{A}'$ be a sub-σ-algebra of $\mathscr{A}$. For any $A \in \mathscr{A}'$, $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = F([\boldsymbol{p}_1\restriction_{\mathscr{A}'}], ..., [\boldsymbol{p}_n\restriction_{\mathscr{A}'}])(A)$.

Below, we will state an analogue of another of McConway's results. That result says that MP is equivalent to the *weak setwise function property* (WSFP) (1981, Theorem 3.1). Instead of $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$ depending just on the $\boldsymbol{p}_i(A), i = 1, ...n$, those pooling functions merely satisfying WSFP depend on both $\boldsymbol{p}_i(A)$ *and* the event, A. The difference is that a profile in $[0, 1]^n$ may be mapped to more than one output, so long as the associated event differs.

Weak Setwise Function Property. There exists a function $G : \mathscr{A} \times [0, 1]^n \to [0, 1]$ such that, for any event $A \in \mathscr{A}$, $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = G(A, \boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A))$ for each profile in the domain of F.

Probabilistic independence is another natural candidate property for preservation under pooling. In the precise setting, there are a number of equivalent formulations of probabilistic independence. For example, two events, A and B, are said to be *stochastically independent* according to $\boldsymbol{p}$ if $\boldsymbol{p}(A \cap B) = \boldsymbol{p}(A)\boldsymbol{p}(B)$. Dividing both sides by $\boldsymbol{p}(B)$, provided $\boldsymbol{p}(B) > 0$, yields $\frac{\boldsymbol{p}(A\cap B)}{\boldsymbol{p}(B)} = \boldsymbol{p}(A)$ when A and B are independent. But the lefthand side of the equation is a standard definition of the probability of A *conditional* on B: $\boldsymbol{p}(A|B) = \frac{\boldsymbol{p}(A\cap B)}{\boldsymbol{p}(B)}$, when $\boldsymbol{p}(B) > 0$. This observation allows us to state another standard formulation of probabilistic independence. A and B are independent according to $\boldsymbol{p}$ if $\boldsymbol{p}(A|B) = \boldsymbol{p}(A)$. The *conditionalization* of $\boldsymbol{p}$ with respect to an event B, $\boldsymbol{p}^B$, is given by setting $\boldsymbol{p}^B(A) = \boldsymbol{p}(A|B)$ for all A. We will return to stochastic independence below, but it will be convenient for us to adopt the definition in terms of conditional probabilities.

Probabilistic Independence Preservation. If $\boldsymbol{p}_i(A|B) = \boldsymbol{p}_i(A)$ for $i = 1, ..., n$, then $F^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$.

This axiom says that two events that are probabilistically independent according to every individual probability function are independent according to the pool.

Another preservation axiom is *unanimity preservation*, which requires that, if all of the functions being pooled are identical, then the output of the pooling function is that probability function. So if all the individual opinions are the same, the group opinion is identical to that common distribution.

Unanimity Preservation. For every opinion profile $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$, if all $\boldsymbol{p}_i$ are identical, then $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \boldsymbol{p}_i$.

Other sorts of pooling axioms, like MP above, demand that some operation or other commutes with pooling. A very interesting example of such an operation is a type of Bayesian updating.

[3] $\mathscr{A}'$ is a boolean subalgebra of $\mathscr{A}$ if $\mathscr{A}' \subseteq \mathscr{A}$ and $\mathscr{A}'$, with the distinguished elements and operations of $\mathscr{A}$, is a boolean algebra. That is, the operations must be the restrictions of the operations of the whole algebra; being a subset that is a boolean algebra is not sufficient for being a subalgebra of $\mathscr{A}$ (Halmos, 1963).

Standard Bayesian conditionalization goes *via* Bayes' theorem:

$$\boldsymbol{p}^B(A) = \boldsymbol{p}(A|B) = \frac{\boldsymbol{p}(A)\boldsymbol{p}(B|A)}{\boldsymbol{p}(B)}, \text{ when } \boldsymbol{p}(B) > 0.$$

By the law of total probability, the denominator, $\boldsymbol{p}(B)$, can be rewritten. Where $\{C_j : j = 1, 2, ...\}$ is a partition of Ω, $\boldsymbol{p}(B) = \sum_j \boldsymbol{p}(B|C_j)\boldsymbol{p}(C_j)$.

External Bayesianity is a mild generalization of commutativity with Bayesian conditionalization. The requirement is that updating the individual probabilities on a common *likelihood function* (as opposed to updating on an event) and then pooling is the same as pooling and then updating the pool on that likelihood function. The likelihood function, $\lambda : \Omega \to [0, \infty)$, is defined on elements of the sample space. In conditionalizing, $\lambda(\cdot)$ serves the same role as the conditional probability $\boldsymbol{p}(B|\cdot)$ in Bayes' theorem above, expressing the degree to which some fixed evidence B is expected on various events. Put roughly, updating on λ results from substituting the likelihood function in for the conditional probabilities on the right hand side of Bayes' theorem. For every $\omega \in \Omega$,

$$\boldsymbol{p}^\lambda(\omega) = \frac{\boldsymbol{p}(\omega)\lambda(\omega)}{\sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega')}, \text{when } \sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega') > 0$$

If $\boldsymbol{p}$ is a probability measure, it must be defined on an algebra including the elements of Ω. Otherwise, take $\boldsymbol{p}$ to be a pmf and obtain a probabiltiy measure on a given algebra by summing over the elements of Ω in each event to obtain the probability of events in the algebra. Comparing the above formula with the version of Bayes' theorem in which the denominator is expanded by the law of total probability makes the relation between $\lambda(\omega)$ and $\boldsymbol{p}(B|\cdot)$ apparent. While not itself a probability distribution, $\lambda(\omega)$ is proportional to $\boldsymbol{p}(B|\omega)$, for fixed data B. And though not a function of general events in $\mathscr{A}$, the likelihood of an event A can be obtained by summing the likelihoods of all $\omega \in A$. Updating on a likelihood function reduces to standard conditionalization on some event, B, when

$$\lambda(\omega) = \begin{cases} 1, & \text{if } \omega \in B \\ 0, & \text{otherwise.} \end{cases}$$

(This reduction holds generally for imprecise probabilities also (Stewart and Ojea Quintana, MS2, Proposition 2).)

Crucially, the likelihood function is assumed to be *common* in the external Bayesianity axiom. So while disagreement concerning the prior is permitted by pooling functions satisfying external Bayesianity, the commutativity of pooling and updating is guaranteed only when there is agreement on the likelihood funciton.

External Bayesianity. For every profile $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ in the domain of F and every likelihood function λ such that $(\boldsymbol{p}^\lambda, ..., \boldsymbol{p}_n^\lambda)$ remains in the domain of F, $F(\boldsymbol{p}_1^\lambda, ..., \boldsymbol{p}_n^\lambda) = F^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$.

A similar axiom requires that a single individual conditionalizing on λ before pooling is the same as conditionalizing the pool on λ (Dietrich and List, 2014).

Individualwise Bayesianity. For every profile $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ in the domain of F and every individual k such that $(\boldsymbol{p}_1, ..., \boldsymbol{p}_k^\lambda, ..., \boldsymbol{p}_n)$ remains in the domain, $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_k^\lambda, ..., \boldsymbol{p}_n) = F^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$

We also have more to say about individualwise Bayesianity below.

2.2. **Types of Pooling Functions.** Various concrete pooling functions have been studied in the literature. These functions fare differently on the criteria reviewed just above. Of the commonly discussed pooling operators, linear pooling functions may be the most common and obvious proposal (Stone, 1961; McConway, 1981; Lehrer and Wagner, 1981).

Linear Opinion Pools. $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \sum_{i=1}^n w_i \boldsymbol{p}_i$, where $w_i \geqslant 0$ and $\sum_{i=1}^n w_i = 1$.

$w_1, ..., w_n$ are fixed non-negative weights summing to 1 that are associated with the n individuals. Linear pooling, then, takes a weighted average of the individual probabilities. Equal weights for the n probability functions specifies one linear pooling function; a *dictatorship* specifies another linear pooling function. In the latter case, all of the weight is accorded to a single individual ($w_i = 1$ for some i) with the result that the pooled probability for any event A is that individual's probability for A: $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \boldsymbol{p}_i(A)$. Interestingly, weights $w_i = \frac{1}{n}$ were used in a U.S. Nuclear Regulatory Commission study of the frequency of nuclear reactor accidents (Ouchi, 2004, p. 5).

Another proposal is to take a weighted *geometric* instead of a weighted arithmetic average of the n probability functions (Madansky, 1964; Bacharach, 1972; Genest et al., 1986).[4]

Geometric Opinion Pools. $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = c\prod_{i=1}^{n} \boldsymbol{p}_i^{w_i}$, where $w_i \geqslant 0$ and $\sum_{i=1}^{n} w_i = 1$, and $c = \frac{1}{\sum_{\omega' \in \Omega} [\boldsymbol{p}_1(\omega')]^{w_1} \cdots [\boldsymbol{p}_n(\omega')]^{w_n}}$ is a normalization factor.

Unlike linear pools, geometric pools specify the collective probabilities of elements of Ω instead of events in general. But as with the likelihood functions above, the probability of any event A is determined by summing the probabilities of $\omega \in A$. Because of the way in which multiplication figures into the geometric pooling recipe, there are profiles for which $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(\omega) = 0$ for all $\omega \in \Omega$, in violation of the probability axioms. If for each $\omega \in \Omega$ there is a $\boldsymbol{p}_i \in (\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ such that $\boldsymbol{p}_i(\omega) = 0$ we have such a violation. To avoid this worry, the domain of geometric pooling operators is restricted to profiles of *regular* pmfs, i.e., those $\boldsymbol{p}$ such that $\boldsymbol{p}(\omega) > 0$ for all $\omega \in \Omega$. We denote the set of regular pmfs $\mathbb{P}'$ making the relevant domain $\mathbb{P}'^n$.[5]

A third, more recent proposal from Dietrich (2010) is given by the following formula.

