Forthcoming in *The Review of Symbolic Logic*

Support for Geometric Pooling

Jean Baccelli and Rush T. Stewart*

MCMP, LMU Munich

October 9, 2020

Abstract

Supra-Bayesianism is the Bayesian response to learning the opinions of others. Proba-bility pooling constitutes an alternative response. One natural question is whether there are cases where probability pooling gives the supra-Bayesian result. This has been called the problem of Bayes-compatibility for pooling functions. It is known that in a com-mon prior setting, under standard assumptions, linear pooling cannot be non-trivially Bayes-compatible. We show by contrast that geometric pooling can be non-trivially Bayes-compatible. Indeed, we show that, under certain assumptions, geometric and Bayes-compatible pooling are equivalent. Granting supra-Bayesianism its usual norma-tive status, one upshot of our study is thus that, in a certain class of epistemic contexts, geometric pooling enjoys a normative advantage over linear pooling as a social learning mechanism. We discuss the philosophical ramifications of this advantage, which we show to be robust to variations in our statement of the Bayes-compatibility problem.

Keywords: supra-Bayesianism; linear pooling; geometric pooling; Bayes-compatibility; common prior; total evidence; deference; synergy

1 Introduction

Consider two or more Bayesian agents, endowed with a common prior. Assume that each agent privately receives some information about the true state of nature, updates her prior accordingly, and publicly announces her posterior beliefs. Now, contrast the following two responses. The first corresponds to so-called supra-Bayesianism (see Genest and Zidek, 1986,

Sec. 4, and the references therein). It consists in treating the announcements of the posterior probability values as further information on which the common prior is to be updated. The second consists simply in pooling—specifically, taking a weighted average of—the announced posterior probability values (for a review of the most important pooling methods, see, e.g., Genest and Zidek, 1986, Sec. 3). One interesting question is whether there are pooling methods such that the two responses lead to the same result, i.e., whether some pooling methods can deliver the supra-Bayesian response. This has been called the problem of Bayes-compatibility for pooling subjective probabilities, and it has been investigated even beyond the common prior setting (e.g., Genest and Schervish, 1985). The literature has so far essentially focused on investigating the problem under linear pooling. In the common prior setting, an impossibility result has been established, to the effect that linear pooling with strictly positive weights cannot be non-trivially Bayes-compatible (Dawid et al., 1995; Bradley, 2018 as qualified by Dawid and Mortera, 2020; also Steele, 2012). Under weighted geometric pooling, by contrast, the problem has not hitherto been systematically studied in this or any other context, to the best of our knowledge. Dawid et al. touch upon the issue, but they do not elaborate (Dawid et al., 1995, p. 280). The question addressed in Easwaran et al., 2016 is related but distinct, if only because the authors focus on an unweighted variant of geometric pooling.

Our paper aims at filling the above gap in the literature. We identify and study the form of Bayes-compatibility that is most relevant for geometric pooling in a common prior setting. Judging from the seemingly trivial informational circumstances under which it applies, the relevant condition, which we dub "full-revelation Bayes-compatibility," encapsulates a very minimal notion of Bayes-compatibility. But by the same token, it expresses an extremely basic requirement for any pooling function aspiring to the supra-Bayesian ideal. In a nutshell, for any such pooling function, our full-revelation Bayes-compatibility requirement should be a low bar to clear. Equipped with our new notion of Bayes-compatibility, we depart from the existing literature—that has essentially focused on the aforementioned impossibility obtaining under linear pooling—in establishing the possibility result that geometric pooling can, in a common prior setting, be non-trivially Bayes-compatible. The issue of the support of the measures being pooled, traditionally considered a weakness of geometric pooling, plays a prominent role in this possibility; hence our paper's title. In fact, our main finding is the stronger result that geometric pooling is the only pooling method displaying the form of Bayes-compatibility which we introduce. Setting aside the trivial cases where linear and geometric pooling coincide, linear pooling systematically violates the very basic Bayes-compatibility requirement which our notion expresses; i.e., linear pooling fails to clear the low bar for Bayes-compatibility set by this requirement. Positively speaking, our main result delivers at once a characterization of a very fundamental notion of Bayes-compatibility and, under some restrictions (including the domain restriction induced by the common prior assumption), a characterization of geometric pooling. Given that supra-Bayesianism has the status of a normative ideal, our result thus indicates one dimension along which geometric pooling enjoys a normative advantage over linear pooling. We discuss the philosophical ramifications of this advantage. Perhaps most important in this respect, we explain how it relates to respect for a pillar of Bayesian epistemology: the principle of total evidence (Carnap, 1947; Good, 1967). While this completes the conceptual core of our contribution, for interested readers, we also examine whether this advantage carries over to more gen-

eral settings than the one in which we initially investigate the Bayes-compatibility problem. Specifically, we examine the effects of relaxing the common prior assumption. We also examine the effects of relaxing the assumption which we will initially impose on the amount of information shared by the agents. We find that even when such generalizations have an impact on the Bayes-compatibility of geometric pooling or the Bayes-incompatibility of linear pooling, they do not eliminate the comparative advantage which the former enjoys over the latter.

Zooming out from the set of papers specifically dedicated to analyzing Bayes-compatibility, our study is situated, more generally, at the intersection of several strands of literature. It belongs, first, to the literature on supra-Bayesianism (see Morris, 1974 and the other references in Genest and Zidek, 1986, Sec. 4). On this score, we present a restrictive set of conditions under which the consequences of the supra-Bayesian ideal are unusually tractable, given that they coincide with those of geometric pooling. Second, our paper belongs to the common prior literature (Aumann, 1976; Geanakoplos and Polemarchakis, 1982). Indeed, our main result consists in deriving a logical implication of the classical common prior assumption. Most of the other assumptions used in this derivation are standard in the common prior literature. Our paper contributes, third, to the axiomatic analysis of pooling functions (surveyed in Genest and Zidek, 1986, Sec. 3; more recently Dietrich and List, 2016). We introduce a simple Bayes-compatibility property, which we characterize by two preexisting axioms and use to differentiate geometric from linear pooling over suitably restricted domains. Fourth, our paper belongs to a large stream of formal social epistemology using pooling functions to answer a variety of questions concerning consensus formation, social learning, and related topics (see DeGroot, 1974; Lehrer and Wagner, 1981; and a review of the recent literature on social networks in Golub and Sadler, 2016). One difference is worth mentioning upfront, however. In this literature, opinion pooling is typically introduced as an informationally parsimonious, boundedly rational alternative to supra-Bayesianism—a view which, in the case of linear pooling, has been given some axiomatic foundations (see Golub and Sadler, 2016, Sec. 3.1.4). We mention this justification but avoid using it to motivate our own study. This is because, as we will observe, it would apply to geometric pooling with more qualifications than to linear pooling. One of our main conclusions will be precisely that, in our setting, there is a general trade-off between the informational parsimony and the Bayes-compatibility of a pooling function. Fifth and finally, our study touches upon some established concepts in social epistemology, such as testimony (Bradley, 2007; Steele, 2012), deference (Gaifman, 1988; Elga, 2007), and synergy (Christensen, 2009; Easwaran et al., 2016)—three ideas which we define and discuss below. Our paper as a whole is, in effect, a study in peer disagreement (Goldman, 2001), though the scope of disagreement in our case is obviously limited by the common prior restriction and cognate structural assumptions.

The rest of the paper is organized as follows. Section 2 gathers the necessary preliminaries on pooling functions. Section 3 introduces the Bayes-compatibility problem as it arises under linear pooling. Section 4 states and studies the problem as it arises under geometric pooling. This is where the central result of the paper is to be found. The next two sections are dedicated to better understanding that result. To this end, Section 5 emphasizes possible philosophical interpretations. Section 6, which may be skipped by those wishing to focus on our core conceptual take-aways, emphasizes possible mathematical generalizations. Section 7 concludes.

2 Preliminaries on Pooling Functions

Let Ω be a set of states of nature. For simplicity, we assume throughout the paper that Ω is finite and we take the maximal algebra of events $\mathcal{A} = 2^{\Omega}$. Let p denote a probability measure on $\mathcal{A}$, and define the the support of p as $supp(p) = \{\omega \in \Omega : p(\omega) > 0\}$. Let Δ be the set of all probability measures on $\mathcal{A}$. For any $n \in \mathbb{N}$, Δ^n is the set of all profiles $(p_1, ..., p_n)$ of length n, with $N = \{1, ..., n\}$ the set of profile members, and $n \geq 2$ for non-triviality. Of particular interest to this study will be the set $\Delta^{n'}$ of all profiles of n probability measures with intersecting supports, i.e., $\Delta^{n'} = \{(p_1, ..., p_n) \in \Delta^n : \bigcap_{i \in N} supp(p_i) \neq \emptyset\}$. The profiles displaying this property have been called "coherent" in the literature (Dietrich, 2017, p. 3).

A pooling function $F : \Delta^n \to \Delta$ associates a probability measure $F(p_1, ..., p_n)$ with a profile $(p_1, ..., p_n)$. The two most prominent pooling functions in the literature are the weighted linear and the weighted geometric pooling functions. Assume for now that both pooling functions apply with a set of weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \in (0, 1)$ for all i in N and $\sum_{i=1}^{n} \alpha_i = 1$.[1] The first property is natural, for instance, whenever the weights reflect respective trust in an epistemic community and none of the profile members has reasons to entirely disregard any other's opinions. The second amounts to a kind of normalization. Then, for any $(p_1, ..., p_n) \in \Delta^n$ and any $A \in \mathcal{A}$, the weighted linear pool of $(p_1, ..., p_n)$ is given by

$$F(p_1, ..., p_n)(A) = \sum_{i=1}^{n} \alpha_i p_i(A), \tag{1}$$

or equivalently by

$$F(p_1, ..., p_n)(A) = \sum_{\omega \in A} \sum_{i=1}^{n} \alpha_i p_i(\omega).$$

By contrast, for any $(p_1, ..., p_n) \in \Delta^{n'}$, the weighted geometric pool of $(p_1, ..., p_n)$ is given by

$$F(p_1, ..., p_n)(A) = c \sum_{\omega \in A} \prod_{i=1}^{n} p_i(\omega)^{\alpha_i}, \tag{2}$$

with c the normalizing constant given by

$$c = \frac{1}{\sum_{\omega' \in \Omega} \prod_{i=1}^{n} p_i(\omega')^{\alpha_i}}. \tag{3}$$

We highlight that in the case of geometric pooling, a domain restriction (to $\Delta^{n'}$) is necessary to ensure that pooling results in a probability measure. This is due to the following property of geometric pooling: If one profile member gives probability value zero to a state, then the geometric pool gives it probability zero as well. This is typically considered a major defect of geometric pooling, inasmuch as "zeros [...] constitute vetos" (Genest and Zidek, 1986, p. 120;

---

[1]This assumption is to initially facilitate the comparison of linear and geometric pooling. However, neither is restricted to this case. Indeed, we will later consider more general cases for both. Furthermore, in the case of geometric pooling, we will provide more primitive assumptions about pooling functions under which, over some domains, the above restrictions on the weights can be imposed without loss of generality.

see also French, 1985, p. 186). In what follows, we revisit the negative evaluation of this property of geometric pooling, which we call *the support-veto property*. At this stage, we simply observe that the entailed domain restriction is natural in a wide range of epistemic contexts.[2] In particular, it is satisfied whenever all profile members give strictly positive probability to the true state of nature. At a later stage, we will show that there are cases where the support-veto property of geometric pooling can be exploited further for epistemic gain.

Both linear and geometric pooling have prima facie desirable properties that, under additional assumptions, can be used to characterize them. Unlike geometric pooling, linear pooling has the so-called strong setwise function property (McConway, 1981; also Mongin, 1995, Sec. 2; Bonnay and Cozic, 2018, Sec. 5). This is essentially the requirement that the pooling function be separable event by event in the algebra $\mathcal{A}$, i.e., that the value of the pool on $A \in \mathcal{A}$ depend on the values taken by the p_i only on A, and on no other event in $\mathcal{A}$. On the other hand, unlike linear pooling, geometric pooling has the so-called external Bayesianity property (Genest et al., 1986; also Russell et al., 2015; Dietrich, 2017). This is essentially the requirement that pooling updated probability measures leads to the same result as updating the pool of the probability measures. Each of these properties amounts to a commutativity requirement on, and can be interpreted as a form of non-manipulability of, the pooling function (Dietrich and List, 2016). In what follows, we compare linear and geometric pooling from a related but distinct point of view, namely, their respective merits as mechanisms for *learning from others*. To this end, we assume the prevailing normative view that, in the contexts where it applies, Bayesianism is the gold standard for learning evidence in general, and "supra-Bayesianism", to which we now turn, is the gold standard for learning from others in particular.

3 The Problem of Bayes-Compatibility

3.1 A Baseline Setting for Bayes-Compatibility

Supra-Bayesianism (e.g., Morris, 1974) is nothing other than what Bayesianism demands in a social context, i.e., when the available evidence includes the beliefs of other agents. More specifically, by definition, a supra-Bayesian agent has a prior not only about the state of nature, but also about any probabilistic opinion the other agents in her epistemic community might have about the state of nature. Additionally, when she learns these opinions, for example through verbal reports, she updates her prior by Bayesian conditionalization, just like she does when she learns any event. Such reports simply provide a special kind of evidence, sometimes called *testimonial* evidence (Bradley, 2007; Steele, 2012), the receipt of which her posterior should reflect. To illustrate this idea, let us assume that the supra-Bayesian agent is interested in the probability of some event $A \in \mathcal{A}$. Suppose that there is an event $E \in \mathcal{A}$ that represents the reports of the other agents' probabilistic opinions on A. (This is in fact an important assumption, and we will shortly offer one way to substantiate it.) Under this assumption, with p her prior and provided that $p(E) > 0$,[3] the agent responds

---

[2]We later derive this domain restriction from more primitive assumptions about information; see especially fn. 7.

[3]Throughout the paper, whenever we write expressions like "$p(A|E)$" or refer to "the conditionalization of p on information E", we assume that $p(E) > 0$ so that $p(A|E)$ is well-defined.

in a supra-Bayesian way to learning E just in case she adopts the posterior $p(A|E)$ given by

$$p(A|E) = \frac{p(A \cap E)}{p(E)}. \tag{4}$$

More specifically, we will follow the stream of literature (Dawid et al., 1995; Bradley, 2018) examining the links between supra-Bayesianism and opinion pooling in the simple analytical setting characterized by the following assumptions.[4] First, all the agents start, and know that they start, from a common prior over $\mathcal{A}$. For simplicity, we will assume throughout that this prior has full support over Ω.[5] Second, each agent i receives about the true state of nature ω some private information $E_i(\omega)$. This private information is "representable" (Dietrich, 2017, p. 2), i.e., for all i, $E_i(\omega) \in \mathcal{A}$. This excludes more general forms of learning, such as the ones considered in the theory of Jeffrey conditioning (Jeffrey, 2004). Additionally, the private information is what we call "factive", i.e., whatever the true state of nature ω, for all i, $\omega \in E_i(\omega)$. Intuitively speaking, this excludes receiving false information about the true state of nature. Third, each agent i updates the common prior by Bayesian conditioning on her private information $E_i(\omega)$ and publicly announces her posterior probability value for some event A of interest. The n updated probability measures will constitute the profile of interest from the point of view of pooling. To these three established assumptions we add a fourth one that is likewise established in the common prior literature (Aumann, 1976; Geanakoplos and Polemarchakis, 1982) and comparable to assumptions standardly made in the literature on supra-Bayesianism and opinion pooling.[6] To wit, we will assume that each agent i is equipped with an information partition $\mathcal{E}_i$ over Ω, that any piece of private information $E_i(\omega)$ she might receive is taken from that partition $\mathcal{E}_i$, and that the information partitions $\{\mathcal{E}_i\}_{i \in N}$ are known by all the agents. This allows us to work only with the primitive space of states of nature Ω, instead of having to introduce (like in, e.g., Romeijn and Roy, 2018) an additional space of epistemic states corresponding to the possible probability announcements.

These assumptions call for some preliminary comments. The second assumption, about the nature of information, will be maintained throughout the paper. We want to focus here on the simplest kinds of learning, and these involve only representable and factive pieces of information. From now on, "information" will always be shorthand for "representable and factive information." As regards the third assumption, about social information, the focus on absolute probability values—i.e., single probability values of the form $p(A)$—will be maintained through most of the paper. But we will also briefly consider relative probability values—i.e., probability ratios, of the form $p(A)/p(B)$. The first assumption, i.e., the common prior assumption, is evidently restrictive. We will eventually explore weakening it, but

---

[4]Not all the literature on supra-Bayesianism and opinion pooling adopts this analytical setting, however, especially not the common prior condition, nor even the more general assumption that some prior is fully specified from the outset (and thus, that Bayes-compatibility is eventually to be appreciated with respect to that specific prior). See Genest and Schervish, 1985, for example, with a recent philosophical application in Romeijn, 2019a; see also Bonnay and Cozic, 2019.

[5]Lindley proposed to call "Cromwell's rule" the "requirement (...) that for quantities taking only a finite set of values, no probability be zero". This is because "[Cromwell] suggested its equivalent when he advised the Church of Scotland to remember that it might be wrong" (Lindley, 1982a, Sec. 6).