Multiplicative Opinion Pools. $F(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(\omega) = c\prod_{i=0}^{n} \boldsymbol{p}_i$, where $\boldsymbol{p}_0$ is a fixed "calibrating" probability function, and $c = \frac{1}{\sum_{\omega' \in \Omega} [\boldsymbol{p}_0(\omega)] \cdot [\boldsymbol{p}_1(\omega')] \cdots [\boldsymbol{p}_n(\omega')]}$ is a normalization factor.

As with geometric pooling functions, the domain of multiplicative pooling functions will be restricted to $\mathbb{P}'^n$. Comments on the interpretation and choice of $\boldsymbol{p}_0$ can be found in (Dietrich and List, 2014, pp. 17-19)

Various results, both characterization and limitative, for the different pooling operations and axioms have been obtained. For example, SSFP characterizes linear pooling.

Theorem 1. *(McConway, 1981, Theorem 3.3; Lehrer and Wagner, 1981, Theorem 6.7) Given that the algebra contains at least three disjoint events, a pooling function satisfies SSFP iff it is a linear pooling function.*

Theorem 2. *(McConway, 1981, Corollary 3.4) Given that the algebra contains at least three disjoint events, F satisfies WSFP and ZPP iff F is a linear pooling function.*

McConway has shown that a pooling function has the WSFP *iff* it has the MP. So linear pooling functions satisfy MP and ZPP.

Theorem 3. *(Genest, 1984, p. 1104) The geometric pooling functions are externally Bayesian and preserve unanimity.*

Other sorts of pooling functions, such as a certain generalization of geometric pooling, satisfy the conditions of Theorem 3. A characterization of externally Bayesian pooling functions is given in (Genest et al., 1986). Dietrich and List provide a characterization of multiplicative pooling.

---

[4]An unweighted geometric pool of n numerical values is given by $\sqrt[n]{x_1 \cdots x_n} = x_1^{\frac{1}{n}} \cdots x_n^{\frac{1}{n}}$.

[5]Rather than assuming regularity or that the algebra contains the elements of Ω, we could make the weaker restriction to the domain of profiles of pmfs such that there is some $\omega \in \Omega$ for which $\boldsymbol{p}_i(\omega) > 0$ for all $i = 1, ..., n$. And pmfs allow us to obtain measures defined on general algebras on Ω.

Theorem 4. *(Dietrich and List, 2014, Theorem 3) The multiplicative pooling functions are the only individualwise Bayesian pooling functions (with domain $\mathbb{P}'^n$).*

There are many Arrovian limitative theorems in the pooling literature. As Robert Nau notes, none of the pooling methods discussed satisfy even unanimity, external Bayesianity, and the marginalization property (2002, p. 266). (As we show below, our proposal in this paper *does* satisfy those properties.) One result that we have occasion to appeal to below follows from results due to Lehrer and Wagner (1983, Theorems 1 and 2) in conjunction with Theorem 1 above:

Theorem 5. *(Cf. Lehrer and Wagner, 1983) Given that the algebra contains at least three pairwise disjoint events, the only pooling functions that preserve probabilistic independence and satisfy SSFP are dictatorial.*

It follows that non-dictatorial *linear pools* do not preserve probabilistic independence. In general, non-dictatorial pooling methods struggle with independence preservation (Genest and Wagner, 1987). (Here, too, we claim to do better.)

3. Motivations for IP

In general terms, imprecise probabilities (IP) models do not require representing an agent's or group's judgments of subjective probability as numerically precise. Instead, such judgments could be represented by *sets* of probability functions (Kyburg and Pittarelli, 1996), for example, or by *intervals* (Kyburg, 1998).

There are a number of motivations for working with IP models. These include the potential to resolve some of the "paradoxes of decision" (Ellsberg, 1963; Levi, 1986b), allowing for more flexible and less arbitrary models of uncertainty (Gärdenfors and Sahlin, 1982; Walley, 1991), allowing for incomplete preferences (and hence judgments of incomparability) in the subjective expected utility setting (Levi, 1986a; Seidenfeld, 1993; Kaplan, 1996), and increased descriptive realism (Arló-Costa and Helzner, 2010). An overview of these and other motivations for IP can be found in (Bradley, 2014).

Most important for present purposes, IP allows for—what we consider—a very interesting and philosophically well-motivated account of consensus (Levi, 1985; Seidenfeld et al., 1989). Our goal in this section is to present this account of consensus for explicit consideration in the context of pooling. It may help to first consider the case of *full* or *plain* belief. At the outset of inquiry, inquirers may seek consensus as *shared agreement* in their beliefs. This could be achieved by retaining whatever beliefs are common to all parties and suspending judgment on those that are controversial thereby avoiding question-begging. Importantly, the consensus is generally a *weaker* state of belief. Since inquiry initiating from the consensus view proceeds without begging questions against parties to the consensus, various hypotheses of concern can receive a fair hearing. Such a consensus constitutes a neutral or non-controversial starting point for subsequent inquiry.

The idea that parties to a joint effort in inquiry or decision making should restrict themselves to their shared agreements—as a compromise or as genuine consensus—can be extended to judgments of probability. An analogous sense of suspending judgment concerning what is controversial is available in the IP setting. To suspend judgment among some number of probability distributions is to not rule them out for the purposes of inference and decision making. Put another way, to suspend judgment among some number of distributions is to regard each as *permissible* to use in inference and decision making. If the parties seeking consensus all agree that $\boldsymbol{p}$ is *not* permissible, then the consensus position reflects that agreement and rules it out (this will have to be finessed when we come to the question of convexity below). A *set* of probability functions represents the shared agreements among the group concerning which probability functions *are not* permissible to use in inference and decision making. For example, it is consensus that the probability of some event is not below the minimum of individual assignments.

Many authors refer to the output of a pooling function as a *consensus* (Lehrer and Wagner, 1981; McConway, 1981; Genest and Zidek, 1986). In what way is a precise pool a consensus? Isaac Levi draws a distinction between consensus as the *outcome* of inquiry and consensus at the *outset* of inquiry (1985). At the outset of inquiry, agents may seek common ground upon which to pursue joint inquiry. This is consensus as shared agreement, discussed just above. Disagreement among the parties to the consensus may then be resolved (in the best case) through joint efforts in inquiry—consensus as the outcome of inquiry. Given the restriction that consensus must be representable by a unique probability function, outside of the special case when all individuals are in total agreement, finding consensus as shared agreement that suspends judgment on unshared probabilistic views is a foreclosed possibility.

The individuals could assume some common, precise probability distribution, but Levi argues this is not consensus as common ground:

> there can be no analogue in contexts of probability judgment of the two senses of consensus I identify. If two or more agents differ in probability judgment, they can all switch either to the distribution adopted by one of them or to some other distribution which is, in a sense, a potential resolution of the conflict between their differing distributions. There is only one kind of consensus to be recognized—namely the resolution of conflict reached through revolution, conversion, voting, bargaining or some other psychological or social process. (1985, pp. 5-6)

Wagner likewise distinguishes between a *compromise* adopted to perform an exercise in group decision making and a *consensus* to which the individuals revise their own beliefs (Section 2). Levi's point in the quotation above is that a precise pool is neither a consensus as shared agreement since, in general, it is not restricted to just the shared probabilistic views; nor is it justified on the basis of inquiry, understood as designing and performing experiments, obtaining evidence, etc. A precise pool might represent the sort of political consensus that a vote does in the case of preferences, or that the output of a judgment aggregation function does in the case of qualitative belief. That is, consensus as bargaining or compromise. Of course, at least one sort of compromise *is* a consensus adopted *for the sake of the argument* rather than genuinely. That is, parties to the compromise could assume the consensus position as Levi identifies it—namely, a convex IP set—for the sake of the argument, or for carrying out some group deliberation or inquiry so long as the consensus position is strong enough for the group's purposes. It must be admitted that there are other compromise positions, including precise pools, that the group might assume.

But Levi's view distinguishes between political and rational consensus. Returning again to the case of full belief, Levi requires that revisions be decomposable into a sequence of contractions and expansions. An inquirer's set of full beliefs constitute her "standard for serious possibility" in the following sense: if A is among her full beliefs, $\neg A$ is not a serious possibility. To change her mind, the inquirer must first suspend judgment on A by contracting A if she cares to avoid error (where error is judged by her own lights). From the contraction, both A and $\neg A$ are serious possibilities. A *direct* revision to include $\neg A$ involves deliberately importing error from the point of view that rules $\neg A$ out as a serious possibility.

Unlike full beliefs, however, judgments of subjective probability do not bear truth values. So how might one suspend judgment among candidate distributions before changing points of view? As discussed just above, subjective probabilities are used in determining expectations for available acts. Levi's proposal is that to suspend judgment among some number of distributions is to not rule them out for use in the functions that they perform in inquiry and deliberation. Coming to regard a probability distribution as *permissible* is the analogue of opening one's mind to the (serious) *possibility* of $\neg A$ in the case of full belief. Just as assuming a weaker position in full belief avoids begging questions, retreating to a superset of distributions avoids prejudging the issue of determining which distributions are permissible for use in inquiry and deliberation. Moving from a set of probability functions (including a singleton) to a superset is the probabilistic analogue of contraction. Employing sets of probability functions avoids demanding direct revisions to probabilistic judgments the agent regards as *impermissible* from the standpoint of her current probabilistic judgments without first "contracting" to a neutral position that suspends judgment among the relevant probabilistic views (Levi, 1974). As he emphasizes, both reaching common ground at the outset of inquiry and subsequent reasoned changes in probabilistic views are available to groups in the IP setting.

Why has most work on probabilistic aggregation restricted itself to consideration of representations in terms of a single probability function? One reason is that such representation is the standard for individuals and, since we are treating groups as agents in a sense, that representation should extend to groups as well. Genest and Zidek write, "it would be natural to express the consensus judgment in the same form as the originals" (1986, p. 115). But the authors of the present essay are not moved by this convention (or in Walley's terminology, by this "Bayesian dogma of precision") for some of the very reasons as discussed just above. We urge, in what follows, that theorizing concerning IP should be extended to accounts of probabilistic opinion pooling and *vice versa*. Even if the motivations for IP in general, including at the level of individuals, are found less than convincing, one might think that the case for IP at the level of group opinion is more persuasive, say as an account of consensus. We take the motivations above and the propositions that follow as recommending further consideration of IP in the context of pooling and consensus.

4. IP Pooling Formats

We want to make a case that the out-of-the-gate restriction of the codomain of F to $\mathbb{P}$ is unwarranted (just as many have argued that the standard Bayesian assumption that rational individuals are committed to determinate probabilistic judgments is unwarranted). Our strategy is to point to a sensible account that abandons that restriction. Here we assume a representation in terms of a set of probability functions. We make use of *set-valued* functions or *correspondences*. Where F refers to a pooling function that outputs a single probability function, we will use $\mathcal{F}$ to refer to pooling correspondences outputting sets of probability functions:

$$\mathcal{F} : \mathbb{P}^n \to \mathscr{P}(\mathbb{P})$$

4.1. **Convex or Not?**

Much of the work with IP assumes that IP sets of probabilities are *convex* (Smith, 1961; Levi, 1974; Girón and Ríos, 1980; Gilboa and Schmeidler, 1989; Walley, 1991; Moral and Del Sagrado, 1998). A set of probability functions, $\boldsymbol{P}$, is convex if, for any two functions in the set, the set includes every convex combination of those functions.

Convexity. If $\boldsymbol{p}_1, \boldsymbol{p}_2 \in \boldsymbol{P}$, then $\alpha\boldsymbol{p}_1 + (1-\alpha)\boldsymbol{p}_2 \in \boldsymbol{P}$ for $\alpha \in [0,1]$.

Besides some handy computational properties of convex sets of probability functions (Girón and Ríos, 1980), convexity can be motivated philosophically. A set of probability functions represents the shared agreements among the group members concerning which distributions are ruled out for use in deliberation and inquiry. But is it not common ground that the convex combinations of individual probability functions are ruled out? The idea is that convexity recommends a *weaker* attitude in suspending judgment among some number of probability distributions; fewer distributions are ruled out. Convexity requires keeping an open mind concerning potential compromises or resolutions of conflict (the convex combinations) between various probabilistic views. Levi argues that convex combinations have "all the earmarks of potential resolutions of the conflict; and, given the assumption that one should not preclude potential resolutions when suspending judgment between rival systems [...] all weighted averages of the two functions are thus taken into account" (1980, p. 192).

The normative status of convexity is the subject of outstanding controversy. Seidenfeld, Schervish, and Kadane make a case against convexity in the context of group decision making (1989). They observe that if two Bayesian agents differ in both probability and utility, any compromise position in probability besides *trivial* convex combinations entails a violation of a Pareto constraint on preference. Levi responds in his (1990), arguing against the Pareto condition. Kyburg and Pittarelli lodge some complaints about the property in "Some Problems for Convex Bayesians" (1992). In "Set-Based Bayesianism," they explore relaxing convexity to allow for IP sets in general. Seidenfeld et al.'s theory of coherent choice does not require convexity (2010). They offer a variation of one of Kyburg and Pittarelli's criticisms of convexity, registering a counterexample that exploits the failure of convex combinations to preserve probabilistic independence (but see (Levi, 2009, pp. 373-375) for a response).

Depending on the decision theory, distinctions between pooling formats may or may not be of importance. For example, there are decision rules that cannot distinguish between certain convex and non-convex sets of probabilities (Gilboa and Schmeidler, 1989; Walley, 1991). Such distinctions are meaningful according to other decision rules (Levi, 1980). And there are decision rules that distinguish between any two sets of probabilities (Seidenfeld et al., 2010). The important point here is that disputes over the format of pooling functions are idle if such distinctions are not decision-theoretically meaningful. On decision theories that cannot distinguish a convex set of probabilities from its extreme points, for example, there is nothing at stake in arguments over whether an IP opinion pool is convex or not.

4.2. **Convex Pooling Functions.** As a proof of concept, we will investigate aggregation functions that output sets of probability functions. The aggregate is formed by taking the *convex hull* of the n probability distributions:

$$\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = conv\{\boldsymbol{p}_i : i = 1, ..., n\}$$

The convex hull of a set of points is the smallest convex set containing those points. We write $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$ as shorthand for the set of probability assignments to A:

$$\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}$$

We work with convexity, not because we presume to know of decisive arguments in its favor, but because convexity is a broadly customary assumption, we do not yet feel compelled to dismiss it, and it allows us to make a proof of concept argument for IP pooling. In effect, assuming convexity amounts to making it slightly harder to show some of the propositions below, though the propositions also hold for IP aggregation methods that relax convexity. We return to the issue of convexity below to make good on our earlier promise to clarify how our case for IP in the context of opinion aggregation does not depend entirely on *convex* IP pools (Section 7.3, Proposition 7).

5. Extending Pooling Axioms to the IP Setting

Convex IP pooling functions satisfy the extensions of those axioms to the IP setting. For the SSFP, we replace G with a set-valued function or correspondence: $\mathcal{G} : [0,1]^n \to \mathscr{P}([0,1])$. $\mathcal{G}$ is a map from n numerical values in $[0,1]$ to a set of probability values, $\mathcal{G}(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot))$. The constraint becomes that there exists a function $\mathcal{G}$ such that, for any event A, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{G}(\boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A))$. WSFP, then, requires a function $\mathcal{G} : \mathscr{A} \times [0,1]^n \to \mathscr{P}([0,1])$. For unanimity preservation, we do not distinguish between $\boldsymbol{p}$ and $\{\boldsymbol{p}\}$. ZPP is generalized analogously. If $\boldsymbol{p}_i(A) = 0$ for all $i = 1, ..., n$, then $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \{0\}$. The MP has a straightforward extension to sets of probability functions: $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}([\boldsymbol{p}_1 \upharpoonright_{\mathscr{A}'}], ..., [\boldsymbol{p}_n \upharpoonright_{\mathscr{A}'}])(A)$, for any $A \in \mathscr{A}'$ There are many ways to generalize constraints. While we offer conservative and natural modifications of the axioms in order to extend them to the imprecise setting, the crucial question is whether we have modified what is *compelling* about the axioms. For example, is representation in terms of a unique probability function crucial to what makes commutativity of conditionalization and pooling compelling, or what is appealing about preserving shared judgments of independence? In each case, we submit, the attractiveness of the axiom does not hinge on whether the output of the aggregation function is a single probability function or a set of them.

First, we note that an analogue of McConway's result holds for IP pooling functions in general.

Proposition 1. *Let $\mathcal{F} : \mathbb{P}^n \to \mathscr{P}(\mathbb{P})$ be an IP pooling function (not necessarily convex). $\mathcal{F}$ satisfies WSFP iff $\mathcal{F}$ satisfies MP.*

Before stating the next proposition, we record a fact about convex sets of probabilities (simple and familiar to those with a background in geometry) that we will make use of in the proof.[6] Proofs for the lemmas and propositions are recorded in the appendix.

Lemma 1. *Let $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = conv\{\boldsymbol{p}_i : i = 1, ..., n\}$ for any profile $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ in the domain of $\mathcal{F}$. Any $\boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ can be expressed as a convex combination of the n probability functions, i.e. $\boldsymbol{p} = \sum_{i=1}^n \alpha_i \boldsymbol{p}_i$, where $\alpha_i \geqslant 0$ for $i = 1, ..., n$ and $\sum_{i=1}^n \alpha_i = 1$.*

Proposition 2. *Convex IP pooling functions satisfy SWFP, WSFP, MP, ZPP, and unanimity preservation.*

As indicated in the proof, SSFP entails both WSFP and ZPP.

While linear pooling functions are not externally Bayesian, convex IP pooling functions satisfy the extension of external Bayesianity to the IP setting. The *convex* or *prior-by-prior* conditionalization of a convex set of probability functions, $\mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, results from conditionalizing each member of the set. Updating a convex set of probability functions on a common likelihood function is defined analogously:

$$\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \left\{\boldsymbol{p}^\lambda : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n), \sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega') > 0, \text{ and } \boldsymbol{p}^\lambda(\cdot) = \frac{\boldsymbol{p}(\cdot)\lambda(\cdot)}{\sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega')}\right\}$$

To show that convex IP pooling functions are externally Bayesian, we first state another observation.

Lemma 2. *(Cf. Levi, 1978; Girón and Ríos, 1980) Convexity is preserved under updating on a likelihood function, i.e., $\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is convex.*

Proposition 3. *Convex IP pooling functions are externally Bayesian.*

Dietrich and List argue that while geometric pooling is justified on epistemic grounds when individual opinions are based on the same information, multiplicative pooling is justified in cases of *asymmetric* information, when individual opinions are in part based on private information. Their case is built around the individualwise Bayesianity axiom and the fact that multiplicative pooling satisfies it (Theorem 4).

Proposition 4. *Convex IP pooling functions are* ***not*** *individualwise Bayesian.*

We regard Proposition 4, however, as stating a feature and not a bug of convex IP aggregation. At least insofar as the idea is to reach a consensus, it is not desirable for features of one individual's probability distribution to be unilaterally imposed on the group. In the case of full belief, the initial consensus does not adopt just any belief of any member. Better, in our view, for group opinion to change through efforts in intelligently conducted inquiry from initial common ground (in inquiry, a group may designate a subgroup as an information source on a given topic, but this process requires a richer representation). Dietrich and List motivate individualwise Bayesianity by pointing out that if the constraint is not satisfied, then it makes a difference if an individual first learns some information and opinions are then pooled, or if the opinions are pooled and then the information is acquired by the group as a whole. But for consensus, this is as it should be. If the opinions of the group members do not reflect some piece of information, that information is not

[6]We include a proof of the observation because we appeal to it several times in the other proofs, because it is a simple special case (but all we need) of a more general result concerning convexity (Rockafellar, 1970), and because some readers may not have a conceptual handle on the property.

common ground. The consensus among group members depends on the probabilistic opinions of the members.