[6]The literature standardly assumes that from the point of view of a distinguished agent i, the report of each j constitutes a random variable for i (Dawid et al., 1995, p. 267-268; Bradley, 2018, p. 9-10). This is tantamount to i equipping each j with an information partition. More on the random variable construction later in the present section.

we will impose it through most of the paper, and will weaken it only mildly. This calls for two immediate clarifications. First, although such claims have been made in the literature (see Morris, 1995 for a review), we make no general normative claim for the common prior assumption. We simply invoke it for a baseline investigation of the links between supra-Bayesianism and opinion pooling. Second, we observe that from a subjectivist perspective at least, this is a rather natural assumption to make in the context of opinion pooling. Indeed, starting from a common prior, differences in beliefs boil down to differences in information (since we assume that only representable and factive pieces of information are received). And from a subjectivist perspective, as Aumann argues, trying to "reconcil[e] subjective probabilities [by pooling them] makes sense if it is a question of implicitly exchanging information, but not if we are talking about 'innate' differences in priors" (Aumann, 1976, p. 1238). Finally, the fourth assumption, about information partitions, is also evidently restrictive. We highlight that in a common prior context, it can be interpreted (as elsewhere in the literature; see Bonanno and Nehring, 1997) in terms of a model in which scientists start from a common hypothesis (the common prior), distribute various experiments between themselves, and then share the results. Such an epistemic community constitutes a natural setting for thinking about learning from others. It can be used for making sense of our main results.

To formally state the problem of Bayes-compatibility, we need to make one last mathematical object explicit. Recall our supposition that, for each $i \in N$, there is an information partition $\mathcal{E}_i$, and that these partitions are known to all agents. Accordingly, let $p(\cdot|E_i(\cdot)) : \mathcal{A} \times \Omega \to \mathbb{R}$ be a random variable such that, for all $\omega \in \Omega$, $p(\cdot|E_i(\omega)) \in \Delta$ and $E_i(\omega)$ is the cell of $\mathcal{E}_i$ that contains ω.[7] In words, $p(\cdot|E_i(\cdot))$ is the random variable whose possible values are the possible posterior opinions of agent i. Since we are assuming that the common prior p has full support, each cell of each partition has positive probability, so that $p(\cdot|E_i(\cdot))$ assumes distinct values on each cell of $\mathcal{E}_i$. Exploiting the full support and factive learning assumptions, we can identify the information that i receives with the support of the posterior of i: for all $\omega \in \Omega$, $\{\omega' \in \Omega : p(\cdot|E_i(\omega')) = p(\cdot|E_i(\omega))\} = E_i(\omega) = supp(p(\cdot|E_i(\omega)))$.[8] Importantly, it then follows that, for all $\omega \in \Omega$, $\bigcap_{i \in N}\{\omega' \in \Omega : p(\cdot|E_i(\omega')) = p(\cdot|E_i(\omega))\} = \bigcap_{i \in N} supp(p(\cdot|E_i(\omega)))$, which is non-empty by the assumption that agents learn only events containing ω. Therefore, in our setting, learning the n distributions is equivalent to learning the intersection of the supports of the probability distributions. This fact, which will play a central role in our paper, can be immediately verified as follows. Since, for all $\omega \in \Omega$, $supp(p(\cdot|E_i(\omega))) = E_i(\omega) \in \mathcal{E}_i$, we can recover the support of $p(\cdot|E_i(\omega))$ from knowledge of both $\bigcap_{i \in N} supp(p(\cdot|E_i(\omega)))$ and $\mathcal{E}_i$. To do so, it suffices to find the unique cell of $\mathcal{E}_i$ that has non-empty intersection with $\bigcap_{i \in N} supp(p(\cdot|E_i(\omega)))$. For simplicity, in what follows, unless we say otherwise (as will be necessary, for example, when we will consider relaxing the full support assumption), we will let $E_i(\omega)$ denote both the conditioning information and the support of $p(\cdot|E_i(\omega))$. Since those events are identical under the above assumptions, no confusion can arise.

Studies of Bayes-compatibility have hitherto focused on cases where the agents report not their full posteriors, but only their posterior probability value for some distinguished event

---

[7]Notice that a profile $(p(\cdot|E_1(\omega)), \ldots, p(\cdot|E_n(\omega)))$ generated by Bayesian updating p on n pieces of private evidence taken from n information partitions is in $\Delta^{n'}$, for with ω the unknown true state of nature, $\omega \in \bigcap_{i \in N} supp(p(\cdot|E_i(\omega))) = \bigcap_{i \in N} E_i(\omega)$, the last equality following from the full support assumption.

[8]Without the full support assumption, the support of p_i could be a strict subset of $E_i(\omega)$.

$A \in \mathcal{A}$. One possible motivation for this is that the agents have only limited information about each other's beliefs. Another can be that the epistemic inquiry is conducted event by event, perhaps so as to take full advantage of event-specific expertise of the agents, perhaps for some other reason. Either way, such cases can be modeled using *real-valued* random variables of the form $p(A|E_i(\cdot)) : \Omega \to [0,1]$. In general, it is possible that $p(A|E) = p(A|E')$ for $E, E' \in \mathcal{E}_i$, $E \neq E'$. Consequently, the report of $p(A|E_i(\omega))$ does not generally inform the other agents of exactly one cell of $\mathcal{E}_i$. The partition corresponding to $p(A|E_i(\cdot))$ will be a coarsening of the information partition $\mathcal{E}_i$, with cells formed by taking unions of all the cells in $\mathcal{E}_i$ that yield the same probability value for A. This is how the form of testimonial evidence most often discussed in the literature can be represented in our setting.

We can now formally state the baseline, standard version of the Bayes-compatibility problem. Let p be the common prior. With A some event of interest in $\mathcal{A}$, let $\{p(A|E_i(\cdot))\}_{i \in N}$ be real-valued random variables corresponding to the updated probability value agent $i = 1, \ldots, n$ might report for event A. The report of agent i, then, is represented as $p(A|E_i(\cdot)) = p(A|E_i(\omega))$. Studies of Bayes-compatibility typically adopt the perspective of some distinguished agent $i \in N$. Call a pooling function F *Bayes-compatible with respect to agent i* if, whatever the true state of nature $\omega \in \Omega$, for any event $A \in \mathcal{A}$, the following equality holds:[9]

$$F\left(p|E_1(\omega), \ldots, p|E_n(\omega)\right)(A) = p\left(A \middle| p(A|E_1(\cdot)) = p(A|E_1(\omega)), \ldots, E_i(\omega), \ldots, p(A|E_n(\cdot)) = p(A|E_n(\omega))\right). \tag{5}$$

It is not always relevant to single out some particular agent $i \in N$ to think about Bayes-compatibility, however. Bayes-compatibility can also be investigated from the point of view of all agents at once, in the following sense. Call a pooling function F *Bayes-compatible* if, whatever the true state of nature $\omega \in \Omega$, for any event $A \in \mathcal{A}$, the following equality holds:

$$F\left(p|E_1(\omega), \ldots, p|E_n(\omega)\right)(A) = p\left(A \middle| p(A|E_1(\cdot)) = p(A|E_1(\omega)), \ldots, p(A|E_n(\cdot)) = p(A|E_n(\omega))\right). \tag{6}$$

Equation (6) is still naturally interpreted from the perspective of any specific $i \in N$, provided one considers i before she receives information $E_i(\omega)$ about the true state of nature ω. Indeed, ex ante in this sense, i is uncertain about her own posterior just like she is uncertain about any other j's posterior. Accordingly, to appreciate Bayes-compatibility from such ex ante perspective, it is not relevant to single out i, or j, or any specific agent. Alternatively, equation (6) can be interpreted from the point of view of some external observer that—without having any private information of her own about the true state of nature—would be uncertain about the opinions of a group of agents. We leave the interpretation open at this

---

[9]Notice that our definition requires equality to hold surely (deterministically), and not just almost surely (in expectation). By contrast, most of the literature focuses on almost sure equality. Sure equality implies almost sure equality, so our requirement is stronger. Furthermore, sure equality can be considered more natural than almost sure equality when, like in our setting, finite state spaces are at stake. By way of anticipation, the equalities in our main finding (Theorem 1) and most subsequent results—even when the full support assumption is relaxed (e.g., Proposition 4)—will also hold surely, not just almost surely. More on the logical relationships between our approach and the traditional approach in fn. 18.

stage. For reasons to be discussed in more detail below (see especially the discussion after (11)), we will henceforth focus on (6) rather than (5). The problem of Bayes-compatibility that we are interested in, then, is whether any pooling function can be Bayes-compatible in the sense displayed in (6).[10]

3.2 The Possibility of Bayes-Compatibility

As defined in the previous section, the problem of Bayes-compatibility has already been studied in the particular case of linear pooling functions. For linear pooling, the equality in (6) reads as follows:

$$\sum_{i=1}^{n} \alpha_i p(A|E_i(\omega)) = p\left(A \middle| p(A|E_1(\cdot)) = p(A|E_1(\omega)), \ldots, p(A|E_n(\cdot)) = p(A|E_n(\omega))\right). \quad (7)$$

It has been established in the literature that, for any A, the equality in (7) holds almost surely only if, for all $i, j \in N$, $p(A|E_i(\cdot)) = p(A|E_j(\cdot))$ holds almost surely (Dawid et al., 1995, Sec. 4.1; Bradley, 2018, Sec. 4.2;[11] see also Steele, 2012). This can be interpreted as an impossibility result—into which we later offer a new, simple insight. Indeed, this means that under linear pooling, Bayes-compatibility holds only if all agents are expected to report the same updated probability value for the event A of interest. Under the assumption that these probability values come from updating the common prior on typically different pieces of private information, this is impossible.

To fully appreciate the impossibility, however, one would need to know whether it affects only linear or all pooling functions. The following preliminary example (parallel to the one in Bradley, 2018, Sec. 4.2) suggests that geometric pooling affords more possibilities than linear pooling. To understand the example, notice the following elementary implication of the common prior and the private information assumptions. Assume that agent 1, say, reports an updated probability value of zero for the event A of interest.[12] Then, upon learning this report, no matter what the other reports are, a supra-Bayesian $i \in N$ will set her posterior probability value for that event to zero. That is to say, $p\left(A|p(A|E_1(\cdot)) = 0, p(A|E_2(\cdot)) = p(A|E_2(\omega)), \ldots, p(A|E_n(\cdot)) = p(A|E_n(\omega))\right) = 0$. But, for any ω, this particular constraint cannot be satisfied under linear pooling unless $p(A|E_i(\omega)) = 0$ for all i. Under geometric pooling, by contrast, this constraint will always be satisfied. This follows from the support-veto property of geometric pooling, which we highlighted when introducing (2). Indeed, for any $\omega \in \Omega$, the following equality holds:

---

[10]We are aware of the fact that supra-Bayesianism itself could be formally construed as a pooling function (e.g., Genest and Zidek, 1986, p. 120). But we are interested in non-trivial Bayes-compatibility, so we consider only pooling functions that differ from the supra-Bayesian one in pooling-functional form.

[11]Dawid and Mortera (Dawid and Mortera, 2020) identify a problem in the proof of one of Bradley's theorems. But they grant—and indeed they prove in their own work (Dawid et al., 1995)—that linear pooling with strictly positive weights cannot be non-trivially Bayes-compatible. This corresponds to the above statement.

[12]The trivial case $E_i(\omega) = \Omega$ put aside, under our assumptions about information, each $i \in N$ will have some non-empty event—namely, $E_i(\omega)^c$—for which to report an updated probability value of zero.

$$c \sum_{\omega' \in A} \prod_{i=1}^{n} p(\omega'|E_i(\omega))^{\alpha_i} = 0 = p\left(A \middle| p(A|E_1(\cdot)) = 0, p(A|E_2(\cdot)) = p(A|E_2(\omega)), \ldots, p(A|E_n(\cdot)) = p(A|E_n(\omega))\right).$$

$$(8)$$

This raises several questions, among which are the following. Can geometric pooling achieve Bayes-compatibility also when none of the agents attributes posterior probability value zero to the event of interest? And when geometric pooling achieves Bayes-compatibility, what is the role played by the support-veto property?

4   Full-Revelation Bayes-Compatibility

4.1   Introducing Full-Revelation Bayes-Compatibility

Under the preceding definitions and assumptions, geometric pooling is Bayes-compatible if, whatever the true state of nature $\omega \in \Omega$, for any event $A \in \mathcal{A}$, the following equality holds:

$$c \sum_{\omega' \in A} \prod_{i=1}^{n} p(\omega'|E_i(\omega))^{\alpha_i} = p\left(A \middle| p(A|E_1(\cdot)) = p(A|E_1(\omega)), \ldots, p(A|E_n(\cdot)) = p(A|E_n(\omega))\right).$$

$$(9)$$

Notice first that, as indicated by the denominator of the normalizing factor c defined in (3),[13] the geometric variant of the Bayes-compatibility problem calls for stronger assumptions about social information than the linear variant. This is because, unlike linear pooling, geometric pooling is not separable event by event in the algebra $\mathcal{A}$. The value of the geometric pool on A depends on the values taken by the $p(\cdot|E_i(\cdot))$ not only on A, but also on events other than A. Thus, under geometric pooling, unlike under linear pooling, it makes little—if any—sense to maintain the assumption that the agents report only their updated probability value for the event A of interest. Put differently, the very definition of Bayes-compatibility in (6) proves tailored to, if not biased towards, linear pooling. For investigating the Bayes-compatibility of geometric pooling, then, an alternative approach must be taken. Now, the definition of geometric pooling in (2) makes it clear that to compute the value of the geometric pool for A, one needs to know no less than the full posterior distributions of the agents. Accordingly, in what follows, we require that the full posterior distributions of the agents be reported. We call this requirement *the full revelation assumption*. Admittedly, this is a very strong informational requirement. This is why we will later present a variation on the problem of Bayes-compatibility in which it can be relaxed. However, under geometric pooling, this requirement is necessary in order to study the problem of Bayes-compatibility as initially stated, i.e., with reference to a given event and a fixed underlying measurable

---

[13]Since the profile considered in (9) depends on ω, so does the normalizing factor c defined in (3).

space.[14] Furthermore, the minimal kind of Bayes-compatibility to which the full revelation assumption leads is worth investigating not only under geometric pooling, but under any type of pooling. In particular, as the next section will confirm, even linear pooling is well worth considering anew under the full revelation assumption. Indeed, as much as one should recognize that, under full revelation, Bayes-compatibility looks like a low hurdle, one cannot simply presume, but must check, that this low hurdle is actually cleared by any pooling mechanism, starting with the linear one.

As a preliminary to such an inquiry, notice that the full revelation assumption allows us to state a new version of the equality in (9). Let $\{p(\cdot|E_i(\cdot)\}_{i \in N}$ be random variables corresponding to the possible updated distributions of experts $i = 1, \ldots, n$. Then, indicating a particular full-revelation report by $p|E_i(\cdot) = p|E_i(\omega)$, the new equality reads as follows:

$$c \sum_{\omega' \in A} \prod_{i=1}^{n} p(\omega'|E_i(\omega))^{\alpha_i} = p\left(A \middle| p|E_1(\cdot) = p|E_1(\omega), \ldots, p|E_n(\cdot) = p|E_n(\omega)\right). \quad (10)$$

In our setting, by announcing her full posterior, each agent i effectively reveals her conditioning private information $E_i(\omega)$. Under our assumptions about information, this statement holds without qualification if, as we are assuming now and will be assuming through most of this paper, the prior p has full support over Ω. We will later invoke the fact that, even when this assumption is not imposed, a qualified version of the statement remains true.[15] Accordingly, we propose to model i's announcement that her posterior distribution is $p|E_i(\omega)$ simply as the event $E_i(\omega)$. Admittedly, in general, the two events are fundamentally distinct. However, under the assumption that the $E_i(\omega)$ are taken from information partitions $\mathcal{E}_i$ that are known to all, as we explained in Section 3.1, the two events are identical.[16]

The foregoing motivates introducing and investigating what we call *full-revelation* Bayes-compatibility. Call a pooling function F full-revelation Bayes-compatible if, whatever the true state of nature $\omega \in \Omega$, for any event $A \in \mathcal{A}$, the following equality holds:

---

[14]Rather than requiring agents to report their full posterior distributions, one could envisage requiring that the agents report their updated probability values only for A and A^c. This is essentially the strategy followed in Dawid et al., 1995, Sec. 4.3 and Easwaran et al., 2016, Sec. 9.1. First, notice that this would amount to coarsening the algebra, while the original pooling problem was stated for the algebra $\mathcal{A}$. Second, given the properties of geometric pooling, the only way to secure probabilistically consistent results across all such coarsenings would consist in requiring that the reported posterior values for both A and A^c be decomposed with respect to the finest of all the partitions of Ω—which would lead back to the full revelation assumption.

[15]Even when the full support assumption is relaxed, under our assumption that $\mathcal{A}$ contains all the states in Ω, one can still think of the agents as revealing their private information in the following sense. Assume for instance that ω_0 is the only state not in the support of p. Then, for any $\omega \in \Omega$, upon learning the full posterior of j, i cannot tell whether j's private information was $E_j(\omega)$ or $E_j(\omega) \cup \{\omega_0\}$. Such underdetermination generalizes to when there is more than one state in the complement of the support of p. In this sense, one may say that the agents reveal their private information up to zero-probability events.