None of this is to say, of course, that features of individual probability distributions are irrelevant to group consensus. On the convex IP view, the kernel of truth in individualwise Bayesianity can be formulated by the inequalities below, stated here for standard conditionalization.

$$\begin{aligned} & min\{\boldsymbol{p}(B) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\} \\ \leqslant \; & min\{\boldsymbol{p}(B) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_i^B, ..., \boldsymbol{p}_n)\} \\ \leqslant \; & min\{\boldsymbol{p}(B) : \boldsymbol{p} \in \mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\} \end{aligned}$$

And similarly,

$$\begin{aligned} & max\{\boldsymbol{p}(B) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\} \\ \leqslant \; & max\{\boldsymbol{p}(B) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_i^B, ..., \boldsymbol{p}_n)\} \\ \leqslant \; & max\{\boldsymbol{p}(B) : \boldsymbol{p} \in \mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}. \end{aligned}$$

The consensus probabilities for B shift up (at least not down, more precisely) if one individual conditionalizes on B, and shift more if the consensus itself conditionalizes on B. But these inequalities simply reflect facts about what the common ground is and do not reflect group "learning" from one individual's probability function.

Convex IP pooling also inherits some of the challenges facing linear pooling. SSFP conflicts with probabilistic independence preservation. As Theorem 5 states, the only pooling functions that preserve probabilistic independence and satisfy SSFP are dictatorial. The loss of probabilistic independence presents both epistemic challenges as well as decision theoretic ones (Kyburg and Pittarelli, 1992; Seidenfeld et al., 2010).

In the case of convex IP pooling, however, there is more leeway to address the challenges. Several generalizations of the concept of independence for IP have been proposed and studied (de Campos and Moral, 1995; Cozman, 1998). We consider Levi's notion of *confirmational irrelevance*.

Confirmational Irrelevance Preservation. If $\boldsymbol{p}_i(A|B) = \boldsymbol{p}_i(A)$ for $i = 1, ..., n$, then $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$.[7]

Irrelevance preservation is a generalization of probabilistic independence preservation. It is clear that when $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is a single probability function, irrelevance preservation reduces to independence preservation. According to some decision theories for IP, it is the whole set $\boldsymbol{P}$ that is relevant for inquiry and decision making (Levi, 1980; Seidenfeld et al., 2010). Irrelevance is a sensible generalization of independence because it allows us to identify when some information will not make a difference to certain inquiries or deliberations, namely, those inquiries and deliberations concerning events to which the information is irrelevant.

It also does not take much work to show that confirmational irrelevance preservation is satisfied by any IP pooling function (not necessarily convex) that satisfies *stochastic independence preservation.*

Stochastic Independence Preservation If $\boldsymbol{p}_i(A \cap B) = \boldsymbol{p}_i(A)\boldsymbol{p}_i(B)$, for $i = 1, ..., n$, then, for all $\boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, $\boldsymbol{p}(A \cap B) = \boldsymbol{p}(A)\boldsymbol{p}(B)$.

It turns out that confirmational irrelevance *is*, but stochastic independence is *not*, preserved by convex IP pooling functions. Suppose that A and B are probabilistically independent according to $\boldsymbol{p}_i$, $i = 1, ..., n$. Since linear pooling does not preserve independence, independence is not preserved at some of the interior, non-extreme points of $\boldsymbol{P}$. However, the whole set of probability *values for* A, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$, is the same before and after conditionalizing on B.

Proposition 5. *Convex IP pooling functions satisfy irrelevance preservation.*

[7]This binary case of irrelevance can be generalized to non-binary partitions. Let $A_1, ..., A_k$ be a partition of Ω. In Levi's setup, a question is represented as a partion, each element of which is a potential answer. Information B is pairwise irrelevant to $A_1, ..., A_k$ if B is irrelevant to each cell of the partition.

So, while stochastic independence and confirmational irrelevance are equivalent in the precise setting (when $\boldsymbol{p}(B) > 0$), they come apart in the IP context.[8] Because irrelevance preservation reduces to probabilistic independence preservation when the output of the pooling function is a unique probability function, and linear, geometric, and multiplicative pooling functions do not satisfy probabilistic independence preservation in general, we have that linear, geometric, and multiplicative pooling functions do not satisfy irrelevance preservation either. If there are good reasons to require IP pooling functions to satisfy the stronger stochastic independence preservation property, then convex IP pool does not deliver (though there are IP formats that do (Proposition 7)).

Finally, convex IP pooling admits of a simple characterization in terms of the set of *universally admissible means*.[9] We call a function $\mathfrak{m} : [0,1]^n \to [0,1]$ a *mean* on the interval $[0,1]$. We first define a mapping $\mathfrak{M}_n : \mathbb{P}^n \to \mathscr{P}([0,1]^{[0,1]^n})$ by setting for every $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$:

$$\mathfrak{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \Big\{\mathfrak{m} \in [0,1]^{[0,1]^n} : \mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) \in \mathbb{P}\Big\}.$$

Call a mean *admissible* for $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ if $\mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) \in \mathbb{P}$. Then, $\mathfrak{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is the set of admissible means for $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. Using $\mathfrak{M}_n$, we define another mapping $\mathcal{M}_n : \mathbb{P}^n \to \mathscr{P}(\mathbb{P})$ by setting for every $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$:

$$\mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \Big\{\mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) : \mathfrak{m} \in \bigcap_{\vec{\boldsymbol{q}} \in \mathbb{P}^n} \mathfrak{M}_n(\vec{\boldsymbol{q}})\Big\}.$$

$\mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is the set of probability functions that results from composing each *universally* admissible mean with $\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big)$.

Proposition 6. *Suppose that $\mathscr{A}$ contains at least three pairwise incompatible events. A mapping $\mathcal{F} : \mathbb{P}^n \to \mathscr{P}(\mathbb{P})$ is a convex IP pooling function—that is, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = conv(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$—if and only if $\mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ for all $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$.*

An interesting line of research to pursue would be to consider pooling from the perspective of an analysis of means (e.g., Kolmogorov, 1930; de Finetti, 1931; Aczél, 1948). Perhaps such an analysis could shed light on issues like the propriety of qualitative conditions on pooling rules, or the function of convexity in reaching a consensus in inquiry and deliberation.

6. Epistemic and Procedural Grounds for IP Accounts

In their review article, Dietrich and List claim that the question of how probabilities should be aggregated admits of no obvious answer, and, ultimately, the appropriateness of the pooling method depends on the purpose and context of aggregation (2014). They raise the question of whether a pooling method should be justified on *epistemic* or *procedural* grounds. To be justified on epistemic grounds, "the pooling function should generate collective opinions that [...] respect the relevant evidence or track the truth, for example." In order to be justified on procedural grounds, a pooling method should yield a collective opinion that is a "fair representation of the individual opinions" (2014, p. 2). Dietrich and List claim that, while linear pooling can be justified on procedural grounds, it cannot be justified on epistemic grounds. By satisfying WSFP, linear pooling functions reflect "the democratic idea that the collective opinion on any issue should be determined by individual opinions on that issue" (2014, p. 6). Geometric pooling, however, can be justified in

[8]Pedersen and Wheeler show how logically distinct independence concepts are teased apart in the context of imprecise probabilities. They write, "there are several distinct concepts of probabilistic independence and [...] they only become extensionally equivalent within a standard, numerically determinate probability model. This means that some sound principles of reasoning about probabilistic independence within determinate probability models are invalid within imprecise probability models" (2014, p. 1307). So IP provides a more subtle setting in which to investigate independence concepts.

[9]We thank the referee for suggesting that we include a result along these lines.

epistemic terms "by invoking the axiom of external Bayesianity" (2014, p. 13). The idea seems to be that since updating is a response to the evidence, a pooling method that is well-behaved in the sense of commuting with updating "respects the relevant evidence" by not allowing the order of operations to distort evidential impact.

Convex IP pooling, then, can be justified on Dietrich and List's procedural grounds in virtue of satisfying WSFP. Concerning procedural grounds in general, it is difficult to think of a more fair or democratic representation of individual opinions than a representation that *includes* each opinion and all of the compromises between opinions. But since convex IP pooling functions also satisfy external Bayesianity, it would thus appear that the alleged tension between epistemic and procedural criteria for probabilistic opinion aggregation can be resolved by simply moving to an IP account.

We endorse the basic motivation for Dietirch and List's discussion. Like other deliberate activities, pooling is *goal-directed*. How one should approach pooling depends on, among other things, one's goals. One may have multiple goals, in which case, tensions in jointly satisfying them may require tradeoffs. Nevertheless, we find it rather opaque precisely how WSFP encodes an intuitive procedural constraint on pooling. Similarly, how commutativity of pooling and updating ensures that the collective opinion respects the relevant evidence or tracks the truth stands in need of further clarification. Even if the *philosophical* distinction and interpretation of WSFP and external Bayesianity does not admit of further clarification, however, our point stands that the tension between satisfying the *formal* desiderata can be resolved in the IP setting.

7. Objections to IP Pooling

Perhaps the relative neglect of IP in discussions of pooling can be explained in part by a skepticism concerning the *use* to which IP sets can be put. In their very nice overview of work on pooling, Genest and Zidek write, "the jury remains out on the theory of Walley [...] In particular, it is unclear how [the IP set] could be used 'at the end of the day'" (1986, p. 124). There are essentially two types of uses to which an account of probabilities may be put: those concerning epistemic issues like inference, and those concerning issues in decision making.

7.1. **Epistemology.** Some degree of the skepticism about the epistemic usefulness of IP may be dispelled by considering recent work. For instance, after Genest and Zidek's article, Walley published his magisterial book addressing applications of IP to issues in statistical reasoning (1991). Fabio Cozman explores the application of IP to issues in Bayesian networks in a number of papers (1998; 2000).