[16]Furthermore, even when one cannot identify the announced posteriors with the received private information, $E_i(\omega)$ still is the most specific event representable in $\mathcal{A}$ which $j \neq i$ learns when receiving the probabilistic report that $p|E_i(\cdot) = p|E_i(\omega)$. This implies the following. Assume that (unlike in, e.g., Romeijn and Roy, 2018) one chooses *not* to explicitly model an additional algebra $\mathcal{B}$ for the probabilistic reports over $\mathcal{A}$. This permits zooming in on the simplest implications of supra-Bayesianism, i.e., those reflected within the basic algebra $\mathcal{A}$. Then, the only $\mathcal{A}$-representable option is to model the probabilistic report $p|E_i(\cdot) = p|E_i(\omega)$ as $E_i(\omega)$.

$$F\left(p|E_1(\omega),\ldots,p|E_n(\omega)\right)(A) = p\left(A\middle|\bigcap_{i\in N} E_i(\omega)\right). \tag{11}$$

To start with, the equality in (11) can be compared with the equality in (6). The former equality is evidently simpler than the latter. Furthermore, under the assumptions leading to the simplification of (6) into (11), we need not distinguish between (5) and (6). That is to say, it is unnecessary to define full-revelation Bayes-compatibility with respect to some agent i, rather than simpliciter. This is because, under the full-revelation assumption, a pooling function is Bayes-compatible with respect to some $i \in N$ if and only if it is Bayes-compatible with respect to all $i \in N$. We started directly by considering (6), rather than (5), largely because of this equivalence which we can now state.

Second, the very definition of full-revelation Bayes-compatibility can be conveniently simplified. Under our assumptions of a common prior and individual information partitions, any $\omega \in \Omega$ defines a profile $(p|E_1(\omega),\ldots,p|E_n(\omega))$. Accordingly, let Δ_{CP}^n stand for the set of all profiles generated from some common prior in accordance with all our assumptions.[17] Therefore, for any $A \in \mathcal{A}$, the claim that the equality in (11) holds for any n-tuple $\{p(\cdot|E_i(\cdot))\}_{i\in N}$ and all $\omega \in \Omega$ is equivalent to the claim that, for all profiles $(p|E_1,\ldots,p|E_n)$ in Δ_{CP}^n,

$$F\left(p|E_1,\ldots,p|E_n\right)(A) = p\left(A\middle|\bigcap_{i\in N} E_i\right). \tag{12}$$

Henceforth in this paper, we will adopt the alternative approach of quantifying over profiles, like in (12) and unlike in (11), that features realizations of random variables. While being equivalent to the traditional random variable approach when construed as above,[18] this alternative domain approach allows one to substantially simplify notation. It also brings out the continuity, which our results will exploit, between our question and some of the concepts and tools in the axiomatic theory of pooling functions.

The problem of full-revelation Bayes-compatibility, then, is whether any pooling function can be full-revelation Bayes-compatible as expressed in (12). Before thoroughly studying that problem, it is worth pausing to emphasize what it is and what it is not. The problem is *not* what the supra-Bayesian response would be under the full revelation assumption. The answer to this question is almost transparent. It is already given in the right-hand side of

---

[17]In more detail, the domain Δ_{CP}^n can be defined as follows. Any full-support common prior p and family of information partitions $\{\mathcal{E}_i\}_{i\in N}$ define an n-tuple of random measures $(p|E_i(\cdot),\ldots,p|E_n(\cdot))$ in accordance with our factive learning assumption. Let Λ index the set of all such n-tuples of random measures, $\{(p|E_1(\cdot),\ldots,p|E_n(\cdot))_\lambda : \lambda \in \Lambda\}$. For each $\lambda \in \Lambda$ and $\omega \in \Omega$, $\left(p|E_1(\cdot),\ldots,p|E_n(\cdot)\right)_\lambda$ induces a profile $\left(p|E_1(\omega),\ldots,p|E_n(\omega)\right)_\lambda$. Thus, the set of all profiles that are generated by such an n-tuple of random measures is $\Delta_{\mathrm{CP}}^n = \left\{\left(p|E_1(\omega),\ldots,p|E_n(\omega)\right)_\lambda : \lambda \in \Lambda \text{ and } \omega \in \Omega\right\}$.

[18]More fundamentally, our alternative approach indicates that the traditional conception of Bayes-compatibility as a property that holds for each profile in a set of probability 1 can be *generalized* to the notion of Bayes-compatibility at each profile in a certain domain. This is a generalization since, unlike what the traditional approach implies and a full unpacking of (12) would accordingly reflect, nothing of principle in the new approach necessitates that the domain in question be defined with reference to a set of states of measure 1. For illustration and by way of anticipation, some of the generalizations we will present in Section 6 will be with reference to domains of profiles that need not have unit measure.

the equality in (12). The problem is whether any pooling function can capture this supra-Bayesian response, i.e., whether there exists a left-hand side to realize, for all $(p|E_1, \ldots, p|E_n)$ in Δ^n_{CP}, the equality in (12). The answer to this question is, by contrast, not transparent. We now turn to providing it.

4.2 Characterizing Full-Revelation Bayes-Compatibility

We now proceed to stating our results. As a preliminary check, we start with the observation that, degenerate cases aside, linear pooling cannot be full-revelation Bayes-compatible. We recapitulate the most relevant assumptions made so far. The proofs are in the Appendix.

Proposition 1. *Let* $(p|E_1, \ldots, p|E_n)$ *be any profile in* Δ^n_{CP}. *Let* $\{\alpha_i\}_{i \in N}$ *be any weights such that* $\alpha_i \in (0,1)$ *for all* i *and* $\sum_{i=1}^n \alpha_i = 1$. *The following are equivalent:*

1. $\sum_{i=1}^n \alpha_i p(A|E_i) = p\left(A \middle| \bigcap_{i \in N} E_i\right)$ *for all* $A \in \mathcal{A}$.
2. $E_j = E_k$ *for all* $j, k \in N$.

The above proposition can be related to the impossibility result mentioned after equation (7). Given their heterogeneous assumptions about social information, the two results are not directly comparable. This fact notwithstanding, the above proposition offers an appreciably simpler insight into the Bayes-incompatibility of linear pooling than currently available in the literature—compare the respective proofs of Proposition 1 and Corollary 1 in Bradley, 2018, in particular. The gain in simplicity crucially hinges on the strength of the full revelation assumption. This assumption, which is motivated by our interest in geometric pooling, is not imposed in the rest of the literature. There, in line with a prevailing interest in linear pooling, it is assumed that agents reveal only their posterior probability value for one event. Nevertheless, what can be taken from Proposition 1 is that the Bayes-incompatibility of linear pooling does not depend on that restriction. Indeed, under linear pooling, Bayes-incompatibility prevails even when full revelation is assumed. To this extent, linear pooling can be deemed radically Bayes-incompatible.

Now to our main result. It establishes, among other things, that a pooling function is full-revelation Bayes-compatible if and only if it is geometric. Once again, in line with the caveat with which we ended last section, we wish to stress upfront that while it is almost immediate that, under our assumptions, the n probabilistic reports should correspond to the event $\bigcap_{i \in N} E_i$, it is not immediately transparent that conditioning on this event would coincide with geometric pooling. Now, to state the entire proposition, we introduce two preexisting pooling axioms. The first property is arguably the simplest of all unanimity preservation constraints. The second property is a particular kind of commutativity with Bayesian conditioning, named here as in Dietrich, 2017 (p. 6). It essentially requires that anything learned by one single agent is reflected in the pooled probability.[19]

Unanimity. For any probability measure $p \in \Delta$, $F(p, \ldots, p) = p$.

Private Bayesianism. For any profile $(p_1, \ldots, p_i, \ldots, p_n) \in \Delta^n$ and any event $E \in \mathcal{A}$, $F(p_1, \ldots, p_i|E, \ldots, p_n) = F(p_1, \ldots, p_i, \ldots, p_n)|E$.

The theorem can now be stated.

---

[19]As regards the second axiom, recall that, provided $F(p_1, \ldots, p_n)(E) > 0$, $F(p_1, \ldots, p_n)|E$ is simply the conditionalization of $F(p_1, \ldots, p_n)$ on information $E \in \mathcal{A}$.

Theorem 1. *Let F denote any pooling function on Δ_{CP}^n. The following are equivalent:*

1. *For any profile $(p|E_1, \ldots, p|E_n) \in \Delta_{\mathrm{CP}}^n$ and all $A \in \mathcal{A}$: $F(p|E_1, \ldots, p|E_n)(A) = p\left(A| \bigcap_{i \in N} E_i\right)$.*

2. *For any profile $(p|E_1, \ldots, p|E_n) \in \Delta_{\mathrm{CP}}^n$, any weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \neq 0$ for all $i \in N$ and $\sum_{i=1}^{n} \alpha_i = 1$, with c induced as in (3), for all $\omega \in \Omega$: $F(p|E_1, \ldots, p|E_n)(\omega) = c \prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i}$.*

3. *F satisfies Unanimity and Private Bayesianism.*

While the Appendix contains a full proof of the theorem, it may be helpful to convey here the main intuition behind the equivalence of 1 and 2. Two cases need to be considered. If $\omega \notin \bigcap_{i \in N} E_i$, then, for some $i \in N$, $p(\omega|E_i) = 0 = p(\omega|\bigcap_{i \in N} E_i)$. By the support-veto property of geometric pooling, this implies that, for any admissible set of weights, $c \prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i} = 0$. On the other hand, if $\omega \in \bigcap_{i \in N} E_i$, then simple algebra shows that, for any admissible set of weights, $c \prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i} = p\left(\omega \Big| \bigcap_{i \in N} E_i\right)$. Thus, it must be that a pooling function F is Bayes-compatible if and only if it is geometric with non-zero normalized weights, i.e., 1 and 2 are equivalent.

Stepping back, several comments are in order. First, the equivalence between parts 2 and 3 of Theorem 1 amounts to a characterization of geometric pooling with weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$. This characterization is interesting for its simplicity. But it is important to realize, to start with, that it is here proved to hold only over the domain Δ_{CP}^n. Indeed, the common prior assumption induces a domain restriction to the set of all profiles coming by Bayesian updating from some common prior. For instance, with the state space $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and the probability measures p_1 and p_2 such that $p_1(\{\omega_1, \omega_2\}) = 1/2$ and $p_2(\{\omega_1, \omega_2\}) = 1/3$, the profile (p_1, p_2) is not contained in this set, and so does not fall under the scope of Theorem 1.

Furthermore, it is equally important to note that the weights in the above characterization of geometric pooling are not unique, and this non-uniqueness cannot be improved on the domain over which the characterization has been proved to hold. To be sure, the weights $\{\alpha_i\}_{i \in N}$ can all be taken in $(0, 1)$, or even equal to $1/n$, without altering the result of the pool. One interpretation of this flexibility is that, for the particular epistemic purpose under consideration, geometric pooling is robust against expert miscalibration. It delivers the supra-Bayesian solution even when the agent with the most specific information is given the smallest weight in the pool.[20] That being said, one should not lose track of the fact that taking the weights to be all equal, or even positive, can only be a matter of convention in our setting. In fact, negative weights are acceptable, too, provided the weights sum up to 1 over N. At this point, in all fairness, it might be asked whether allowing weights to be negative would open up new possibilities for the Bayes-compatibility of linear pooling. However, it is well-known that linear pooling cannot accommodate negative weights if—as indicated here by the

---

[20]More generally, geometric pooling thus proves immune, in our setting, to any kind of manipulation of the pooling result that one would want to implement through a manipulation of the pooling weights.

notation—the pooling weights are required to be profile-independent.[21]

Second, as the notation also indicates, the weights $\{\alpha_i\}_{i \in N}$ featured in Theorem 1 are event-independent. Unlike linear pooling, geometric pooling can accommodate at least some forms of event-dependence in the pooling weights. Consider, for example, the limiting case of state-dependent weights, that still induce a well-defined probability measure under geometric pooling. However, nothing guarantees the Bayes-compatibility of geometric pooling in those cases, and we will not explore them any further here.

Third, Proposition 1 follows from Theorem 1 as a corollary. It corresponds to the degenerate case where linear and geometric pooling coincide on all events. In a common prior setting, this happens exactly on the unanimous profiles, i.e., when all agents announce the same posterior probability measure.

Fourth, Theorem 1 highlights the links between Bayes-compatibility and commutativity with Bayesian conditionalization (here expressed, specifically, by the Private Bayesianism axiom). Whatever the exact assumption one wishes to impose regarding social information, this connection is directly relevant to understanding why linear pooling fails to be Bayes-compatible. However, the relevance of commutativity with Bayesian conditionalization to understanding Bayes-compatibility appears less clearly in the previous treatments of the problem than in ours. This, too, is made possible by the fact that we are investigating the problem of Bayes-compatibility under the restrictive assumption of full-revelation.

5 Interpreting Full-Revelation Bayes-Compatibility

This section is dedicated to putting full-revelation Bayes-compatibility, and its characterization by geometric pooling in Theorem 1, in broader philosophical perspective. We first detail how requiring full-revelation Bayes-compatibility can be related to respecting the "principle of total evidence," a pillar of Bayesian epistemology. Next, we explain how this principle and Theorem 1 shed light on the phenomenon of "synergy" that has recently received some attention in the social epistemology literature. Finally, we discuss another recent central topic of this literature, "epistemic deference," in connection to Bayes-compatibility, geometric pooling, the principle of total evidence, and synergy.

5.1 Total Evidence

Consider, say, agent 1. She privately receives information E_1 and subsequently learns, through public announcements, that the posteriors of the other agents are $p|E_2, \ldots, p|E_n$. Under the assumptions spelled out in the previous section, $\bigcap_{i \in N} E_i$ corresponds to all the evidence she is directly or indirectly given regarding the true state of nature. Similarly, for any $i \in N$, i's total evidence is given by $\bigcap_{i \in N} E_i$. This is why there is no need to distinguish between (5) and (6) under our assumptions. An often implicit pillar of Bayesian epistemology known as

---

[21] For some profiles and profile-dependent weights, if negative weights are allowed, linear pooling can behave in a Bayes-compatible way. Consider for example $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, p be the uniform distribution on Ω, $p_1 = p|\{\omega_1, \omega_2\}$, $p_2 = p_3 = p|\{\omega_1, \omega_2, \omega_3\}$. Then, with profile-dependent weights $\alpha_1 = \alpha_2 = 1$, $\alpha_3 = -1$, the linear pool of this profile is well-defined and full-revelation Bayes-compatible. Notice by contrast that, to secure Bayes-compatibility, geometric pooling requires no profile-dependence or sign condition on the weights. Moreover, allowing negative weights for linear pooling would require introducing still further coherence constraints on weights to ensure that the pool itself is a probability function (see, for example, Genest, 1984).

the principle of total evidence (Carnap, 1947; Good, 1967) requires that a Bayesian agent conditionalize on all of her evidence. What Proposition 1 and Theorem 1 show is that, in our setting, geometric pooling, as a mechanism for learning from others, always respects this principle while linear pooling systematically violates it. This is illustrated particularly vividly by how each pooling method treats the states to which some, but not all, agents give posterior probability zero. By putting some probability mass on such states, linear pooling disregards the part of the available evidence entailing that these states should be left out of the support of any Bayesian posterior. By contrast, because of the support-veto property, geometric pooling efficiently aggregates all of the evidence to which the group collectively has access.

Furthermore, it has been remarked in the literature (Bradley, 2006, p. 145-146) that if individuals are required to respect the principle of total evidence, pooling problems will generally be, in one important respect, *unstable*. The key observation behind this assessment is that the very statement of a pooling problem, which includes the profile of opinions, conveys information to the agents whose opinions are being pooled. Thus, whenever the agents are informed of the profile, they are required by the principle of total evidence to revise their opinions. This means that the initial pooling problem transforms into a new one, generally leading, through the pooling function of interest, to a different solution. But the new problem similarly contains new information, thus leading to still another problem and solution, and so on. In light of this, it is worth highlighting that in our common prior setting and under geometric pooling, no such instability arises. In our setting, the aggregation problem informs the agents of $\bigcap_{i \in N} E_i$, which is identical to the intersection of the supports. Suppose that the agents update on that information, in accordance with the principle of total evidence. Then, Theorem 1 ensures that the new pool does *not* differ from the initial one under geometric pooling. To see this, observe that, for any profile in Δ_{CP}^n, for all $i \in N$, and any $\omega \in \Omega$, we have the following chain of equalities.

$$c \prod_{i=1}^n p_i(\omega)^{\alpha_i} = p\left(\omega \,\middle|\, \bigcap_{i \in N} E_i\right) = p_i\left(\omega \,\middle|\, \bigcap_{i \in N} E_i\right) = c \prod_{i=1}^n p_i\left(\omega \,\middle|\, \bigcap_{i \in N} E_i\right)^{\alpha_i} \tag{13}$$

5.2 Synergy

Another implication of Theorem 1 is worth singling out. Say that a pooling function F displays *synergy* (see Christensen, 2009; Easwaran et al., 2016, Sec. 6; Dietrich, 2010, Sec. 1) if for some event A, $F(p_1, ..., p_n)(A) \notin [\min_{i \in N} p_i(A), \max_{i \in N} p_i(A)]$. Synergy thus violates the unanimity preservation constraint requiring that if $p_i(A) \in [x, y]$ for all $i \in N$, then $F(p_1, ..., p_n)(A) \in [x, y]$. Notice that this is a strengthening of a standard unanimity preservation constraint (weaker than the one mentioned in Section 4.2) to the effect that if $p_i(A) = x$ for all $i \in N$, then $F(p_1, ..., p_n)(A) = x$ (e.g., Mongin, 1995, p. 314). Synergy is of interest in the present paper because, mutatis mutandis, it can be displayed by supra-Bayesianism.[22] The following is an informal and elementary illustration. Assume that, based on independent sources of information, n agents report the same very high probability value x for event A. Then it would be conceivable for at least one of the agents to consider that the independent

---

reports *confirm* the truth of A, and raise her probability for A from x to $y > x$. This would be a simple instance of synergy. A related phenomenon has been called *extremizing* and discussed in the recent statistical literature on expert forecasting (see esp. Baron et al., 2014; Tetlock and Gardner, 2016, Chap. 9; and Romeijn, 2019b for a philosophical elaboration). We henceforth focus on synergy while noting that the links with extremizing would deserve further study.

Assuming weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$, linear pooling cannot display synergy. By construction, it will always be the case that, for any A, $\sum_{i=1}^{n} \alpha_i p_i(A) \in [\min_{i \in N} p_i(A), \max_{i \in N} p_i(A)]$. While this can be considered an advantage for certain purposes,[23] it is a limitation for a learning heuristic. Geometric pooling can, by contrast, display synergy. We illustrate this with the following two examples. In these examples, in order to to facilitate comparison with linear pooling, we restrict attention to geometric pooling with weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$.

Example 1. *Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and let p be the uniform distribution on Ω. Let p_1 and p_2 be the probability measures obtained from p by Bayesian conditioning on $E_1 = \{\omega_1, \omega_2, \omega_3\}$ and $E_2 = \{\omega_1, \omega_4\}$, respectively. Then, the profile (p_1, p_2) is given in the table below.*

	ω_1	ω_2	ω_3	ω_4
p_1	1/3	1/3	1/3	0
p_2	1/2	0	0	1/2

Consider the event $A = \{\omega_1, \omega_3\}$. Notice that $[\min_{i \in N} p_i(A), \max_{i \in N} p_i(A)] = [1/2, 2/3]$ and $[\min_{i \in N} p_i(A^c), \max_{i \in N} p_i(A^c)] = [1/3, 1/2]$. For any geometric pool F with weights in $(0, 1)$, it will be the case that $F(p_1, p_2)(A) = 1 > \max_{i \in N} p_i(A)$ and $F(p_1, p_2)(A^c) = 0 < \min_{i \in N} p_i(A^c)$.

The next example immediately clarifies that, contrary to what one might be tempted to conclude on the basis of Example 1, different supports for the p_i is not a necessary condition for geometric pooling to display synergy. It also serves as a reminder of the fact that geometric pooling can violate the standard unanimity preservation constraint mentioned earlier in this section, in the special case when for some A, $\min_{i \in N} p_i(A) = \max_{i \in N} p_i(A)$.

Example 2. *Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and consider the profile given in the table below.*

	ω_1	ω_2	ω_3
p_1	1/4	1/4	1/2
p_2	1/4	1/2	1/4

Observe that $\min_{i \in N} p_i(\{\omega_1\}) = \max_{i \in N} p_i(\{\omega_1\}) = 1/4$. However, with weights $\alpha_1 = \alpha_2 = 1/2$, the geometric pool $F(p_1, p_2)(\{\omega_1\}) = 1/(1 + 2\sqrt{2}) > 1/4$. Similarly, while $\min_{i \in N} p_i(\{\omega_1\}^c) = \max_{i \in N} p_i(\{\omega_1\}^c) = 3/4$, with weights $\alpha_1 = \alpha_2 = 1/2$, the geometric pool $F(p_1, p_2)(\{\omega_1\}^c) = 2\sqrt{2}/(1 + 2\sqrt{2}) < 3/4$.

[23]See, e.g., Lehrer and Wagner, 1981, p. 127: it "has the desirable feature of maintaining disagreement within its original bounds."

Nevertheless, in a common prior setting, the form of synergy illustrated by Example 2 cannot arise; only the form illustrated by Example 1 can. Indeed, in such a setting and under the assumption that the pooling function is geometric, one can show that synergy is closely related to the fact that the p_i have different supports. This is the content of the following characterization, which we provide for the general case of geometric pooling with weights $\{\alpha_i\}_{i\in N}$ such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^n \alpha_i = 1$. Here, as elsewhere, we write $\subsetneq$ (respectively, $\subseteq$) for the strict (respectively, weak) subset relation.

Proposition 2. *Let $(p|E_1, \ldots, p|E_n)$ be any profile in Δ_{CP}^n. Let F be a geometric pooling function with any weights $\{\alpha_i\}_{i\in N}$ such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^n \alpha_i = 1$. The following are equivalent:*

  1. *There exists $A \in \mathcal{A}$ such that $F(p|E_1, ..., p|E_n)(A) \notin [\min_{i\in N} p(A|E_i), \max_{i\in N} p(A|E_i)]$.*
  2. *For all $j \in N$, $\bigcap_{i\in N} E_i \subsetneq E_j$.*

Proposition 2 can be interpreted as follows. In a common prior setting, geometric pooling displays synergy if and only if there is more information in the group $(\bigcap_{i\in N} E_i)$, than any single member has access to (E_j). This is an intuitive result. Under the characterizing condition, geometric pooling amounts to conditioning the prior on an event that is stronger than any of the events a single agent has conditioned on. As a result, the geometric pool gives unit probability to an event that no individual probability measure does.

Finally, we note that the above interpretation sheds light on the stability property highlighted in the previous section. Specifically, the stability of our aggregation problem under geometric pooling relies on the fact that geometric pooling aggregates all the information distributed in the group.

5.3   Deference

We now discuss the links between "epistemic deference" (Gaifman, 1988; Elga, 2007), full-revelation Bayes-compatibility, and geometric pooling. In general terms, epistemic deference is the phenomenon whereby one agent constrains her subjective probabilities to take certain values in light of some other probabilistic assessment. This other assessment could be a relative frequency (Reichenbach, 1949), an "objective" chance distribution (Levi, 1977), the beliefs of another agent (Bradley, 2018), or the beliefs of the same agent at some future time (van Fraassen, 1984). Adopting another's beliefs as one's own is one of the most straightforward illustrations. An agent has motivation to defer to another's probabilistic judgment if she considers the other agent, in some sense, an *expert*. An agent may be considered an expert because she has more information, or because she is more skilled at assessing information. We will be concerned here with the former motivation, called "information deference" in the philosophical literature (Bradley, 2018, p. 10; on the latter motivation, see Bradley, 2018, Sec. 3.2). The assumption of a common prior provides a natural setting in which to investigate this kind of epistemic deference. Starting from a common prior, a less-informed agent has grounds to adopt the posterior beliefs of any more-informed agent if she learns them. In the statistical literature, this kind of epistemic deference has been investigated—not necessarily with a common prior restriction—under the label "probability calibration" (Lindley, 1982b, Sec. 1; French, 1986, Sec. 4; Dawid et al., 1995, Sec. 2.1).[24]

---

[24] This is not to be confused with calibration understood as matching subjective probabilities to frequencies.

The analysis of information deference can be refined by distinguishing between "full" and "partial" deference (Bradley, 2018, Sec. 4). This is best explained by temporarily reverting, till the end of the present section, to the language of random variables. Taking the reported beliefs of another agent as one's own amounts to *fully* deferring to her. The simplest illustration is a two-agent common prior case in which only one agent—say, j—has received new information. Then, the less-informed agent i has grounds to adopt the reported posterior beliefs of the more-informed agent j. More generally, even if both agents have received new information, for any $\omega \in \Omega$, if $E_j(\omega) \subseteq E_i(\omega)$, then agent j has (weakly) more evidence than i. Thus, if $E_j(\omega) \subseteq E_i(\omega)$, $p_i\left(\cdot \left| p|E_j(\cdot) = p|E_j(\omega)\right.\right) = p_i\left(\cdot \left| E_j(\omega)\right.\right) = p\left(\cdot \left| E_i(\omega) \cap E_j(\omega)\right.\right) = p|E_j(\omega)$, which is the very definition of full information deference as we understand it. Therefore, deferring to j is the same as conditioning p on the intersection of the supports of p_i and p_j. By Theorem 1, in our setting, this is equivalent to geometrically pooling i's and j's posteriors.

But consider now a three-agent common prior case in which only two agents—say, j and k—have received new information. Assume that neither of the better-informed agents has more information than the other, in the sense that given $\omega \in \Omega$, neither $E_j(\omega) \subseteq E_k(\omega)$ nor $E_k(\omega) \subseteq E_j(\omega)$ holds. Then, upon learning the updated beliefs of both j and k, the less-informed agent i should not fully defer to either j or k—for example, to j. Indeed, this would amount to disregarding some of the evidence available to her, to wit (in the example), whatever information is contained in k's report. Rather than fully deferring to any single expert, *partially* deferring to both seems to be in order. At least when the full posteriors are reported, i.e., under the full revelation assumption, the content of the Bayesian response remains clear, which is what we wish to emphasize here (see also Bradley, 2018, Eq. 8). Agent i should just exploit the information that the more informed agents have by conditioning on the intersection of the supports of their reported posteriors: for any $\omega \in \Omega$, $p_i\left(\cdot \left| p|E_j(\cdot) = p|E_j(\omega), p|E_k(\cdot) = p|E_k(\omega)\right.\right) = p\left(\cdot \left| E_i(\omega) \cap E_j(\omega) \cap E_k(\omega)\right.\right)$. Thus, while full deference amounts to conditioning the common prior on the evidence received by the most-informed agent, when such an individual exists, partial deference amounts to conditioning the prior on the total evidence collectively received by all agents, which applies even when information is scattered across the group. As Theorem 1 attests, in our setting, geometric pooling secures the latter, more general form of deference. Therefore, geometric pooling constitutes an attractive model for partial deference under the full-revelation assumption.

By contrast, when only probabilities for a fixed event and not entire distributions are reported, agents do not reveal which cells of their respective information partitions they have learned. As a result, admittedly, what partial deference precisely amounts to is much less straightforward. We take no stance on how it should then be modelled.[25]

We conclude this section by observing that the concept of information deference also affords another take on Proposition 2 which characterizes synergy under geometric pooling in a common prior setting. This proposition can be interpreted as stating that geometric pooling displays synergy if and only if there is no particular member of the epistemic community to whom everyone should fully defer. In other words, it is only to the group, and not to any particular member thereof, that each member should fully defer. Indeed, for any $\omega \in \Omega$, all members of the epistemic community N should fully defer to member j if and only if for any

---

[25]As a referee suggested, with the benefit of hindsight, one may doubt whether, by focusing directly and almost exclusively on this case, the philosophical literature has taken the most natural route to understanding deference.

$i \in N$, $E_j(\omega) \subseteq E_i(\omega)$, i.e., $E_j(\omega) = \bigcap_{i \in N} E_i(\omega)$. Mutatis mutandis, this is exactly what is excluded by condition 2 in Proposition 2.

6 Generalizing Full-Revelation Bayes-Compatibility

The central result of our paper is the characterization of geometric pooling by Bayes-compatibility given in Theorem 1. The goal of this section—which may be skipped by those wishing to restrict themselves to our core conceptual take-aways—is to better understand this characterization by answering two questions, considered first separately from one another, then together. First, Theorem 1 characterizes geometric pooling by Bayes-compatibility under (the full revelation assumption and) the common prior assumption. This assumption restricts the domain of the pooling function to Δ_{CP}^n. Can the characterization be extended to a larger subdomain of $\Delta^{n'}$ by relaxing the common prior assumption to some extent? Second, Theorem 1 characterizes geometric pooling by Bayes-compatibility under (the common prior assumption and) the full revelation assumption. This is a strong requirement on how much of their private information agents reveal. Can a supra-Bayesian case for geometric pooling still be made under weaker informational demands than the full revelation assumption? Finally, in case both of the above constraints can be interestingly relaxed, one would also like to know whether they can be relaxed simultaneously. The main question being here: Does the comparative advantage of geometric pooling over linear pooling documented in Theorem 1 persist both under a weaker informational requirement than the full revelation assumption and over a larger domain than Δ_{CP}^n? As we now explain and discuss in detail, the answer to each of the above questions is a qualified "Yes."

To the various qualifications we will detail along the way, we add the following general caveat. In this section, we engage in further axiomatic analysis of the properties in Theorem 1 even when this puts a strain on the epistemic interpretation of the Bayes-compatibility properties under examination. For instance (recalling our remarks following Equation (6)), some of our results will make good epistemic sense only when one takes the perspective of some external observer of the group, not the ex ante perspective of any specific member. Admittedly, and as we will highlight, this narrows the epistemic implications of our results. But unlike in the previous section, our main goal in the present section is to explore the robustness of our main characterization. Complementing our previous emphasis on possible philosophical interpretations, our emphasis now is on possible mathematical generalizations.

6.1 Beyond the Common Prior Assumption?

We focus first on the common prior assumption. Recall that Δ_{CP}^n is the domain of all profiles coming by Bayesian updating from some common prior p, i.e., profiles of the form $(p_1, \ldots, p_n) = (p|E_1, \ldots, p|E_n)$. Notice a key property of such profiles. Say that p_j is proportional to p_k on E, written $p_j \propto p_k$ on E, if $p_i(\omega)/p_i(\omega') = p_j(\omega)/p_j(\omega')$ for all $\omega, \omega' \in E$ such that these ratios are defined. Now, by the definition of conditional probability, all profiles $\{p_i\}_{i \in N}$ in Δ_{CP}^n are such that $p_j \propto p_k$ on $\bigcap_{i \in N} supp(p_i)$ for all $j, k \in N$. However, the profiles in Δ_{CP}^n are not the only ones to display such proportionality. Indeed, to assume that the agents start from a common prior $p = p_j = p_k$ for all $j, k \in N$ is tantamount to assuming that, *whatever* pieces of evidence $\{E_i\}_{i \in N}$ the agents receive, their resulting posteriors

20

$\{p_i|E_i\}_{i\in N}$ are proportional on the intersection of their supports. But such proportionality can hold in some non-common prior cases as well. That is to say, for *some* different priors $\{p_i\}_{i\in N}$ and *some* possible pieces of evidence $\{E_i\}_{i\in N}$, the resulting posteriors $\{p_i|E_i\}_{i\in N}$ are proportional on the intersection of their supports. And, as we now explain, this is in fact the property most tightly connected to the characterization in Theorem 1.

Let $\Delta^n_\propto$ be the domain of all profiles such that $p_j \propto p_k$ on $\bigcap_{i\in N} supp(p_i)$ for all $j,k \in N$.[26] As explained above, $\Delta^n_\propto$ is a strictly larger domain than Δ^n_{CP}. The following result characterizes the domain $\Delta^n_\propto$, and is preliminary to the main result of this section, stated next.

Proposition 3. *Let $\{p_i\}_{i\in N}$ be any profile in $\Delta^{n'}$. Let $\{\alpha_i\}_{i\in N}$ be any weights such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$. The following are equivalent:*

1. *$c\sum_{\omega\in A}\prod_{i=1}^{n} p_i(\omega)^{\alpha_i} = p_i\left(A\middle|\bigcap_{i\in N} supp(p_i)\right)$ for any $i \in N$, $A \in \mathcal{A}$.*

2. *For all $j,k \in N$, $p_j \propto p_k$ on $\bigcap_{i\in N} supp(p_i)$.*

Proposition 3 establishes that a profile is in $\Delta^n_\propto$ if and only if geometric pooling with non-zero normalized weights behaves in a full-revelation Bayes-compatible way with respect to it.[27] Thus, $\Delta^n_\propto$ is exactly the domain over which geometric pooling with non-zero normalized weights behaves in a full-revelation Bayes-compatible way. Now, it is worth explaining why the characterization cannot be given for a larger, simpler, domain. First, it is only informally that a profile in $\Delta^n_\propto$ can be considered a profile of "posteriors" in the context of interest in Proposition 3. The characterization holds directly and only for $\Delta^n_\propto$.[28] Second, the full-revelation Bayes-compatibility condition 1 in Proposition 3 is phrased with respect to event $\bigcap_{i\in N} supp(p_i)$, rather than $\bigcap_{i\in N} E_i$ as in Theorem 1 and other previously stated propositions. This is because (as we stressed in section 4.1) without the assumption of full support for the prior, the announcement of one's posterior support need not coincide with the announcement of one's private information. As the profiles in $\Delta^n_\propto$ cannot be rigorously considered posterior profiles in the context of Proposition 3, it a fortiori holds that the support of the measures in those profiles cannot be rigorously identified with the pieces of information the agents might have received.