But let us consider some epistemological challenges of a general sort. In reviewing some difficulties for the few available accounts of pooling IP sets of probabilities (accounts allowing imprecision at the individual level), Robert Nau claims that neither taking the union nor the intersection of convex sets of imprecise probabilities yields a satisfactory account of pooling. He writes,

> As more opinions are pooled, the union can only get larger, and it reflects only the least informative opinions, whereas intuitively there ought to be (at least the possibility of) an increase in precision as the pool gets larger. On the other hand, the intersection of convex sets of measures may be empty if experts are mutually incoherent, and it generally yields too tight a representation of aggregate uncertainty. As more opinions are pooled, the intersection can only shrink, and it reflects only the most extreme among those opinions, whereas intuitively there should be some convergence to an average opinon when the pool gets sufficiently large. Moreover, neither the union nor the intersection provides an opportunity for the differential weighting of opinions, which would be desireable in cases where one individual is

> considered (either by herslef or by an external evaluator) to be better or worse informed than another individual about a particular event under consideration. (2002, p. 267)

Similar concerns could be expressed about the account of pooling under examination in this essay since uncertainty never decreases by mere pooling on our account. But we think they would misplaced. The appropriateness of the behavior of a pooling function cannot be assessed in abstract, without specifying the *point* of pooling probabilities in the first place. If the point is to find common ground among the opinions being pooled, increasing uncertainty is to be expected. In general, the more opinions among which we try to find common ground, the less common ground there will be.[10] One might not wish to seek consensus among certain opinions, but that is a different matter. On our account uncertainty *can* be reduced, but through inquiry and not through pooling. As the group acquires sufficient information, conditionalization generally leads to a reduction of imprecision. In the IP setting, it is also possible, however, for conditionalization to *increase* imprecision in the short run, a phenomenon known as *dilation* (Seidenfeld and Wasserman, 1993; Wasserman and Seidenfeld, 1994; Herron et al., 1997; Pedersen and Wheeler, 2014, 2015). But our point here is not that conditionalization invariably decreases uncertainty, but that it can and that decreasing uncertainty through conditionalization has familiar Bayesian "learning" foundations whereas pooling (averaging) does not.

Furthermore, in the case of pooling imprecise probabilities, we would not endorse taking intersections for the purpose of finding consensus. In the case of mutual incoherence, intersections yield the empty set. But the lack of any consensus concerning which probability functions can be ruled out means that the group in consensus cannot rule any probability functions out. Taking the convex hull of the union would reflect this, yielding complete uncertainty.[11]

We think it is important to distinguish between finding consensus among some opinions and taking those opinions as evidence. In the latter case, if an agent outside the group considers some members of the group to be less informed than others, *that* opinion should be reflected in conditionalization through the likelihood for the experts' opinions (Cf. the *Supra-Bayesian* approach to pooling (Genest and Zidek, 1986, p. 120)). In the former case, if a group member is considered, by herself or other group members, to be less informed, consensus is often not sought. Finding what common ground the group members share is unproblematic when consensus is sought, regardless of the social, political, or intellectual clout members accord each other. It is also open to, and perhaps rationally obligatory for, the modest group member to allow her opinion concerning her relative informedness to be reflected in her probabilistic opinion before pooling.

Finally, one might object that IP pooling amounts to declining to really aggregate. In a sense, that is true, if pooling is restricted to taking some sort of average of individual probabilities. But, again, what is the theoretical basis for only considering precise averages of subjective probabilities? An IP set clearly *represents* group opinion, and can be employed in inference and decision making.

7.2. **Decision Theory.** Because decision theory is a very involved topic and we do not treat it in this paper, we limit ourselves to pointing out that sophisticated decision theories for IP have been developed and extensively studied. These include Levi's E-admissibility and tie-breaking decision rule (1980), Girón and Rios' quasi-Bayesian decision theory (1980), Gilboa and Schmeidler's Γ-Maximin (1989), and Walley's *Maximality* (1991). Seidenfeld, Schervish, and Kadane axiomatize their theory of coherent choice under uncertainty in the framework of set-valued choice functions (2010). Though saying so overcommits us for our project in this essay, we hold the view that

[10]We suspect that Nau is not targeting consensus because his models of pooling involve game-theoretic bargaining scenarios pitting the opinions to be aggregated against each other.

[11]See Larry Wasserman's review of Walley's book for objections to this representation of complete uncertainty (1993), and Levi's concept of *confirmational commitment* as a potential means of addressing the objections (1974).

theoretical disputes about probability cannot be adjudicated without thorough decision theoretic considerations.

7.3. **Convexity Revisited.** How much do the results in this paper depend on the convexity of the IP set? Not much! To see why, consider the following very simple IP pooling function, $\mathcal{F} : \mathbb{P}^n \to \mathscr{P}(\mathbb{P})$, such that

$$\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \{\boldsymbol{p}_i : i = 1, ..., n\}$$

So defined, $\mathcal{F}$ takes as input a profile of probability functions and returns the set of functions in that profile.

Proposition 7. *Let $\mathcal{F} : \mathbb{P}^n \to \mathscr{P}(\mathbb{P})$ be an IP pooling function such that, for each profile in $\mathbb{P}^n$, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \{\boldsymbol{p}_i : i = 1, ..., n\}$. Then, $\mathcal{F}$ satisfies SSFP, WSFP, ZPP, MP, unanimity preservation, external Bayesianity, and confirmational irrelevance preservation. Moreover, $\mathcal{F}$ satisfies stochastic independence preservation.*

The proof of Proposition 7 is straightforward and so is omitted here. The upshot is that pooling with imprecise probabilities is promising in a robust sense. So, while the convex IP pooling model is the chief subject of our philosophical approbation and mathematical analysis in this essay, our case for the consideration of IP in the context of pooling does not rest exclusively with that model.

7.4. **Dynamics.** As we have presented convex IP pooling functions, the input is a profile of individual probability functions and the output is a convex set of probability functions. What if individual probabilities are themselves imprecise? Or what happens if we attempt to pool the probabilistic opinions of agents that are themselves groups?[12] As it stands, our account is silent. There is, however, a natural extension of the account on offer. Consonant with the philosophical position staked out here, the idea is to convexify the *union* of sets of probability functions.

$$\mathcal{F} : \mathscr{P}(\mathbb{P})^n \to \mathscr{P}(\mathbb{P})$$

Where the profile consists of *n sets* of probability functions, $(\boldsymbol{P}_1, ..., \boldsymbol{P}_n)$, the pool is given by $\mathcal{F}(\boldsymbol{P}_1, ..., \boldsymbol{P}_n) = conv\{\bigcup_i \boldsymbol{P}_i\}$. We leave examination of this more complete account to future work.

8. Conclusion

According to standard Bayesian theory, personal probabilities are *subjective*. One route that has been explored for recovering some objectivity is establishing intersubjective agreement. There are, for example, the famous convergence theorems to the effect that, given non-extreme priors and a suitably large amount of evidence upon which to conditionalize, posteriors converge (Savage, 1954; Gaifman and Snir, 1982). Consensus in the (*very*) long-run, however, is not the only kind of consensus we may seek. Prior to inquiry, consensus as *shared agreement* is still possible, and desirable for joint efforts in inquiry. Convex IP pooling can be philosophically motivated as an account of such consensus.

Our objective has been to undermine the preconception that probabilistic opinion pooling should result in a representative probability function for the group. Our tack has been to explore another option, arguing that, even by the very lights of those working in pooling, this option is promising. We have the following summary (an "X" means the pooling method does not generally satisfy the property):

[12]The problem being raised is similar to one in the literature on AGM belief revision. The *principle of categorical matching* requires that the output of a belief revision operator be of the same format as the input. Otherwise, the account of belief revision, constructed for a certain input format, is silent about iterated belief revision (Gärdenfors and Rott, 1995). In the case of convex IP pooling functions, dynamics *of pooling* are defined so long as we are never pooling sets of probabilities.

Table 1. Pooling Method Report Card

	Linear	**Geometric**	**Multiplicative**	**Convex IP**
SSFP	✓	X	X	✓
ZPP	✓	✓	✓	✓
MP	✓	X	X	✓
WSFP	✓	X	X	✓
Unanimity Preservation	✓	✓	X	✓
External Bayesianity	X	✓	X	✓
Individualwise Bayesianity	X	X	✓	X
Irrelevance Preservation	X	X	X	✓

Perhaps the most sensible representation of group opinion, especially when pooling is interpreted as reaching consensus, is not in terms of a single probability function. At the very least, the arguments and results above may be read both as an exploration of extending the mathematical framework of opinion pooling to cover IP pooling, and as a plea for liberalism about pooling formats.

Appendix: Proofs

Proof of Proposition 1

Proof. We carry out McConway's proof with minimal adjustments made for our framework (pp. 411-412 1981, Theorem 3.1).

WSFP $\Rightarrow$ MP. Assume that $\mathcal{F}$ has the WSFP, i.e., there is a function $\mathcal{G} : \mathscr{A} \times [0,1]^n \to \mathscr{P}([0,1])$ such that $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{G}(A, \boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A))$. By WSFP, we have $\mathcal{F}([\boldsymbol{p}_1 \upharpoonright_{\mathscr{A}'}], ..., [\boldsymbol{p}_n \upharpoonright_{\mathscr{A}'}])(A) = \mathcal{G}(A, [\boldsymbol{p}_1 \upharpoonright_{\mathscr{A}'}](A), ..., [\boldsymbol{p}_n \upharpoonright_{\mathscr{A}'}](A))$. Since $\mathcal{G}$ is a function and $\boldsymbol{p}_i(A) = [\boldsymbol{p}_i \upharpoonright_{\mathscr{A}'}](A)$ for any $A \in \mathscr{A}'$ (all such $A \in \mathscr{A}'$ are also in $\mathscr{A}$), it follows that $\mathcal{G}(A, [\boldsymbol{p}_1 \upharpoonright_{\mathscr{A}'}](A), ..., [\boldsymbol{p}_n \upharpoonright_{\mathscr{A}'}](A)) = \mathcal{G}(A, \boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A)) = \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$. Hence, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}([\boldsymbol{p}_1 \upharpoonright_{\mathscr{A}'}], ..., [\boldsymbol{p}_n \upharpoonright_{\mathscr{A}'}])(A)$.

MP $\Rightarrow$ WSFP. Assume that $\mathcal{F}$ has the MP. Let $A \in \mathscr{A}$. We want to show that $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$ depends only on A and $\boldsymbol{p}_i(A), i = 1, ..., n$.