Now, recall the first question asked at the beginning of Section 6: Can the characterization in Theorem 1 be extended to a larger subdomain of $\Delta^{n'}$ than Δ^n_{CP}? Proposition 3 is

---

[26]To facilitate comparison with our initial approach in terms of random variables, this domain may be defined as follows. For all $i \in N$, let $p_i(\cdot,\cdot) : \mathcal{A}\times\Omega \to \mathbb{R}$ be a random probability measure. Such a random measure can be interpreted as representing, for example, the uncertainty of an external agent about the opinions of in-profile agents. For any $\omega \in \Omega$, $p_i(\cdot,\omega) \in \Delta$ is a probability measure, just like $p|E_i(\cdot)$ before. However, unlike before, we do not assume that the p_i are random posteriors coming from a common prior p and individual information partitions for each i. Instead, we merely assume that, for all $i \in N$ and any $\omega \in \Omega$, $(p_1(\cdot,\omega),\ldots,p_n(\cdot,\omega)) \in \Delta^{n'}$ and, for all $j,k \in N$, $p_j(\cdot,\omega) \propto p_k(\cdot,\omega)$ on $\bigcap_{i\in N} supp(p_i(\cdot,\omega))$. Let B be the index set for all such n-tuples of random measures. The profiles that are possible values of $(p_1(\cdot,\cdot),\ldots,p_n(\cdot,\cdot))_\beta$ for any $\beta \in B$ define a new domain, and it is this domain that we denote by $\Delta^n_\propto = \left\{(p_1(\cdot,\omega),\ldots,p_n(\cdot,\omega))_\beta \in \Delta^{n'} : \beta \in B \text{ and } \omega \in \Omega\right\}$.

[27]The domain considered in Proposition 3 is the full domain $\Delta^{n'}$, over which geometric pooling will behave in a Bayes-compatible way for some profiles, and fail to do so for others. Thus, when defined over $\Delta^{n'}$, geometric pooling is certainly not Bayes-compatible according to our definition. However, based on Proposition 3, one can define the partial domain $\Delta^n_\propto$ over which geometric pooling will always behave in Bayes-compatible way.

[28]More generally, it would be interesting to furnish $\Delta^n_\propto$ with an epistemic interpretation that could parallel and generalize the one already presented (in terms of a team of scientists distributing various experiments between themselves) for Δ^n_{CP}. To our knowledge, no such interpretation is currently available in the literature.

instrumental in providing a qualified positive answer to this question. The qualification is that in what follows, like in Proposition 3, the property of full-revelation Bayes-compatibility is generalized to refer to $\bigcap_{i \in N} supp(p_i)$, rather than $\bigcap_{i \in N} E_i$. Subject to this qualification regarding how the full revelation condition is cashed out, the following proposition constitutes a positive answer.

Proposition 4. *Let F be a pooling function on $\Delta_\propto^n$. The following are equivalent:*

1. *For any profile $\{p_i\}_{i \in N} \in \Delta_\propto^n$, all $A \in \mathcal{A}$, $i \in N$: $F(p_1, \ldots, p_n)(A) = p_i\left(A \middle| \bigcap_{i \in N} supp(p_i)\right)$.*

2. *For any profile $\{p_i\}_{i \in N} \in \Delta_\propto^n$, any weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \neq 0$ for all $i \in N$ and $\sum_{i=1}^n \alpha_i = 1$, with c induced as in (3), for all $\omega \in \Omega$: $F(p_1, \ldots, p_n)(\omega) = c\prod_{i=1}^n p_i(\omega)^{\alpha_i}$.*

3. *F satisfies Unanimity and Private Bayesianism.*

Like Theorem 1, Proposition 4 offers a simple characterization of geometric pooling. Remarks similar to those made after Theorem 1 apply to Proposition 4 with regard to the limited domain of this characterization, and the non-uniqueness of the pooling weights. But most importantly for our purposes, with appropriate adjustments to the epistemic interpretations and under the aforementioned qualification regarding how the full revelation condition is cashed out, Proposition 4 shows that the equivalence between geometric and Bayes-compatible pooling presented in Theorem 1 generalizes beyond Δ_{CP}^n, namely, to $\Delta_\propto^n$. To this extent, the equivalence does not crucially depend on the common prior assumption.

As we initially did for Δ_{CP}^n, we can contrast the behavior of geometric and linear pooling over $\Delta_\propto^n$. The following result thus parallels Proposition 1.

Proposition 5. *Let $\{p_i\}_{i \in N}$ be any profile in $\Delta_\propto^n$. Let $\{\alpha_i\}_{i \in N}$ be any weights such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$. The following are equivalent:*

1. *$\sum_{i=1}^n \alpha_i p_i(A) = p_i\left(A \middle| \bigcap_{i \in N} supp(p_i)\right)$ for all $i \in N$, $A \in \mathcal{A}$.*

2. *$p_j = p_k$ for all $j, k \in N$.*

Subject to the qualification highlighted before Proposition 4, Proposition 5 establishes that the impossibility result in Proposition 1 generalizes from Δ_{CP}^n to $\Delta_\propto^n$. Therefore, this impossibility does not crucially depend on the common prior assumption. A joint implication of Propositions 4 and 5 is that, from a supra-Bayesian perspective, geometric pooling conserves a clear-cut comparative advantage over linear pooling beyond the traditional common prior setting.

Our final question in this section will be whether this comparative advantage of geometric pooling over linear pooling could generalize, perhaps in a less clear-cut way, even beyond $\Delta_\propto^n$. Now, it is clear from Proposition 3 that geometric pooling will not be full-revelation Bayes-compatible beyond $\Delta_\propto^n$. However, as we now explain, a weaker notion of full-revelation Bayes-compatibility is still worth investigating over a larger domain. So far, we have construed Bayes-compatibility as holding at every profile in a certain domain and, given any profile, for all agents. Instead of these two universal quantifications, consider now defining Bayes-compatibility *for some $i \in N$, at some $\{p_i\}_{i \in N} \in \Delta^{n'}$.* Setting aside relevant discontinuities

in the epistemic interpretations, this very weak notion of Bayes-compatibility is, in fact, the building block of all our previous definitions. Formally, with $\{p_i\}_{i \in N}$ some profile in $\Delta^{n'}$, call a pooling function F full-revelation Bayes-compatible *at profile $\{p_i\}_{i \in N}$ with respect to agent i* if, for any event $A \in \mathcal{A}$, the following equality holds:

$$F\left(p_1, \ldots, p_n\right)(A) = p_i\left(A \middle| \bigcap_{i \in N} supp(p_i)\right). \tag{14}$$

The following proposition provides a partial answer to the final question raised above. In the proposition, $\Delta_{\ll}^n$ denotes the domain of profiles containing only mutually absolutely continuous measures, i.e., measures with common support.[29]

Proposition 6. *Let $\{p_i\}_{i \in N}$ be any profile in $\Delta^{n'}$. Let $\{\alpha_i\}_{i \in N}$ be any weights such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$.*

1. *$c \sum_{\omega \in A} \prod_{i=1}^n p_i(\omega)^{\alpha_i} = p_j(A | \bigcap_{i \in N} supp(p_i))$ for some $j \in N$ and all $A \in \mathcal{A}$ if and only if $p_j \propto c \prod_{i=1}^n p_i^{\alpha_i}$ on $\bigcap_{i \in N} supp(p_i)$.*

2. *If $\sum_{i=1}^n \alpha_i p_i(A) = p_j(A | \bigcap_{i \in N} supp(p_i))$ for some $j \in N$ and all $A \in \mathcal{A}$, then $p_j = \sum_{i=1}^n \alpha_i p_i$. Furthermore, when $\Delta^{n'}$ is restricted to $\Delta_{\ll}^n$, the consequent in the previous sentence also implies the antecedent.*

Several comments are in order. First, it is worth highlighting that the condition investigated by Proposition 6, viz. Property (14), is only a Bayes-compatibility-like condition that structurally differs from the Bayes-compatibility with respect to agent i defined in Equation (5). The crucial difference being that, as indicated by the very name given to Property (14), the condition may hold at certain profiles of the domain, and fail at others. Second, it also bears emphasis that in Proposition 6, be it for geometric or for linear pooling, the pooling weights matter in the following sense. Given a profile in $\Delta^{n'}$, be it under geometric or linear pooling, the Bayes-compatibility-like condition in (14) may hold for some sets of weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$, and fail for others. This is another contrast with the possibility (respectively, impossibility) results in Theorem 1 and Proposition 4 (respectively, Propositions 1 and 5), where only general constraints on the weights matter.

Third, the righthand side conditions in Proposition 6 are strong and will generally not be met by either geometric or linear pooling. Indeed, normalizing the weights, one can rephrase these conditions as stating the mathematically restrictive fact that j's probability function is in fact already a pool of the profile, i.e., j is already an "aggregator". In the context of linear pooling, "aggregator" is thus shorthand for "linear aggregator with a normalization of the weights used in the linear pooling function." A similar explication applies regarding geometric

---

[29]To facilitate comparison with our initial random variable approach, for $i \in N$, let $p_i(\cdot, \cdot) : \mathcal{A} \times \Omega \to \mathbb{R}$ be a random probability measure. Thus, for all $i \in N$ and any $\omega \in \Omega$, $p_i(\cdot, \omega) \in \Delta$. We now further assume that, for all $\omega \in \Omega$ and any $i, j \in N$, $p_i(\cdot, \omega) \ll p_j(\cdot, \omega)$. Let Γ be the index set for all n-tuples satisfying those assumptions. The possible values of $(p_1(\cdot, \cdot), \ldots, p_n(\cdot, \cdot))_\gamma$ for $\gamma \in \Gamma$ and $\omega \in \Omega$ define a new domain of profiles of probability functions. We let this domain be denoted by $\Delta_{\ll}^n = \left\{ (p_1(\cdot, \omega), \ldots, p_n(\cdot, \omega))_\gamma \in \Delta^{n'} : \gamma \in \Gamma \text{ and } \omega \in \Omega \right\}$.

pooling. The condition that a given profile already features either kind of aggregator will, when some richness conditions are satisfied, generically be violated.[30]

Fourth, there is nonetheless a sense in which geometric pooling achieves the relevant form of Bayes-compatibility under less restrictive conditions than linear pooling. The equality constraints necessary under linear pooling are stronger than the relevant proportionality constraints for geometric pooling. Indeed, under geometric pooling, it can be that $\bigcap_{i \in N} supp(p_i) \subsetneq supp(p_j)$ and that on $supp(p_j) \setminus \bigcap_{i \in N} supp(p_i)$, j is not an aggregator. Thus, the statement that j must be an aggregator holds with the qualification that it applies only up to proportionality on $\bigcap_{i \in N} supp(p_i)$. Under linear pooling, by contrast, the statement must hold not only up to proportionality, but equality is required on $\bigcap_{i \in N} supp(p_i)$. This is the respect in which one can cautiously claim that, as far as supra-Bayesianism is concerned, geometric pooling conserves some comparative advantage over linear pooling even beyond $\Delta_{\propto}^n$.

6.2 Beyond the Full-Revelation Assumption?

We now temporarily revert to the common prior assumption to focus, instead, on the full revelation assumption. To investigate the Bayes-compatibility of geometric pooling, we have so far assumed that the n agents report their full posterior probability distributions. As we now explain, this particularly demanding informational requirement can be somewhat relaxed. Indeed, under some conditions, one can require that the n agents report only the restriction of their posterior probability distributions to a pair of events, A and B. This means that the agents report their posterior probability values for all states in A and B, but not for any other state. We call this requirement *the partial revelation assumption*. The partial revelation assumption is informationally less demanding than requiring that the entire posterior distributions be known, as in the full-revelation Bayes-compatibility problem. But it bears emphasis that it is still more demanding than requiring that only the posterior probability value for some event be known, as in the baseline version of the Bayes-compatibility problem.

The informational basis thus provided proves tractable enough if, *at the same time*, the definition of Bayes-compatibility is generalized to target not absolute, but relative probability values—i.e., ratios of the form $p(A)/p(B)$, rather than single values like $p(A)$ or $p(B)$. Put differently, one moves from $p(A)/p(B)$ with B the total event Ω, to $p(A)/p(B)$ with B an arbitrary event such that $p(B) > 0$. This is a natural generalization to consider because in many epistemic or practical situations, the question of interest is not what respective probability values to attribute to A and B, but only whether $p(A) \geq p(B)$, i.e. (assuming that $p(B) > 0$), whether $p(A)/p(B) \geq 1$. Examples include determining which of two hypothesis

---

[30]More specifically, let m be the number of agents and n be the number of states in $\bigcap_{i \in N} supp(p_i)$. Consider $\boldsymbol{A}$, the $m \times n$ real matrix constituted by the restriction of the m posteriors to the n states in $\bigcap_{i \in N} supp(p_i)$. If $n \geq m$, it is generically the case, in the sense explained next, that the restrictions of the m posteriors to the n states in $\bigcap_{i \in N} supp(p_i)$ are linearly independent. Consider all the values which the m posteriors can conjointly take over the n states. Finding one set of values such that the m posteriors are *not* linearly independent over the n states is akin to drawing one particular number from a real interval. Similar reasoning (based on the observation that any entry of $\boldsymbol{A}$ is strictly positive, so that its logarithm exists) shows that, generically, the restriction of the m posteriors to the n states in $\bigcap_{i \in N} supp(p_i)$ are log-linearly independent. Notice however that, when $n < m$, there can be only $n < m$ linearly independent restrictions of the m posteriors to the n states in $\bigcap_{i \in N} supp(p_i)$. Similarly, there can then be only $n < m$ log-linearly independent restrictions of the m posteriors to those n states.

is most likely to be true, deciding on which of two events to bet, and any similar context where qualitative probability assessments may suffice. Crucially for our purposes, under geometric pooling, such situations spare computing the normalizing factor in the denominator of Equation (2).

This leads us to introduce the following variant of the Bayes-compatibility problem, which we call *the partial-revelation Bayes-compatibility problem*. Consider a profile $(p|E_1, \ldots, p|E_n)$ in Δ^n_{CP}. Assume that through public announcements, the agents learn the restrictions of the n posteriors to a given pair of events A and B. Thus, they learn which states in A and B are in the intersection of the supports of the n posteriors. More specifically, under the assumptions in Section 4.1, they learn exactly the event $E(A, B)$, i.e., the complement of the set of states in A and B that have been excluded by the public announcements:

$$E(A, B) = \left(\left(A \setminus \bigcap_{i \in N} E_i \right) \bigcup \left(B \setminus \bigcap_{i \in N} E_i \right) \right)^c. \tag{15}$$

Recall that each agent j also has access to her private information E_j. Such private information is in general not fully disclosed by her revelation of the restriction of her posterior distribution p_j to A and B. Therefore, the total evidence of any agent j is given by the more specific event $E(A, B) \cap E_j$. Accordingly, the relevant revised version of the Bayes-compatibility problem consists in knowing whether there exists any pooling function F with the following property. Given $A, B \in \mathcal{A}$ with B such that $B \cap \left(\bigcap_{i \in N} E_i \right) \neq \emptyset$, for any $j \in N$, the following equality holds:

$$\frac{F(p|E_1, \ldots, p|E_n)(A)}{F(p|E_1, \ldots, p|E_n)(B)} = \frac{p(A|E(A, B) \cap E_j)}{p(B|E(A, B) \cap E_j)}. \tag{16}$$

For future reference, we immediately observe that the equality in (16) can be simplified in virtue of the following equality:[31]

$$\frac{p(A|E(A, B) \cap E_j)}{p(B|E(A, B) \cap E_j)} = \frac{p\left(A| \bigcap_{i \in N} E_i\right)}{p\left(B| \bigcap_{i \in N} E_i\right)}. \tag{17}$$

This clarifies by way of anticipation that, in (16), the expression on the right-hand side of the equality does not really depend on the particular $j \in N$ considered, since the probability ratio in question takes the same value given any individual's total evidence, or the group's total evidence. As we will detail shortly, (17) has more specific epistemic implications for how informationally demanding geometric pooling is.

Indeed, recall the second question asked at the beginning of Section 6: Can a supra-Bayesian case for geometric pooling still be made under weaker informational demands than the full-revelation assumption? Equality (17) proves instrumental in providing an answer to this question. Under the previously explained qualification that Bayes-compatibility is generalized to target not absolute, but relative probability values, and requiring such Bayes-compatibility for all pairs of events, the following proposition constitutes a positive answer to this question.

---

[31]This equality is established within the proof of Proposition 7.