First, if $A = \varnothing$ or $A = \Omega$, then, since the range of $\mathcal{F}$ is $\mathscr{P}(\mathbb{P})$, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$ depends only on A and $\boldsymbol{p}_i(A), i = 1, .., n$, for any profile because, setting $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{G}(A, \boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A))$ and $\mathcal{F}(\boldsymbol{p}'_1, ..., \boldsymbol{p}'_n)(A) = \mathcal{G}(A, \boldsymbol{p}'_1(A), ..., \boldsymbol{p}'_n(A))$, it follows that $\mathcal{G}(A, \boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A)) = \mathcal{G}(A, \boldsymbol{p}'_1(A), ..., \boldsymbol{p}'_n(A))$.

Next, suppose that $\varnothing \neq A \neq \Omega$. Consider the σ-algebra $\mathscr{A}' = \{\varnothing, A, A^c, \Omega\}$. $\mathscr{A}$ contains A and has $\mathscr{A}'$ as a sub-algebra. By MP, then

$$\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}([\boldsymbol{p}_1 \upharpoonright_{\mathscr{A}'}], ..., [\boldsymbol{p}_n \upharpoonright_{\mathscr{A}'}])(A).$$

$\mathscr{A}'$ is uniquely defined by A and any probability over $\mathscr{A}'$ is uniquely determined by the probability of A under that distribution. So the righthand side of the equation above is determined by A and $\boldsymbol{p}_i \upharpoonright_{\mathscr{A}'}(A) = [\boldsymbol{p}_i \upharpoonright_{\mathscr{A}'}](A) = \boldsymbol{p}_i(A)$. □

Proof of Lemma 1

Proof. Let $Y = \{\boldsymbol{p} : \boldsymbol{p} = \sum_{i=1}^n \alpha_i \boldsymbol{p}_i$ such that $\alpha_i \geqslant 0$ for $i = 1, ..., n$ and $\sum_{i=1}^n \alpha_i = 1\}$. We want to show the following:

$$\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = conv\{\boldsymbol{p}_i : i = 1, ..., n\} = Y$$

The first equality we have by definition. In order to show the second equality, we have to show that Y is the smallest convex set containing $\{\boldsymbol{p}_i : i = 1, ..., n\}$. To show convexity, we show that for any two functions in Y, any convex combination of those functions is in Y. Suppose that $\boldsymbol{p}, \boldsymbol{p}' \in Y$. By assumption, $\boldsymbol{p} = \sum_{i=1}^n \alpha_i \boldsymbol{p}_i$ and $\boldsymbol{p}' = \sum_{i=1}^n \beta_i \boldsymbol{p}_i$. Consider $\boldsymbol{p}^\star = \gamma \boldsymbol{p} + (1-\gamma)\boldsymbol{p}' = \gamma(\sum_{i=1}^n \alpha_i \boldsymbol{p}_i) + (1-\gamma)\sum_{i=1}^n \beta_i \boldsymbol{p}_i$.

$$\begin{aligned}
\boldsymbol{p}^\star &= \gamma \textstyle\sum_{i=1}^n \alpha_i \boldsymbol{p}_i + (1-\gamma)\sum_{i=1}^n \beta_i \boldsymbol{p}_i \\
&= \textstyle\sum_{i=1}^n \gamma\alpha_i \boldsymbol{p}_i + \sum_{i=1}^n (1-\gamma)\beta_i \boldsymbol{p}_i \\
&= \textstyle\sum_{i=1}^n [\gamma\alpha_i \boldsymbol{p}_i + (1-\gamma)\beta_i \boldsymbol{p}_i] \\
&= \textstyle\sum_{i=1}^n [\gamma\alpha_i + (1-\gamma)\beta_i] \boldsymbol{p}_i \\
&= \textstyle\sum_{j=1}^n \delta_j \boldsymbol{p}_j
\end{aligned}$$

where $\delta_j = \gamma\alpha_j + (1-\gamma)\beta_j$. $\delta_j \geqslant 0$ for $j = 1, ..., n$ because every term is nonnegative. $\sum_{j=1}^n \delta_j = \sum_{i=1}^n [\gamma\alpha_i + (1-\gamma)\beta_i] = \sum_{i=1}^n \gamma\alpha_i + \sum_{i=1}^n (1-\gamma)\beta_i = \gamma\sum_{i=1}^n \alpha_i + (1-\gamma)\sum_{i=1}^n \beta_i = \gamma(1) + (1-\gamma)1 = 1$. Hence, $\boldsymbol{p}^\star \in Y$, so Y is convex. If Y were not the smallest such set, then there would be some convex $Z \subsetneq Y$ such that $\{\boldsymbol{p}_i : i = 1, ..., n\} \subseteq Z$. But for any $\boldsymbol{p} \in Y, \boldsymbol{p}$ is a convex combination of the elements in $\{\boldsymbol{p}_i : i = 1, ..., n\}$. Since Z is convex and contains the $\boldsymbol{p}_i$, it follows that $\boldsymbol{p} \in Z$, which is a contradiction. □

Proof of Proposition 2

Proof. Since $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}$, we let $\mathcal{G}$ of the SSFP be the convex hull operation applied to $\{\boldsymbol{p}_i(A) : i = 1, ..., n\}$. It is clear that $\mathcal{G}$ depends just on the individual probabilities for A. We need to show that

$$\{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\} = conv\{\boldsymbol{p}_i(A) : i = 1, ..., n\}.$$

Trivially, the lefthand side includes $\{\boldsymbol{p}_i(A) : i = 1, ..., n\}$. Suppose $\boldsymbol{p}(A), \boldsymbol{p}'(A) \in \{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}$. Since $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is convex, it follows immediately that any convex combination of $\boldsymbol{p}(A), \boldsymbol{p}'(A)$ is in $\{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}$. Finally, suppose that there is some convex $Z \subsetneq \{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}$ which contains $\{\boldsymbol{p}_i(A) : i = 1, ..., n\}$. But for any $\boldsymbol{p}(A) \in \{\boldsymbol{p}(A) : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)\}$, $\boldsymbol{p}(A)$ is a convex combination of the $\boldsymbol{p}_i(A)$ since every $\boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is such a convex combination of the $\boldsymbol{p}_i$ (Lemma 1). Hence, $\boldsymbol{p}(A) \in Z$, contrary to our supposition. So, the equality holds and the SWFP is satisfied.

But since SWFP clearly implies WSFP, WSFP is satisfied, too. By Proposition 1, it follows immediately that $\mathcal{F}$ has the MP.

Because $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is a set of probability functions, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(\varnothing) = \{0\}$. Let $\boldsymbol{p}_i(A) = 0$, $i = 1, ..., n$. Since there is a function, $\mathcal{G}$, such that $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{G}(\boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A))$, we have it that

$$\begin{aligned}
\{0\} &= \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(\varnothing) \\
&= \mathcal{G}(\boldsymbol{p}_1(\varnothing), ..., \boldsymbol{p}_n(\varnothing)) \\
&= \mathcal{G}(\boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A)) \\
&= \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)
\end{aligned}$$

So, ZPP follows from SWFP.

For any profile $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$, if all $\boldsymbol{p}_i$ are identical, then the convex hull is just $\{\boldsymbol{p}_i\}$. So $\mathcal{F}$ satisfies *unanimity preservation*.

□

Proof of Lemma 2

We generalize a proof of a result due originally to Girón and Rios and Levi (Levi, 1978; Girón and Ríos, 1980) for updating on an *event* to updating on a common likelihood function.

Proof. We want to show that $\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is convex. That is, given any two members, $\boldsymbol{p}^\lambda, \boldsymbol{p}'^\lambda \in \mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ and $\alpha \in [0, 1]$, $\boldsymbol{p}^\star = \alpha\boldsymbol{p}^\lambda + (1-\alpha)\boldsymbol{p}'^\lambda$ is in $\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. If there is a convex combination of $\boldsymbol{p}$ and $\boldsymbol{p}'$, $\boldsymbol{p}_\star$, such that $\boldsymbol{p}^\lambda_\star = \boldsymbol{p}^\star$, then the convexity of $\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is established as a consequence of the convexity of $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. Where $\boldsymbol{p}^\lambda_\star(\cdot) = \frac{\boldsymbol{p}_\star(\cdot)\lambda(\cdot)}{\sum_{\omega' \in \Omega} \boldsymbol{p}_\star(\omega')\lambda(\omega')} = \frac{\beta\boldsymbol{p}(\cdot)\lambda(\cdot) + (1-\beta)\boldsymbol{p}'(\cdot)\lambda(\cdot)}{\beta\sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega') + (1-\beta)\sum_{\omega' \in \Omega} \boldsymbol{p}'(\omega')\lambda(\omega')}$, for any α we want to find some β such that

$$\boldsymbol{p}^\star(\cdot) = \alpha\boldsymbol{p}^\lambda(\cdot) + (1-\alpha)\boldsymbol{p}'^\lambda(\cdot) = \frac{\beta\boldsymbol{p}(\cdot)\lambda(\cdot) + (1-\beta)\boldsymbol{p}'(\cdot)\lambda(\cdot)}{\beta\sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega') + (1-\beta)\sum_{\omega' \in \Omega} \boldsymbol{p}'(\omega')\lambda(\omega')} = \boldsymbol{p}^\lambda_\star(\cdot).$$

For $\beta = \frac{\alpha\sum_{\omega* \in \Omega} \boldsymbol{p}'(\omega*)\lambda(\omega*)}{\alpha\sum_{\omega* \in \Omega} \boldsymbol{p}'(\omega*)\lambda(\omega*) + (1-\alpha)\sum_{\omega* \in \Omega} \boldsymbol{p}(\omega*)\lambda(\omega*)}$, the equality is verifiable with some tedious algebra. □

Proof of Proposition 3

Proof. We must show that convex IP pooling functions are externally Bayesian, i.e., $\mathcal{F}(\boldsymbol{p}^\lambda_1, ..., \boldsymbol{p}^\lambda_n) = \mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ (provided the relevant profiles are in the domain of $\mathcal{F}$).