Proposition 7. *Let F denote any pooling function on Δ_{CP}^n. The following are equivalent:*

1. *For any profile $(p|E_1, \ldots, p|E_n)$ in Δ_{CP}^n, $\frac{F(p|E_1,\ldots,p|E_n)(A)}{F(p|E_1,\ldots,p|E_n)(B)} = \frac{p(A|E(A,B)\cap E_j)}{p(B|E(A,B)\cap E_j)}$ holds for all $j \in N$, $A \in \mathcal{A}$, and $B \in \mathcal{A}$ such that $B \cap \left(\bigcap_{i \in N} E_i\right) \neq \emptyset$, with $E(A,B)$ the event defined in (15).*

2. *For any profile $(p|E_1, \ldots, p|E_n)$ in Δ_{CP}^n, any weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \neq 0$ for all $i \in N$ and $\sum_{i=1}^n \alpha_i = 1$, with c induced as in (3), for all $\omega \in \Omega$: $F(p|E_1, \ldots, p|E_n)(\omega) = c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i}$.*

Thus, over the common prior domain Δ_{CP}^n, requiring that a pooling function be partial-revelation Bayes-compatible whatever the pair of events is equivalent to requiring it to be geometric with non-zero normalized weights.[32] We stress that the equivalence does not require that the agents reveal their posterior ratios for all pair of events. What it requires is that the partial Bayes-compatibility condition holds whatever the pair of events under consideration. Put differently, what the proposition establishes is that geometric pooling is the only format that is *guaranteed* to achieve, whatever the pair of events, partial-revelation Bayes-compatibility. Thus, the motivation for using geometric pooling as a Bayesian recipe for social learning does not depend on the full revelation assumption.

Subject to the qualification just explained regarding that with respect to which Bayes-compatibility should be appreciated, the next proposition shows that weakening full-revelation to partial-revelation does not open any new possibility for linear pooling.

Proposition 8. *Let $(p|E_1, \ldots, p|E_n)$ be any profile in Δ_{CP}^n. For any $A \in \mathcal{A}$ and any $B \in \mathcal{A}$ such that $B \cap \left(\bigcap_{i \in N} E_i\right) \neq \emptyset$, let $E(A,B)$ denote the event defined in (15). Let $\{\alpha_i\}_{i \in N}$ be any weights such that $\alpha_i \in (0,1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$. The following are equivalent:*

1. *$\frac{\sum_{i=1}^n \alpha_i p(A|E_i)}{\sum_{i=1}^n \alpha_i p(B|E_i)} = \frac{p(A|E(A,B)\cap E_j)}{p(B|E(A,B)\cap E_j)}$ holds for all $j \in N$, $A \in \mathcal{A}$, and all $B \in \mathcal{A}$ meeting the condition above.*

2. *$E_k = E_l$ for all $k, l \in N$.*

The following example immediately clarifies the content of Proposition 8 by showing that although the equality in (16) cannot non-trivially hold *for all* pairs $A, B \in \mathcal{A}$ under linear pooling, it can nonetheless non-trivially hold *for some* such pairs.[33]

Example 3. *Let $\Omega = \{\omega_1, \omega_2, \omega_2, \omega_4, \omega_5\}$ and let p be the uniform distribution on Ω. Let p_1 and p_2 be the probability measures obtained from p by Bayesian conditioning on $E_1 = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $E_2 = \{\omega_2, \omega_3, \omega_5\}$, respectively. Then, $\bigcap_{i \in N} E_i = \{\omega_2, \omega_3\}$ and the profile (p_1, p_2) is given in the table below.*

---

[32]Thus, it also follows from Proposition 7 and Theorem 1 that on Δ_{CP}^n, partial-revelation Bayes-compatibility for a pooling function is equivalent to its respecting Unanimity and Private Bayesianism.

[33]We do not yet know the general conditions under which the equality in (16) holds for some $A, B \in \mathcal{A}$ under linear pooling.

	ω_1	ω_2	ω_3	ω_4	ω_5
p_1	1/4	1/4	1/4	1/4	0
p_2	0	1/3	1/3	0	1/3

Consider the events $A = \{\omega_1, \omega_2\}$ and $B = \{\omega_3, \omega_4\}$. With weights $\alpha_1 = \alpha_2 = 1/2$, although $\sum_{i=1}^2 \alpha_i p(A|E_i) = \sum_{i=1}^2 \alpha_i p(B|E_i) = \frac{5}{12} \neq \frac{1}{2} = p\left(A | \bigcap_{i \in N} E_i\right) = p\left(B | \bigcap_{i \in N} E_i\right)$, we have that $\frac{\sum_{i=1}^2 \alpha_i p(A|E_i)}{\sum_{i=1}^2 \alpha_i p(B|E_i)} = \frac{p\left(A | \bigcap_{i \in N} E_i\right)}{p\left(B | \bigcap_{i \in N} E_i\right)} = 1$.

Together with Example 3, Propositions 7 and 8 call for two series of comments. First, as can be seen by bringing together Example 3 and the preexisting literature discussed in Section 3 (especially Bradley, 2018), linear pooling performs better when the Bayes-compatibility of probability ratios is examined under the partial-revelation assumption, than when the Bayes-compatibility of single probability values is examined under the full-revelation assumption (notice the two simultaneous changes). Naturally, this has to be appreciated in light of the fact that the former problem involves weaker mathematical constraints than the latter. Furthermore, this is of interest only provided one finds worth considering a form of Bayes-compatibility that would hold for some, but not all events. Be that as it may, what Propositions 7 and 8 show is that, unlike geometric pooling, linear pooling cannot nontrivially deliver the desired compatibility whatever the probability ratio. To this extent, we take these propositions overall to confirm, rather than challenge, the Bayesian advantage of geometric pooling over linear pooling in a common prior setting.

Second, Proposition 7 illustrates a useful shortcut for geometric pooling. As (17) makes explicit, when only probability ratios are at stake, learning the evidence of the group restricted to the events of interest leads to the same result as learning all the evidence available to the group. This allows less social information to be elicited for geometric pooling, while preserving its supra-Bayesian qualities in a large class of epistemic and practical situations.

This fact notwithstanding, first and foremost, we take the present section to highlight the strong informational demands of geometric pooling. Indeed, even when only probability ratios of the form $p(A)/p(B)$ are at stake, geometric pooling requires knowing the individual probability values for each state in A and B, and not just the individual aggregate values for A and B. This requirement extends to the whole state space when single probability values, like $p(A)$ or $p(B)$, are at stake. Either way, then, there is a tension with one standard motivation for pooling as a heuristic—or a boundedly rational mechanism, or a shortcut rule—for learning from others, namely, informational parsimony. Linear pooling can claim such parsimony (Dawid et al., 1995; Golub and Jackson, 2010), but at the expense of Bayes-compatibility. In a nutshell, then, geometric pooling illustrates the converse trade-off.

6.3 Beyond Both the Common Prior and the Full-Revelation Assumptions?

Our last observation will be that the qualified generalizations of Theorem 1 presented separately in the two previous subsections could also be carried out—though even more cautiously still—simultaneously. Consider $\{p_i\}_{i \in N}$, a profile in $\Delta_\propto^n$, and not necessarily Δ_{CP}^n. Assume that through public announcements, the agents learn not the entire supports of

the n measures, but only their restrictions to a given pair of events, A and B. Accordingly, let $S(A, B)$—with S for "support"—denote the following event:

$$S(A, B) = \left(\left(A \setminus \bigcap_{i \in N} supp(p_i) \right) \bigcup \left(B \setminus \bigcap_{i \in N} supp(p_i) \right) \right)^c. \tag{18}$$

Under additional qualifications similar to the ones already discussed in subsection 6.1, the more specific event $S(A, B) \cap supp(p_i)$ can then be considered the total evidence of agent i, for any $i \in N$. Furthermore, in this context, like in subsection 6.2, the Bayes-compatibility of a pooling function can be investigated with respect not to absolute, but to relative probability values.

Under these assumptions, the following proposition establishes that the supra-Bayesian case for geometric pooling generalizes beyond both the common prior and the full revelation settings.

Proposition 9. *Let F be a pooling function on $\Delta_\propto^n$. The following are equivalent:*

1. *For any profile $\{p_i\}_{i \in N}$ in $\Delta_\propto^n$, $\frac{F(p_1,\dots,p_n)(A)}{F(p_1,\dots,p_n)(B)} = \frac{p_i(A|S(A,B) \cap supp(p_i))}{p_i(B|S(A,B) \cap supp(p_i))}$ holds for all $i \in N$, $A \in \mathcal{A}$, and $B \in \mathcal{A}$ such that $B \cap \left(\bigcap_{i \in N} supp(p_i) \right) \neq \emptyset$, with $S(A, B)$ the event defined in (18).*

2. *For any profile $\{p_i\}_{i \in N}$ in $\Delta_\propto^n$, any weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \neq 0$ for all $i \in N$ and $\sum_{i=1}^n \alpha_i = 1$, with c induced as in (3), for all $\omega \in \Omega$: $F(p_1, \dots, p_n)(\omega) = c \prod_{i=1}^n p_i(\omega)^{\alpha_i}$.*

Similarly, the next and final proposition states that the Bayes-incompatibility of linear pooling simultaneously generalizes beyond both the common prior and the full revelation settings.

Proposition 10. *Let $\{p_i\}_{i \in N}$ be a profile in $\Delta_\propto^n$. With $A \in \mathcal{A}$ and $B \in \mathcal{A}$ always such that $B \cap \left(\bigcap_{i \in N} supp(p_i) \right) \neq \emptyset$, let $S(A, B)$ denote the event defined in (18). Let $\{\alpha_i\}_{i \in N}$ be any weights such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$. The following are equivalent:*

1. *$\frac{\sum_{i=1}^n \alpha_i p_i(A)}{\sum_{i=1}^n \alpha_i p_i(B)} = \frac{p_i(A|S(A,B) \cap supp(p_i))}{p_i(B|S(A,B) \cap supp(p_i))}$ holds for all $i \in N$, $A \in \mathcal{A}$, $B \in \mathcal{A}$ meeting the condition above.*

2. *For all $j, k \in N$, $supp(p_j) = supp(p_k)$.*

The proofs of Propositions 9 and 10 are analogous to those for Propositions 7 and 8, respectively; so we omit them altogether.

We close by stressing what we mean to be the main take-away of the present section, namely that the key positive and negative results of our paper are robust to important variations in the statement of the Bayes-compatibility problem.

7 Conclusion

The literature investigating the Bayes-compatibility of pooling functions in a common prior setting has so far focused on certain impossibilities that obtain under linear pooling. By contrast, we have established a possibility result to the effect that geometric pooling can be Bayes-compatible in such a setting. Indeed, our main finding is the stronger result that, under certain assumptions, geometric and Bayes-compatible pooling are equivalent. This characterization involves a very basic Bayes-compatibility requirement which, setting aside trivial cases, linear pooling systematically fails to satisfy. We have also shown that our central characterization extends somewhat beyond the common prior setting. As a result, in a distinguished class of epistemic contexts, geometric pooling enjoys a normative advantage over linear pooling as a recipe for social learning. But upon closing, we wish to stress again that this advantage should be appreciated cautiously. First, our study suggests that there is a general trade-off between the informational parsimony and the Bayes-compatibility of a pooling function. Linear pooling can lay claim to such parsimony, but at the expense of Bayes-compatibility; geometric pooling essentially illustrates the converse trade-off. The cost of Bayes-compatibility, then, should not be downplayed. Second, and more generally, there is simply no one-size-fits-all format of the Bayes-compatibility problem. We have moved away from what is perhaps its most standard variant because it proves almost exclusively tailored to linear pooling. We have introduced a different variant of the problem befitting an inquiry into geometric pooling. This format is of interest beyond geometric pooling, as we have indicated; in particular, it also sheds light on the structural properties of linear pooling. But stepping back, first and foremost, what our study highlights is that there are in fact several, non-trivially related Bayes-compatibility properties worth considering. Under those two caveats, we can conclude that the consideration of Bayes-compatibility provides some support for geometric pooling as a social learning mechanism.

Appendix

Proof of Proposition 1

Proof. Let $(p|E_1, \ldots, p|E_n)$ be a profile in Δ_{CP}^n and let $\{\alpha_i\}_{i \in N}$ be weights such that $\alpha_i \in (0,1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$.

($\Rightarrow$) We prove the contrapositive. Suppose that $E_j \neq E_k$ for some $j, k \in N$. By the common prior assumption, it then follows that there is some event $A \neq \emptyset$ such that $A \subseteq E_j$ but $A \cap E_k = \emptyset$ or vice versa. Without loss of generality, assume the former. Since $\alpha_j, \alpha_k > 0$, it follows that $\sum_{i=1}^n \alpha_i p(A|E_i) > 0$. However, since $A \cap \left(\bigcap_{i \in N} E_i \right) = \emptyset$, we have that $p(A | \bigcap_{i \in N} E_i) = 0$, so that $\sum_{i=1}^n \alpha_i p(A|E_i) \neq p(A | \bigcap_{i \in N} E_i)$.

($\Leftarrow$) Trivial. $\qquad \square$

Proof of Theorem 1

Proof. ($1 \Leftrightarrow 2$) Let $(p|E_1, \ldots, p|E_n)$ be a profile in Δ_{CP}^n. Consider first any ω such that $\omega \notin \bigcap_{i \in N} E_i$. Then, it must be the case that $p(\omega|E_j) = 0$ for some $j \in N$. Since $\alpha_i \neq 0$, $c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i} = 0 = p(\omega | \bigcap_{i \in N} E_i)$. It then follows that $F(p|E_1, \ldots, p|E_n)(\omega) = c \prod_{i=1}^n p(\omega|E_i)$

if and only if $F(p|E_1, \ldots p|E_n)(\omega) = p(\omega|\bigcap_{i\in N} E_i)$. Consider next any ω such that $\omega \in \bigcap_{i\in N} E_i$. In this case, the following chain of equalities holds.

$$
\begin{aligned}
c\prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i} &= \frac{\prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i}}{\sum_{\omega'\in\Omega}\prod_{i=1}^{n} p(\omega'|E_i)^{\alpha_i}} \\
&= \frac{\prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i}}{\sum_{\omega'\in\bigcap_{i\in N} E_i}\prod_{i=1}^{n} p(\omega'|E_i)^{\alpha_i}} \\
&= \frac{\prod_{i=1}^{n} \left(\frac{p(\omega)}{p(E_i)}\right)^{\alpha_i}}{\sum_{\omega'\in\bigcap_{i\in N} E_i}\prod_{i=1}^{n} \left(\frac{p(\omega')}{p(E_i)}\right)^{\alpha_i}} \\
&= \frac{\left(\prod_{i=1}^{n} p(E_i)^{-\alpha_i}\right)\left(\prod_{i=1}^{n} p(\omega)^{\alpha_i}\right)}{\sum_{\omega'\in\bigcap_{i\in N} E_i}\left[\left(\prod_{i=1}^{n} p(E_i)^{-\alpha_i}\right)\left(\prod_{i=1}^{n} p(\omega')^{\alpha_i}\right)\right]} \\
&= \frac{\left(\prod_{i=1}^{n} p(E_i)^{-\alpha_i}\right)\left(\prod_{i=1}^{n} p(\omega)^{\alpha_i}\right)}{\left(\prod_{i=1}^{n} p(E_i)^{-\alpha_i}\right)\left(\sum_{\omega'\in\bigcap_{i\in N} E_i}\prod_{i=1}^{n} p(\omega')^{\alpha_i}\right)} \\
&= \frac{\prod_{i=1}^{n} p(\omega)^{\alpha_i}}{\sum_{\omega'\in\bigcap_{i\in N} E_i}\prod_{i=1}^{n} p(\omega')^{\alpha_i}} \\
&= \frac{p(\omega)^{\alpha_1+\cdots+\alpha_n}}{\sum_{\omega'\in\bigcap_{i\in N} E_i} p(\omega')^{\alpha_1+\cdots+\alpha_n}} \\
&= \frac{p(\omega)}{\sum_{\omega'\in\bigcap_{i\in N} E_i} p(\omega')} \\
&= \frac{p(\omega)}{p(\bigcap_{i\in N} E_i)} \\
&= p\left(\omega\,\middle|\,\bigcap_{i\in N} E_i\right).
\end{aligned}
$$

Hence, whether $\omega \in \bigcap_{i\in N} E_i$ or not, we have that, for any profile $(p|E_1, \ldots, p|E_n\}$ in Δ^n_{CP}, $F(p|E_1, \ldots, p|E_n)(\omega) = c\prod_{i=1}^{n} p(\omega|E_i)$ with weights $\{\alpha_i\}_{i\in N}$ such that $\alpha_i \neq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$ and c the constant defined in (3) if and only if $F(p|E_1, \ldots p|E_n)(\omega) = p(\omega|\bigcap_{i\in N} E_i)$. Given the scope of the universal quantifiers, the equivalence thus established is stronger than, therefore implies, the equivalence between 1. and 2. in Theorem 1, viz.: 1. For all profiles in Δ^n_{CP}, $F(p|E_1, \ldots p|E_n)(\omega) = p(\omega|\bigcap_{i\in N} E_i)$ for all $\omega \in \Omega$; 2. For all profiles in Δ^n_{CP}, $F(p|E_1, \ldots, p|E_n)(\omega) = c\prod_{i=1}^{n} p(\omega|E_i)$ for all $\omega \in \Omega$, any weights $\{\alpha_i\}_{i\in N}$ such that $\alpha_i \neq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$, with c the profile-dependent constant defined in (3).

$(1 \Leftarrow 3)$ Assume that F satisfies Unanimity and Private Bayesianism. Let $(p|E_1, \ldots, p|E_n)$ be a profile in Δ^n_{CP}. Dietrich shows (2017, Proposition 1) that Private Bayesianism implies Public Bayesianism.

Public Bayesianism. For any profile $(p_1, \ldots, p_n) \in \Delta^n$ and any (suitable) event $E \in \mathcal{A}$, $F(p_1|E, \ldots, p_n|E) = F(p_1, \ldots, p_n)|E$.

Alternating applications of Private Bayesianism and Public Bayesianism, we get

$$
\begin{aligned}
F(p(\cdot|E_1),\ldots,p(\cdot|E_n)) &= F(p,p(\cdot|E_2),\ldots,p(\cdot|E_n))|E_1 \\
&= F(p(\cdot|E_1),p(\cdot|E_1 \cap E_2),\ldots,p(\cdot|E_1 \cap E_n)) \\
&= F(p(\cdot|E_1),p(\cdot|E_1),p(\cdot|E_1 \cap E_3),\ldots,p(\cdot|E_1 \cap E_n))|E_2 \\
&= F(p(\cdot|E_1 \cap E_2),p(\cdot|E_1 \cap E_2),p(\cdot|E_1 \cap E_2 \cap E_3),\ldots,p(\cdot|E_1 \cap E_2 \cap E_n)) \\
&\vdots \\
&= F\left(p\left(\cdot \,\middle|\, \bigcap_{i \in N} E_i\right),\ldots,p\left(\cdot \,\middle|\, \bigcap_{i \in N} E_i\right)\right).
\end{aligned}
$$

Now, by Unanimity,

$$
F\left(p\left(\cdot \,\middle|\, \bigcap_{i \in N} E_i\right),\ldots,p\left(\cdot \,\middle|\, \bigcap_{i \in N} E_i\right)\right) = p\left(\cdot \,\middle|\, \bigcap_{i \in N} E_i\right),
$$

which completes the argument for the claim.

$(1 \Rightarrow 3)$ Assume—henceforth: "the main assumption"—that $F(p|E_1,\ldots,p|E_n) = p(\cdot|\bigcap_{i \in N} E_i)$ for all profiles in Δ^n_{CP}.

First, let $(p|E_1,\ldots,p|E_n)$ be a profile such that $E_i = E_j$ for all $i,j \in N$. Then, $E_j = \bigcap_{i \in N} E_i$ for all $j \in N$. We then have that $p|E_j = p(\cdot|\bigcap_{i \in N} E_i)$ for all $j \in N$. Accordingly, by the main assumption, Unanimity is satisfied.

Next, consider a profile $(p_1,\ldots,p_n)$ in Δ^n_{CP}. By the main assumption, $F(p_1,\ldots,p_n) = p(\cdot|\bigcap_{i \in N} E_i)$. Next, consider the profile $(p_1,\ldots,p_i|E,\ldots p_n)$. Again by the main assumption, $F(p_1,\ldots,p_i|E,\ldots p_n) = p(\cdot|E \cap \bigcap_{i \in N} E_i)$. By the definition of Bayesian conditioning, we then have that $F(p_1,\ldots,p_i|E,\ldots p_n) = F(p_1,\ldots,p_n)|E$, so that Private Bayesianism is satisfied. $\qquad\square$

Proof of Proposition 2

Proof. Let $(p|E_1,\ldots,p|E_n)$ be a profile in Δ^n_{CP}, and let $\{\alpha_i\}_{i \in N}$ be weights such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^n \alpha_i = 1$.

$(\Rightarrow)$ We show the contrapositive. Suppose that for some $j \in N$, $E_j = \bigcap_{i \in N} E_i$. By Theorem 1, it follows that $c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i} = p(\omega|\bigcap_{i \in N} E_i) = p(\omega|E_j)$ for all $\omega \in \Omega$. Then, for all $\omega \in \Omega$,

$$
\min_{i \in N} p(\omega|E_i) \leq p(\omega|E_j) = c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i} \leq \max_{i \in N} p(\omega|E_i).
$$

This rules out synergy, which establishes the claim.

$(\Leftarrow)$ Suppose that $\bigcap_{i \in N} E_i \subsetneq E_j$ for all $j \in N$. Since $\alpha_i \neq 0$ for all $i \in N$,

$$
\sum_{\omega \in \bigcap_{i \in N} E_i} c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i} = 1.
$$

But since for every $j \in N$ there exists a $\omega \in E_j \setminus \bigcap_{i \in N} E_i$, it follows that $p_j(\bigcap_{i \in N} E_i) < 1$ for all $j \in N$, which establishes synergy. $\square$

Proof of Proposition 3

Proof. Let $\{p_i\}_{i \in N}$ be a profile in $\Delta^{n'}$ and let $\{\alpha_i\}_{i \in N}$ be such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^n \alpha_i = 1$.

$(2 \Rightarrow 1)$ Suppose that $p_j \propto p_k$ on $\bigcap_{i \in N} supp(p_i)$ for all $j, k \in N$. It then follows that $p_j | \bigcap_{i \in N} supp(p_i) = p_k | \bigcap_{i \in N} supp(p_i)$ for all $j, k \in N$. If $\omega \notin \bigcap_{i \in N} supp(p_i)$, then $c \prod_{i=1}^n p_i(\omega)^{\alpha_i} = 0 = p_i(\omega | \bigcap_{i \in N} supp(p_i))$ for all $i \in N$ when $\alpha_i \neq 0$ for all $i \in N$. If $\omega \in \bigcap_{i \in N} supp(p_i)$, we have the following.

$$
c \prod_{i=1}^n p_i(\omega)^{\alpha_i} = \frac{\prod_{i=1}^n p_i(\omega)^{\alpha_i}}{\sum_{\omega' \in \Omega} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}
$$

$$
= \frac{\prod_{i=1}^n p_i(\omega)^{\alpha_i}}{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}
$$

Multiplying by 1 expressed as $\frac{\frac{1}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}$, we obtain the following.

$$
\frac{\prod_{i=1}^n p_i(\omega)^{\alpha_i}}{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}} = \frac{\frac{\prod_{i=1}^n p_i(\omega)^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}{\frac{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}
$$

$$
= \frac{\prod_{i=1}^n \left(\frac{p_i(\omega)}{p_i(\bigcap_{i \in N} supp(p_i))} \right)^{\alpha_i}}{\frac{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}.
$$

Next, it follows from the proportionality assumption that $\frac{p_j(\omega)}{p_j(\bigcap_{i \in N} supp(p_i))} = \frac{p_k(\omega)}{p_k(\bigcap_{i \in N} supp(p_i))}$ for all $j, k \in N$ and any $\omega \in \bigcap_{i \in N} supp(p_i)$. So fix $p^* \in \{p_i\}_{i \in N}$, and observe the following.

$$
\frac{\prod_{i=1}^n \left(\frac{p_i(\omega)}{p_i(\bigcap_{i \in N} supp(p_i))} \right)^{\alpha_i}}{\frac{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}} = \frac{\prod_{i=1}^n \left(\frac{p^*(\omega)}{p^*(\bigcap_{i \in N} supp(p_i))} \right)^{\alpha_i}}{\frac{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}
$$

$$
= \frac{\left(\frac{p^*(\omega)}{p^*(\bigcap_{i \in N} supp(p_i))} \right)^{\alpha_1 + \cdots + \alpha_n}}{\frac{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}
$$

$$
= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i \in N} supp(p_i))}}{\frac{\sum_{\omega' \in \bigcap_{i \in N} supp(p_i)} \prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i \in N} supp(p_i))^{\alpha_i}}}
$$

32

$$= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}}{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} \frac{\prod_{i=1}^n p_i(\omega')^{\alpha_i}}{\prod_{i=1}^n p_i(\bigcap_{i\in N} supp(p_i))^{\alpha_i}}}$$

$$= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}}{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} \prod_{i=1}^n \left(\frac{p_i(\omega')}{p_i(\bigcap_{i\in N} supp(p_i))}\right)^{\alpha_i}}$$

$$= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}}{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} \prod_{i=1}^n \left(\frac{p^*(\omega')}{p^*(\bigcap_{i\in N} supp(p_i))}\right)^{\alpha_i}}$$

$$= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}}{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} \left(\frac{p^*(\omega')}{p^*(\bigcap_{i\in N} supp(p_i))}\right)^{\alpha_1+\cdots+\alpha_n}}$$

$$= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}}{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} \frac{p^*(\omega')}{p^*(\bigcap_{i\in N} supp(p_i))}}$$

$$= \frac{\frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}}{\frac{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} p^*(\omega')}{p^*(\bigcap_{i\in N} supp(p_i))}}$$

$$= \frac{p^*(\omega)}{\sum_{\omega'\in\bigcap_{i\in N} supp(p_i)} p^*(\omega')}$$

$$= \frac{p^*(\omega)}{p^*(\bigcap_{i\in N} supp(p_i))}$$

$$= p^*\left(\omega \,\middle|\, \bigcap_{i\in N} supp(p_i)\right)$$

Hence, for any $i \in N$, $c \prod_{i=1}^n p_i(\omega)^{\alpha_i} = p_i\left(\omega | \bigcap_{i\in N} supp(p_i)\right)$.

$(1 \Rightarrow 2)$. Assume that $c \prod_{i=1}^n p_i(\omega)^{\alpha_i} = p_i(\omega | \bigcap_{i\in N} supp(p_i))$ for all $\omega \in \Omega$ and any $i \in N$. It is then immediate that for all $\omega, \omega' \in \bigcap_{i\in N} supp(p_i)$ and any $i \in N$,

$$\frac{p_i(\omega)}{p_i(\omega')} = \frac{p_i(\omega | \bigcap_{i\in N} supp(p_i))}{p_i(\omega' | \bigcap_{i\in N} supp(p_i))} = \frac{c \prod_{i=1}^n p_i(\omega)^{\alpha_i}}{c \prod_{i=1}^n p_i(\omega')^{\alpha_i}}.$$

It follows that $p_j \propto p_k$ on $\bigcap_{i\in N} supp(p_i)$ for all $j, k \in N$.

$\square$

Proof of Proposition 4

Proof. $(1 \Leftrightarrow 2)$ Let $\{p_i\}_{i\in N}$ be a profile in Δ_∞^n. By Proposition 3, we have

$$c \prod_{i=1}^n p_i(\omega)^{\alpha_i} = p_i\left(\omega \,\middle|\, \bigcap_{i\in N} supp(p_i)\right) \tag{19}$$

for all $i \in N$ and any weights $\{\alpha_i\}_{i\in N}$ such that $\sum_{i=1}^n \alpha_i = 1$ and $\alpha_i \neq 0$ for all $i \in N$. Then, $F(p_1,\ldots,p_n)$ is equal to the lefthand side of (19) for non-zero weights summing to 1 if and

only if $F(p_1, \ldots, p_n)$ is equal to the righthand side of (19) for all $i \in N$. As in the proof of the equivalence of 1. and 2. in Theorem 1, the equivalence of 1. and 2. in Proposition 4 now follows from the stronger equivalence just established between Bayes-compatibility and geometric pooling for any profile in $\Delta_\propto^n$.

$(1 \Leftarrow 3)$ Assume that F satisfies Unanimity and Private Bayesianism. Let $\{p_i\}_{i \in N}$ be a profile in $\Delta_\propto^n$. Noting that $p_i = p_i | supp(p_i)$ and alternating applications of Private Bayesianism and Public Bayesianism (see p. 30), we get the following:

$$
\begin{aligned}
F(p_1, \ldots, p_n) &= F(p_1 | supp(p_1), \ldots, p_n | supp(p_n)) \\
&= F(p_1, p_2 | supp(p_2), \ldots, p_n | supp(p_n)) | supp(p_1) \\
&= F(p_1 | supp(p_1), p_2 | \cap_{i=1,2} supp(p_i), \ldots, p_n | \cap_{i=1,n} supp(p_i)) \\
&= F(p_1 | supp(p_1), p_2 | supp(p_1), p_3 | \cap_{i=1,3} supp(p_i), \ldots, p_n | \cap_{i=1,n} supp(p_i)) | supp(p_2) \\
&= F(p_1 | \cap_{i=1,2} supp(p_i), p_2 | \cap_{i=1,2} supp(p_i), p_3 | \cap_{i=1,2,3} supp(p_i), \ldots, p_n | \cap_{i=1,2,n} supp(p_i)) \\
&\quad \vdots \\
&= F\left(p_1 \bigg| \bigcap_{i \in N} supp(p_i), \ldots, p_n \bigg| \bigcap_{i \in N} supp(p_i)\right).
\end{aligned}
$$

Since $p_j | \bigcap_{i \in N} supp(p_i) \propto p_k | \bigcap_{i \in N} supp(p_i)$ by the domain restriction to $\Delta_\propto^n$ and given that $supp(p_j | \bigcap_{i \in N} supp(p_i)) = supp(p_k | \bigcap_{i \in N} supp(p_i))$ for all $j, k \in N$, it must be the case that $p_j | \bigcap_{i \in n} supp(p_i) = p_k | \bigcap_{i \in N} supp(p_i)$ for all $j, k \in N$. Thus, invoking Unanimity, for all $i \in N$,

$$
F\left(p_1 \bigg| \bigcap_{i \in N} supp(p_i), \ldots, p_n \bigg| \bigcap_{i \in N} supp(p_i)\right) = p_i \bigg| \bigcap_{i \in N} supp(p_i),
$$

which establishes the claim.

$(1 \Rightarrow 3)$ Assume that $F(p_1, \ldots, p_n) = p_i(\cdot | \bigcap_{i \in N} supp(p_i))$ for all $i \in N$ and all profiles in the domain of F.

First, let $(p_1, \ldots, p_n)$ be a profile such that $p_j = p_k$ for all $j, k \in N$. Any such profile is in $\Delta_\propto^n$. Clearly, $p_j = p_j | \bigcap_{i \in N} supp(p_j) = p_k | \bigcap_{i \in N} supp(p_i) = p_k$ for all $i \in N$. Thus, by (1), we can infer that $F(p_1, \ldots, p_n) = p_i | \bigcap_{i \in N} supp(p_i) = p_i$ for all $i \in N$. So, Unanimity is satisfied.

Next, let $(p_1, \ldots, p_n)$ be a profile in $\Delta_\propto^n$. Consider the profile $(p_1, \ldots, p_j | E, \ldots p_n)$. Since $\left(\bigcap_{i \in N} supp(p_i) \cap E\right) \subseteq \bigcap_{i \in N} supp(p_i)$ and $p_j \propto p_k$ on $\bigcap_{i \in N} supp(p_i)$ for all $j, k \in N$, we have that $p_j \propto p_k$ on $\bigcap_{i \in N} supp(p_i) \cap E$ for all $j, k \in N$, where $\bigcap_{i \in N} supp(p_i) \cap E$ is the intersection of the supports of the measures in the profile $(p_1, \ldots, p_j | E, \ldots, p_n)$. Hence, $(p_1, \ldots, p_j | E, \ldots, p_n) \in \Delta_\propto^n$. Moreover, $p_j | \bigcap_{i \in N} supp(p_i) \cap E = p_k | \bigcap_{i \in N} supp(p_i) \cap E$ for all $j, k \in N$. Thus, by a double application of (1), $F(p_1, \ldots, p_j | E, \ldots, p_n) = p_i | \bigcap_{i \in N} supp(p_i) \cap E = F(p_1 | E, \ldots, p_n | E)$. So, Private Bayesianism is satisfied. $\qquad\square$

Proof of Proposition 5

Proof. Let $\{p_i\}_{i \in N}$ be a profile in $\Delta_{\propto}^n$ and let $\{\alpha_i\}_{i \in N}$ be weights such that $\alpha_i \in (0,1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$.

$(1 \Rightarrow 2)$ Assume that $\sum_{i=1}^n \alpha_i p_i(A) = p_i(A | \bigcap_{i \in N} supp(p_i))$ for all $i \in N$ and any $A \in \mathcal{A}$. We want to show that $p_j = p_k$ for all $j, k \in N$. Since $p_i(A | \bigcap_{i \in N} supp(p_i)) = 0$ for any A such that $A \cap (\bigcap_{i \in N} supp(p_i)) = \emptyset$, by the above assumption, it must be the case that $p_i(A) = 0$ for all i and any A disjoint from the intersection of the supports. Hence, $supp(p_i) \subseteq \bigcap_{i \in N} supp(p_i)$ for all $i \in N$. And clearly, $\bigcap_{i \in N} supp(p_i) \subseteq supp(p_i)$ for all $i \in N$. So, $supp(p_i) = \bigcap_{i \in N} supp(p_i)$ for all $i \in N$, so that $supp(p_j) = supp(p_k)$ for all $j, k \in N$. But since $p_j \propto p_k$ on $\bigcap_{i \in N} supp(p_i)$, it follows that $p_j = p_k$ for all $j, k \in N$.

$(2 \Leftarrow 1)$ Trivial. $\qquad\qquad\square$

Proof of Proposition 6

Proof. Let $\{p_i\}_{i \in N}$ be a profile in $\Delta^{n'}$ and let $\{\alpha_i\}_{i \in N}$ be weights such that $\alpha_i \in (0,1)$ for all i and $\sum_{i=1}^n \alpha_i = 1$.