$\mathcal{F}(\boldsymbol{p}_1^\lambda, ..., \boldsymbol{p}_n^\lambda) \subseteq \mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. Trivially, for each $i = 1, ..., n$, $\boldsymbol{p}_i^\lambda \in \mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. By Lemma 2, $\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is convex. It follows that $conv\{\boldsymbol{p}_i^\lambda : i = 1, ..., n\} = \mathcal{F}(\boldsymbol{p}_1^\lambda, ..., \boldsymbol{p}_n^\lambda) \subseteq \mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$.

$\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \subseteq \mathcal{F}(\boldsymbol{p}_1^\lambda, ..., \boldsymbol{p}_n^\lambda)$. By Lemma 1, any $\boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ can be expressed as the convex combination of the n extreme points generating $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, i.e., $\boldsymbol{p} = \sum_{i=1}^n \alpha_i \boldsymbol{p}_i$ where $\alpha_i \geqslant 0$ for $i = 1, ..., n$ and $\sum_{i=1}^n \alpha_i = 1$. By definition,

$$\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \{\boldsymbol{p}^\lambda : \boldsymbol{p} \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \text{ and } \boldsymbol{p}^\lambda(\cdot) = \frac{\boldsymbol{p}(\cdot)\lambda(\cdot)}{\sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega')}\}$$

We show that any member of $\mathcal{F}^\lambda(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is identical to some member of $\mathcal{F}(\boldsymbol{p}_1^\lambda, ..., \boldsymbol{p}_n^\lambda)$.

$$\begin{aligned}
\boldsymbol{p}^\lambda(\omega) &= \frac{\boldsymbol{p}(\omega)\lambda(\omega)}{\sum_{\omega' \in \Omega} \boldsymbol{p}(\omega')\lambda(\omega')} && \text{[Definition]} \\
&= \frac{\sum_{i=1}^n \alpha_i \boldsymbol{p}_i(\omega)\lambda(\omega)}{\sum_{\omega' \in \Omega} \sum_{i=1}^n \alpha_i \boldsymbol{p}_i(\omega')\lambda(\omega')} && \text{[Lemma 1]} \\
&= \frac{\sum_{i=1}^n \alpha_i \boldsymbol{p}_i^\lambda(\omega) \cdot \sum_{\omega' \in \Omega} \boldsymbol{p}_i(\omega')\lambda(\omega')}{\sum_{\omega' \in \Omega} \sum_{i=1}^n \alpha_i \boldsymbol{p}_i(\omega')\lambda(\omega')} && [\boldsymbol{p}_i(\omega)\lambda(\omega) = \boldsymbol{p}_i(\omega)^\lambda \cdot \textstyle\sum_{\omega' \in \Omega} \boldsymbol{p}_i(\omega')\lambda(\omega')] \\
&= \textstyle\sum_{j=1}^n \beta_j \boldsymbol{p}_j^\lambda(\omega) \in \mathcal{F}(\boldsymbol{p}_1^\lambda, ..., \boldsymbol{p}_n^\lambda) && \text{[Algebra]}
\end{aligned}$$

where $\beta_j = \dfrac{\alpha_j \cdot \sum_{\omega' \in \Omega} \boldsymbol{p}_j(\omega')\lambda(\omega')}{\sum_{\omega' \in \Omega} \sum_{i=1}^n \alpha_i \boldsymbol{p}_i(\omega')\lambda(\omega')}$ with $\beta_j \geqslant 0$ for all $j = 1, ..., n$ and $\sum_{j=1}^n \beta_j = 1$.

□

Proof of Proposition 4

Proof. We provide a very simple type of counterexample to individualwise Bayesianity, though counterexamples are plentiful. Consider the profile $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ for $n = 2$ agents such that $\boldsymbol{p}_1 = \boldsymbol{p}_2$. Individualwise Bayesianity requires that $\mathcal{F}(\boldsymbol{p}_1, \boldsymbol{p}_2^\lambda) = \mathcal{F}^\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2)$ (provided both $(\boldsymbol{p}_1, \boldsymbol{p}_2)$ and $(\boldsymbol{p}_1, \boldsymbol{p}_2^\lambda)$ are in the domain of $\mathcal{F}$). By Proposition 3 (external Bayesianity), it follows that $\mathcal{F}^\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2) = \mathcal{F}(\boldsymbol{p}_1^\lambda, \boldsymbol{p}_2^\lambda)$. But since $\boldsymbol{p}_1 = \boldsymbol{p}_2$, it follows that $\boldsymbol{p}_1^\lambda = \boldsymbol{p}_2^\lambda$. By unanimity (Proposition 2), then, we have $\mathcal{F}(\boldsymbol{p}_1^\lambda, \boldsymbol{p}_2^\lambda) = \{\boldsymbol{p_i}^\lambda\}$, where $\boldsymbol{p}_i^\lambda = \boldsymbol{p}_1^\lambda = \boldsymbol{p}_2^\lambda$. However, in general $\boldsymbol{p}_i \neq \boldsymbol{p}_i^\lambda$ and so $\mathcal{F}(\boldsymbol{p}_1, \boldsymbol{p}_2^\lambda)$ is *not* a singleton. It follows that, in general, $\mathcal{F}(\boldsymbol{p}_1, \boldsymbol{p}_2^\lambda) \neq \mathcal{F}^\lambda(\boldsymbol{p}_1, \boldsymbol{p}_2)$.

□

Proof of Proposition 5

Proof. Suppose that $\boldsymbol{p}_i(A|B) = \boldsymbol{p}_i(A)$ for $i = 1, ..., n$. We want to show that $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$. Consider $\boldsymbol{p}^\star(A) \in \mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$ and $\boldsymbol{p}_\star(A) \in \mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$. By Lemma 1, $\boldsymbol{p}^\star(A) = \sum_{i=1}^n \alpha_i \boldsymbol{p}_i(A)$, for appropriate α_i. By Proposition 3 (external Bayesianity), $\mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}(\boldsymbol{p}_1^B, ..., \boldsymbol{p}_n^B)(A)$ (Proposition 3 holds for standard conditionalization since standard conditionalization is a special case of updating on a likelihood function, as noted in the body of the paper). So, we have $\boldsymbol{p}_\star(A) = \sum_{i=1}^n \beta_i \boldsymbol{p}_i^B(A)$, for appropriate β_i, again by Lemma 1. By hypothesis $\boldsymbol{p}_i^B(A) = \boldsymbol{p}_i(A)$ for $i = 1, ..., n$. Hence, $\boldsymbol{p}_\star(A) = \sum_{i=1}^n \beta_i \boldsymbol{p}_i^B(A) = \sum_{i=1}^n \beta_i \boldsymbol{p}_i(A)$. Letting $\alpha_i = \beta_i$, it follows that $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A) = \mathcal{F}^B(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)(A)$. □

Proof of Proposition 6

Proof. We show first that $\mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \subseteq conv(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ for all $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$. Since there are at least three disjoint events, $A_1, A_2, A_3 \in \mathscr{A}$, following (Lehrer and Wagner, 1981, Theorems

6.4, 6.7) and (McConway, 1981, Theorem 3.3), we can exploit techniques and results for functional equations. For any numbers $a_i, b_i \in [0,1]$ with $a_i + b_i \in [0,1]$, define a sequence of probability measures, $\boldsymbol{p}_i,\ i = 1, ..., n$ by setting

$$\begin{aligned} \boldsymbol{p}_i(A_1) &= a_i \\ \boldsymbol{p}_i(A_2) &= b_i \\ \boldsymbol{p}_i(A_3) &= 1 - a_i - b_i \end{aligned}$$

Since it is the case that $\mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) \in \mathbb{P}$ for all $(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \in \mathbb{P}^n$ and every $\mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) \in \mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, we have it that $\mathfrak{m}\big(\boldsymbol{p}_1(A), ..., \boldsymbol{p}_n(A)\big) = \boldsymbol{p}(A)$, for some $\boldsymbol{p} \in \mathbb{P}$ and all $A \in \mathscr{A}$. Now, by the additivity of probability measures, $\boldsymbol{p}(A_1 \cup A_2) = \boldsymbol{p}(A_1) + \boldsymbol{p}(A_2)$. Hence, $\mathfrak{m}(a_1 + b_1, ..., a_n + b_n) = \mathfrak{m}\big(a_1, ..., a_n\big) + \mathfrak{m}\big(b_1, ..., b_n\big)$. So, $\mathfrak{m}$ satisfies Cauchy's multivariable functional equation. For each $i = 1, ..., n$, define $\mathfrak{m}_i(a) = \mathfrak{m}(0, ..., a, ..., 0)$, where a occupies the i-th position of the vector $(0, ..., a, ..., 0)$. It is clear that $\mathfrak{m}_i(a + b) = \mathfrak{m}_i(a) + \mathfrak{m}_i(b)$ for all $a, b \in [0,1]$ with $a + b \in [0,1]$. Because $\mathfrak{m}$ is nonnegative, so is $\mathfrak{m}_i,\ i = 1, ..., n$. By Theorem 3 of (Aczél and Oser, 2006, p. 48), it follows that there exists a nonnegative constant α_i such that $\mathfrak{m}_i(a) = \alpha_i a$ for all $a \in [0,1]$. By the Cauchy equation we have

$$\begin{aligned} \mathfrak{m}(a_1, ..., a_n) &= \mathfrak{m}(a_1, 0, ..., 0) + \mathfrak{m}(0, a_2, ..., a_n) \\ &= \mathfrak{m}(a_1, 0, ..., 0) + \mathfrak{m}(0, a_2, 0, ..., 0) + ... + \mathfrak{m}(0, ..., 0, a_n) \end{aligned}$$

So we have $\mathfrak{m}(a_1, ..., a_n) = \mathfrak{m}_1(a_1) + ... + \mathfrak{m}_n(a_n) = \alpha_1 a_1 + ... + \alpha_n a_n$. And since $\mathfrak{m}(1, ..., 1) = 1$ (by consideration of the probability of Ω), it follows that $\sum_{i=1}^n \alpha_i = 1$. Thus, $\mathfrak{m}$ is a convex combination.

Now, we want to show that $conv(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) \subseteq \mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. Let $\boldsymbol{p}$ be an element of $conv(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. It is clear that there exists an $\mathfrak{m} \in \mathfrak{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ such that $\mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) = \boldsymbol{p}$. And since $\boldsymbol{p}$ is just a convex combination, there exist weights $\alpha_1, ..., \alpha_n \in [0,1]$ such that $\sum_{i=1}^n \alpha_i = 1$ and $\boldsymbol{p} = \sum_{i=1}^n \alpha_i \boldsymbol{p}_i$. But for any other profile $(\boldsymbol{q}_1, ..., \boldsymbol{q}_n) \in \mathbb{P}^n$, taking any convex combination yields a probability measure. In particular, $\sum_{i=1}^n \alpha_i \boldsymbol{q}_i \in \mathbb{P}$. It follows that $\mathfrak{m} \in \bigcap_{\vec{\boldsymbol{q}} \in \mathbb{P}^n} \mathfrak{M}_n(\vec{\boldsymbol{q}})$. So, $\boldsymbol{p} = \mathfrak{m}\big(\boldsymbol{p}_1(\cdot), ..., \boldsymbol{p}_n(\cdot)\big) \in \mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$, as desired.

The two inclusions above show that $\mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = conv(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. Hence, $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = conv(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$ is equivalent to $\mathcal{F}(\boldsymbol{p}_1, ..., \boldsymbol{p}_n) = \mathcal{M}_n(\boldsymbol{p}_1, ..., \boldsymbol{p}_n)$. □

References

Aczél, J. (1948). On mean values. *Bulletin of the American Mathematical Society 54*(4), 392–400.

Aczél, J. and H. Oser (2006). *Lectures on Functional Equations and Their Applications*. Courier Corporation.

Arló-Costa, H. and J. Helzner (2010). Ambiguity aversion: The explanatory power of indeterminate probabilities. *Synthese 172*(1), 37–55.

Bacharach, M. (1972). Scientific disagreement. *Unpublished Manuscript*.

Bordley, R. F. (1982). A multiplicative formula for aggregating probability assessments. *Management Science 28*(10), 1137–1148.

Bradley, S. (2014). Imprecise probabilities. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 ed.).

Clemen, R. T. and R. L. Winkler (1999). Combining probability distributions from experts in risk analysis. *Risk Analysis 19*(2), 187–203.

Cozman, F. (1998). Irrelevance and independence relations in quasi-bayesian networks. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, pp. 89–96. Morgan Kaufmann Publishers Inc.

Cozman, F. G. (2000). Credal networks. *Artificial Intelligence 120*(2), 199–233.

de Campos, L. M. and S. Moral (1995). Independence concepts for convex sets of probabilities. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 108–115. Morgan Kaufmann Publishers Inc.

de Finetti, B. (1931). Sul concetto di media. *Giornale dell'Istituto Italiano degli Attuari 2*, 369–396.

de Finetti, B. (1964). Foresight: Its logical laws, its subjective sources. In H. E. Kyburg and H. E. Smoklery (Eds.), *Studies in Subjective Probability*. Wiley.

Dietrich, F. (2010). Bayesian group belief. *Social Choice and Welfare 35*(4), 595–626.

Dietrich, F. and C. List (2014). Probabilistic opinion pooling. In A. Hájek and C. Hitchcock (Eds.), *Oxford Handbook of Probability and Philosophy*. Oxford University Press.

Ellsberg, D. (1963). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics 77*(2), 327–336.

French, S. (1985, September). Group consensus probability distributions: A critical survey. In D. L. J.M. Bernardo, M.H. DeGroot and A. Smith (Eds.), *Bayesian Statistics: Proceedings of the Second Valencia International Meeting*, Volume 2, pp. 183–201. North-Holland.

Gaifman, H. and M. Snir (1982). Probabilities over rich languages, testing and randomness. *The journal of symbolic logic 47*(03), 495–548.

Gärdenfors, P. and H. Rott (1995). *Handbook of Logic in Artificial Intelligence and Logic Programming: Epistemic and temporal reasoning*, Volume 4, Chapter Belief Revision. Oxford University Press, Oxford.

Gärdenfors, P. and N.-E. Sahlin (1982). Unreliable probabilities, risk taking, and decision making. *Synthese 53*(3), 361–386.

Genest, C. (1984). A characterization theorem for externally bayesian groups. *The Annals of Statistics*, 1100–1105.

Genest, C., K. J. McConway, and M. J. Schervish (1986). Characterization of externally bayesian pooling operators. *The Annals of Statistics*, 487–501.

Genest, C. and C. G. Wagner (1987). Further evidence against independence preservation in expert judgement synthesis. *Aequationes Mathematicae 32*(1), 74–86.

Genest, C. and J. V. Zidek (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 114–135.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Girón, F. J. and S. Ríos (1980). Quasi-bayesian behaviour: A more realistic approach to decision making? *Trabajos de Estadística y de Investigación Operativa 31*(1), 17–38.

Halmos, P. R. (1963). *Lectures on Boolean Algebras*. Van Nostrand, Princeton.

Herron, T., T. Seidenfeld, and L. Wasserman (1997). Divisive conditioning: Further results on dilation. *Philosophy of Science*, 411–444.

Herzberg, F. (2014). Aggregating infinitely many probability measures. *Theory and Decision*, 1–19.

Kaplan, M. (1996). *Decision Theory as Philosophy*. Cambridge University Press.

Kolmogorov, A. N. (1930). Sur la notion de la moyenne. *Atti della R. Academia Nazionale dei Lincei 12*(9), 388–391.

Kyburg, H. E. (1998). Interval-valued probabilities. *Imprecise Probabilities Project*.

Kyburg, H. E. and M. Pittarelli (1992). Some problems for convex bayesians. In *Proceedings of the Eighth international conference on Uncertainty in artificial intelligence*, pp. 149–154. Morgan Kaufmann Publishers Inc.

Kyburg, H. E. and M. Pittarelli (1996). Set-based bayesianism. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on 26*(3), 324–339.

Lehrer, K. and C. Wagner (1981). *Rational Consensus in Science and Society: A Philosophical and Mathematical Study*, Volume 21. Springer.

Lehrer, K. and C. Wagner (1983). Probability amalgamation and the independence issue: A reply to laddaga. *Synthese 55*(3), 339–346.

Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy 71*(13), 391–418.

Levi, I. (1978). Irrelevance. In *Foundations and Applications of Decision Theory*, pp. 263–273. Springer.

Levi, I. (1980). *The Enterprise of Knowledge*. MIT Press, Cambridge, MA.

Levi, I. (1985). Consensus as shared agreement and outcome of inquiry. *Synthese 62*(1), pp. 3–11.

Levi, I. (1986a). *Hard choices: Decision making under unresolved conflict*. Cambridge University Press.

Levi, I. (1986b). The paradoxes of allais and ellsberg. *Economics and Philosophy 2*(1), 23–53.

Levi, I. (1990). Pareto unanimity and consensus. *The Journal of Philosophy 87*(9), 481–492.

Levi, I. (2009). Why indeterminate probability is rational. *Journal of Applied Logic 7*(4), 364–376.

Madansky, A. (1964). Externally bayesian groups.

McConway, K. J. (1981). Marginalization and linear opinion pools. *Journal of the American Statistical Association 76*(374), 410–414.

Mongin, P. (1995). Consistent bayesian aggregation. *Journal of Economic Theory 66*(2), 313–351.

Moral, S. and J. Del Sagrado (1998). Aggregation of imprecise probabilities. In B. Bouchon-Meunier (Ed.), *Aggregation and Fusion of Imperfect Information*, pp. 162–188. Springer.

Nau, R. F. (2002). The aggregation of imprecise probabilities. *Journal of Statistical Planning and Inference 105*(1), 265–282.

Ouchi, F. (2004). A literature review on the use of expert opinion in probabilistic risk analysis.

Pedersen, A. P. and G. Wheeler (2014). Demystifying dilation. *Erkenntnis 79*(6), 1305–1342.

Pedersen, A. P. and G. Wheeler (2015). Dilation, disintegrations, and delayed decisions. In *Proceedings of the 9th*.

Ramsey, F. P. (1931). Truth and probability. *The Foundations of Mathematics and Other Logical Essays*, 156–198.

Rockafellar, R. T. (1970). *Convex Analysis*. Number 28. Princeton University Press.

Savage, L. (1972, originally published in 1954). *The Foundations of Statistics*. New York: John Wiley and Sons.

Seidenfeld, T. (1993). Outline of a theory of partially ordered preferences. *Philosophical Topics 21*(1), 173–189.

Seidenfeld, T., J. B. Kadane, and M. J. Schervish (1989). On the shared preferences of two bayesian decision makers. *The Journal of Philosophy 86*(5), 225–244.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2010). Coherent choice functions under uncertainty. *Synthese 172*(1), 157–176.

Seidenfeld, T. and L. Wasserman (1993). Dilation for sets of probabilities. *The Annals of Statistics 21*(3), 1139–1154.

Smith, C. A. B. (1961). Consistency in statistical inference and decision. *Journal of the Royal Statistical Society. Series B (Methodological) 23*(1), pp. 1–37.

Stewart, R. T. and I. Ojea Quintana (MS2). Learning and pooling, pooling and learning. *Unpublished Manuscript*.

Stone, M. (1961). The opinion pool. *The Annals of Mathematical Statistics 32*(4), 1339–1342.

Wagner, C. (2009). Jeffrey conditioning and external bayesianity. *Logic Journal of IGPL 18*(2), 336–345.

Walley, P. (1982). The elicitation and aggregation of beliefs. Technical Report 23, Department of Statistics, University of Warwick, Coventry CV4 7AL, England.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman and Hall London.

Wasserman, L. (1993). Review: Statistical reasoning with imprecise probabilities by peter walley. *Journal of the American Statistical Association 88*(422), pp. 700–702.

Wasserman, L. and T. Seidenfeld (1994). The dilation phenomenon in robust bayesian inference. *Journal of Statistical Planning and Inference 40*(2), 345–356.