1. $(\Rightarrow)$ Assume that $c \prod_{i=1}^n p_i^{\alpha_i} = p_j | \bigcap_{i \in N} supp(p_i)$ for some $j \in N$ and all $A \in \mathcal{A}$. For any $\omega, \omega' \in \bigcap_{i \in N} supp(p_i)$,

$$\frac{c \prod_{i=1}^n p_i(\omega)^{\alpha_i}}{c \prod_{i=1}^n p_i(\omega')^{\alpha_i}} = \frac{p_j(\omega | \bigcap_{i \in N} supp(p_i))}{p_j(\omega' | \bigcap_{i \in N} supp(p_i))} = \frac{p_j(\omega)}{p_j(\omega')}.$$

Hence, $p_j \propto c \prod_{i=1}^n p_i^{\alpha_i}$ on $\bigcap_{i \in N} supp(p_i)$.

$(\Leftarrow)$ Assume that $p_j \propto c \prod_{i=1}^n p_i^{\alpha_i}$ on $\bigcap_{i \in N} supp(p_i)$. Then, it is also the case that $p_j | \bigcap_{i \in N} supp(p_j) \propto c \prod_{i=1}^n p_i^{\alpha_i}$ since, for any $\omega, \omega' \in \bigcap_{i \in N} supp(p_i)$,

$$\frac{p_j(\omega)}{p_j(\omega')} = \frac{p_j(\omega | \bigcap_{i \in N} supp(p_i))}{p_j(\omega' | \bigcap_{i \in N} supp(p_i))} = \frac{c \prod_{i=1}^n p_i(\omega)^{\alpha_i}}{c \prod_{i=1}^n p_i(\omega')^{\alpha_i}}.$$

This implies that $supp(p_j | \bigcap_{i \in N} supp(p_i)) = supp(c \prod_{i=1}^n p_i^{\alpha_i}) = \bigcap_{i \in N} supp(p_i)$. Given that $p_j | \bigcap_{i \in N} supp(p_j)$ and $c \prod_{i=1}^n p_i^{\alpha_i}$ have the same support and are proportional to one another on this support, we have that $p_j | \bigcap_{i \in N} supp(p_i) = c \prod_{i=1}^n p_i^{\alpha_i}$.

2. $(\Rightarrow)$ Assume that $\sum_{i=1}^n \alpha_i p_i = p_j | \bigcap_{i \in N} supp(p_i)$ for some $j \in N$. Then, since $\alpha_i \in (0,1)$ for all $i \in N$,

$$\bigcup_{i \in N} supp(p_i) = supp\left(\sum_{i=1}^n \alpha_i p_i\right) = supp\left(p_j \middle| \bigcap_{i \in N} supp(p_i)\right) \subseteq \bigcap_{i \in N} supp(p_i).$$

From $\bigcup_{i \in N} supp(p_i) \subseteq \bigcap_{i \in N} supp(p_i)$, it follows that $supp(p_i) = supp(p_k)$ for all $i, k \in N$. As $supp(p_j) = \bigcap_{i \in N} supp(p_i)$, we also have that $p_j = p_j | \bigcap_{i \in N} supp(p_i)$. From the above assumption, it then follows that $\sum_{i=1}^n \alpha_i p_i = p_j$.

(Partial $\Leftarrow$) Let $\{p_i\}_{i \in N}$ be a profile in $\Delta_{\ll}^n$. Since $\Delta_{\ll}^n \subseteq \Delta^{n'}$, Proposition 6.2 holds for any profile in this smaller domain. For the partial converse, suppose that $p_j = \sum_{i=1}^n \alpha_i p_i$ for some $j \in N$. Since $supp(p_j) = \bigcap_{i \in N} supp(p_j)$, it follows that $p_j = p_j | \bigcap_{i \in N} supp(p_j)$. Thus, $p_j = \sum_{i=1}^n \alpha_i p_i$ implies that $p_j | \bigcap_{i \in N} supp(p_i) = \sum_{i=1}^n \alpha_i p_i$. $\qquad\square$

Proof of Proposition 7

Proof. $(2 \Rightarrow 1)$. Suppose that, for all profiles $(p|E_1, \ldots, p|E_n)$ in Δ_{CP}^n and any weights $\{\alpha_i\}_{i \in N}$ such that $\alpha_i \neq 0$ for all i and $\sum_{i=1}^n \alpha_i = 1$, $F(p|E_1, \ldots, p|E_n)(\omega) = c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i}$ for all $\omega \in \Omega$. From Theorem 1, we have that $c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i} = p(\omega| \bigcap_{i \in N} E_i)$. Hence,

$$\frac{\sum_{\omega \in A} c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i}}{\sum_{\omega \in B} c \prod_{i=1}^n p(\omega|E_i)^{\alpha_i}} = \frac{\sum_{\omega \in A} \prod_{i=1}^n p(\omega|E_i)^{\alpha_i}}{\sum_{\omega \in B} \prod_{i=1}^n p(\omega|E_i)^{\alpha_i}} = \frac{p(A| \bigcap_{i \in N} E_i)}{p(B| \bigcap_{i \in N} E_i)}. \tag{20}$$

We will now show that

$$\frac{p(A| \bigcap_{i \in N} E_i)}{p(B| \bigcap_{i \in N} E_i)} = \frac{p\left(A|E(A,B) \cap E_j\right)}{p\left(B|E(A,B) \cap E_j\right)} = \frac{p\left(A|E(A,B)\right)}{p\left(B|E(A,B)\right)}. \tag{21}$$

The conclusion will then follow from (20) and (21).

First, notice that, using the De Morgan's laws, we can simplify the conditioning event $E(A,B) = ((A \setminus \bigcap_{i \in N} E_i) \cup (B \setminus \bigcap_{i \in N} E_i))^c$ in (15) as follows:

$$\begin{aligned}
E(A,B) &= \left(\left(A \setminus \bigcap_{i \in N} E_i \right) \cup \left(B \setminus \bigcap_{i \in N} E_i \right) \right)^c \\
&= \left(A \setminus \bigcap_{i \in N} E_i \right)^c \cap \left(B \setminus \bigcap_{i \in N} E_i \right)^c \\
&= \left(A^c \cup \bigcap_{i \in N} E_i \right) \cap \left(B^c \cup \bigcap_{i \in N} E_i \right) \\
&= \bigcap_{i \in N} E_i \cup (A^c \cap B^c).
\end{aligned}$$

Similarly, for any $j \in N$,

$$\left(\bigcap_{i \in N} E_i \cup (A^c \cap B^c) \right) \cap E_j = \bigcap_{i \in N} E_i \cup (A^c \cap B^c \cap E_j).$$

Then, using the definition of conditional probability,

$$\begin{aligned}
\frac{p(A|E(A,B))}{p(B|E(A,B))} &= \frac{p(A| \bigcap_{i \in N} E_i \cup (A^c \cap B^c))}{p(B| \bigcap_{i \in N} E_i \cup (A^c \cap B^c))} \\
&= \frac{\frac{p(A \cap (\bigcap_{i \in N} E_i \cup (A^c \cap B^c)))}{p(\bigcap_{i \in N} E_i \cup (A^c \cap B^c))}}{\frac{p(B \cap (\bigcap_{i \in N} E_i \cup (A^c \cap B^c)))}{p(\bigcap_{i \in N} E_i \cup (A^c \cap B^c))}} \\
&= \frac{p(A \cap (\bigcap_{i \in N} E_i \cup (A^c \cap B^c)))}{p(B \cap (\bigcap_{i \in N} E_i \cup (A^c \cap B^c)))} \\
&= \frac{p(A \cap (\bigcap_{i \in N} E_i))}{p(B \cap (\bigcap_{i \in N} E_i))}.
\end{aligned}$$

$$\tag{22}$$

By analogous reasoning, for any $j \in N$,

$$\frac{p\left(A|E(A,B) \cap E_j\right)}{p\left(B|E(A,B) \cap E_j\right)} = \frac{p(A|\bigcap_{i \in N} E_i \cup (A^c \cap B^c \cap E_j))}{p(B|\bigcap_{i \in N} E_i \cup (A^c \cap B^c \cap E_j))} = \frac{p(A \cap (\bigcap_{i \in N} E_i))}{p(B \cap (\bigcap_{i \in N} E_i))}. \tag{23}$$

Now, multiplying $\frac{p(A \cap (\bigcap_{i \in N} E_i))}{p(B \cap (\bigcap_{i \in N} E_i))}$ by 1 expressed as $\frac{1/p(\bigcap_{i \in N} E_i)}{1/p(\bigcap_{i \in N} E_i)}$, we obtain

$$\frac{\frac{p(A \cap (\bigcap_{i \in N} E_i))}{p(\bigcap_{i \in N} E_i)}}{\frac{p(B \cap (\bigcap_{i \in N} E_i))}{p(\bigcap_{i \in N} E_i)}} = \frac{p(A|\bigcap_{i \in N} E_i)}{p(B|\bigcap_{i \in N} E_i)}. \tag{24}$$

Hence, (21) follows from (22), (23), and (24), and the conclusion follows from (20) and (21). Notice also that (17) follows from (21), as claimed in the main text.

$(1 \Rightarrow 2)$ Suppose that for all profiles $(p|E_1, \ldots, p|E_n)$ in Δ^n_{CP}, for all $j \in N$, for all $A \in \mathcal{A}$, and for all $B \in \mathcal{A}$ such that $B \cap (\bigcap_{i \in N} E_i) \neq \emptyset$,

$$\frac{F(p|E_1, \ldots, p|E_n)(A)}{F(p|E_1, \ldots, p|E_n)(B)} = \frac{p(A|E(A,B) \cap E_j)}{p(B|E(A,B) \cap E_j)}.$$

By (21),

$$\frac{F(p|E_1, \ldots, p|E_n)(A)}{F(p|E_1, \ldots, p|E_n)(B)} = \frac{p(A|E(A,B) \cap E_j)}{p(B|E(A,B) \cap E_j)} = \frac{p\left(A|\bigcap_{i \in N} E_i\right)}{p\left(B|\bigcap_{i \in N} E_i\right)}.$$

For any $\omega \in \Omega$, since $\omega \in \mathcal{A} = 2^\Omega$, we have that, with $B = \Omega$, the following holds:

$$\frac{F(p|E_1, \ldots, p|E_n)(\omega)}{F(p|E_1, \ldots, p|E_n)(\Omega)} = F(p|E_1, \ldots, p|E_n)(\omega) = p\left(\omega \middle| \bigcap_{i \in N} E_i\right).$$

Now, applying Theorem 1, we obtain

$$F(p|E_1, \ldots, p|E_n)(\omega) = p\left(\omega \middle| \bigcap_{i \in N} E_i\right) = c \prod_{i=1}^{n} p(\omega|E_i)^{\alpha_i},$$

for any $\{\alpha_i\}_{i \in N}$ s.t. $\alpha_i \neq 0$ and $\sum_{i=1}^{n} \alpha_i = 1$. The equality of the first and last terms in the above equation then gives the desired result. $\square$

Proof of Proposition 8

Proof. Let $(p|E_1, \ldots, p|E_n)$ be a profile in Δ^n_{CP} and let $\{\alpha_i\}_{i \in N}$ be weights such that $\alpha_i \in (0, 1)$ for all i and $\sum_{i=1}^{n} \alpha_i = 1$. Let $A \in \mathcal{A}$, and $B \in \mathcal{A}$ be such that $B \cap (\bigcap_{i \in N} E_i) \neq \emptyset$. Let $E(A, B)$ denote the event defined in (15).

$(1 \Rightarrow 2)$ We show the contrapositive. Suppose that $E_k \neq E_l$ for some $k, l \in N$. By the common prior assumption, there must be some event $A \neq \emptyset$ such that $A \subseteq E_k$ and $A \cap E_l = \emptyset$ or $A \subseteq E_l$ and $A \cap E_k = \emptyset$. Without loss of generality, suppose the former. Since $\alpha_k > 0$, it follows that $\sum_{i=1}^{n} \alpha_i p(A|E_i) > 0$. However, $p(A|E(A,B) \cap E_l) = 0$. Now, let B be any event in $\mathcal{A}$ such that $B \cap \bigcap_{i \in N} E_i \neq \emptyset$. We then have

$$\frac{\sum_{i=1}^{n} \alpha_i p(A|E_i)}{\sum_{i=1}^{n} \alpha_i p(B|E_i)} > \frac{p\left(A|E(A,B) \cap E_l\right)}{p\left(B|E(A,B) \cap E_l\right)} = 0,$$

37

which establishes the claim.

$(2 \Leftarrow 1)$ Trivial using (17).

$\square$

References

Aumann, R. (1976). Agreeing to Disagree. *The Annals of Statistics 4*(6), 1236–1239.

Baron, J., B. Mellers, P. Tetlock, E. Stone, and L. Ungar (2014). Two Reasons to Make Aggregated Probability Forecasts More Extreme. *Decision Analysis 11*(2), 133–145.

Bonanno, G. and K. Nehring (1997). Agreeing to Disagree: A Survey. *U.C.-Davis Department of Economics Working Paper No. 97-18*.

Bonnay, D. and M. Cozic (2018). Weighted Averaging, Jeffrey Conditioning and Invariance. *Theory and Decision 85*(1), 21–39.

Bonnay, D. and M. Cozic (2019). Weighted Averaging and Bayesian Conditioning. ms..

Bradley, R. (2006). Taking Advantage of Difference in Opinion. *Episteme 3*(3), 141–155.

Bradley, R. (2007). Reaching a Consensus. *Social Choice and Welfare 29*(4), 609–632.

Bradley, R. (2018). Learning from Others: Conditioning versus Averaging. *Theory and Decision 85*(1), 5–20.

Carnap, R. (1947). On the Application of Inductive Logic. *Philosophy and Phenomenological Research 8*(1), 133–148.

Christensen, D. (2009). Disagreement as Evidence: The Epistemology of Controversy. *Philosophy Compass 4*(5), 756–767.

Dawid, P., M. DeGroot, and J. Mortera (1995). Coherent Combination of Experts' Opinions. *Test 4*(2), 263–313.

Dawid, P. and J. Mortera (2020). Resolving Some Contradictions in the Theory of Linear Opinion Pools. *Theory and Decision 88*, 453–456.

DeGroot, M. (1974). Reaching a Consensus. *Journal of the American Statistical Association 69*(345), 118–121.

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

Dietrich, F. (2017). A Theory of Bayesian Groups. *Noûs*.

Dietrich, F. and C. List (2016). Probabilistic Opinion Pooling. In A. Hájek and C. Hitchcock (Eds.), *Oxford Handbook of Probability and Philosophy*, pp. 519–541. Oxford: Oxford University Press.

Easwaran, K., L. Fenton-Glynn, C. Hitchcock, and J. Velasco (2016). Updating on the Credences of Others: Disagreement, Agreement, and Synergy. *Philosophers' Imprint 16*(11), 1–39.

Elga, A. (2007). Reflection and Disagreement. *Noûs 41*(3), 478–502.

French, S. (1985). Group Consensus Probability Distributions: A Critical Survey. In J. Bernardo, M. DeGroot, D. Lindley, and A. Smith (Eds.), *Bayesian Statistics 2*, pp. 183–201. Amsterdam: North Holland.

French, S. (1986). Calibration and the Expert Problem. *Management Science 32*(3), 315–321.

Gaifman, H. (1988). A Theory of Higher Order Probabilities. In B. Skyrms and W. Harper (Eds.), *Causation, Chance and Credence*, pp. 191–219. New York: Springer.

Geanakoplos, J. and H. Polemarchakis (1982). We Can't Disagree Forever. *Journal of Economic Theory 28*(1), 192–200.

Genest, C. (1984). Pooling Operators with the Marginalization Property. *Canadian Journal of Statistics 12*(2), 153–163.

Genest, C., K. McConway, and M. Schervish (1986). Characterization of Externally Bayesian Pooling Operators. *The Annals of Statistics 14*(3), 487–501.

Genest, C. and M. Schervish (1985). Modeling Expert Judgments for Bayesian Updating. *The Annals of Statistics*, 1198–1212.

Genest, C. and J. Zidek (1986). Combining Probability Distributions: A Critique and an Annotated Bibliography. *Statistical Science 1*(1), 114–135.

Goldman, A. (2001). Experts: Which Ones Should You Trust? *Philosophy and phenomenological research 63*(1), 85–110.

Golub, B. and M. Jackson (2010). Naive Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics 2*(1), 112–49.

Golub, B. and E. Sadler (2016). Learning in Social Networks. In Y. Bramoull, A. Galeotti, and B. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 594–542. Oxford: Oxford University Press.

Good, I. (1967). On the Principle of Total Evidence. *The British Journal for the Philosophy of Science 17*(4), 319–321.

Jeffrey, R. (2004). *Subjective Probability: The Real Thing.* Cambridge: Cambridge University Press.

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

Levi, I. (1977). Direct Inference. *The Journal of Philosophy 74*(1), 5–29.

Lindley, D. (1982a). The Bayesian Approach to Statistics. In J. Tiago de Oliveira and B. Epstein (Eds.), *Some Recent Advances in Statistics*, pp. 65–87. New York: Academic Press.

Lindley, D. (1982b). The Improvement of Probability Judgements. *Journal of the Royal Statistical Society: Series A 145*(1), 117–126.

McConway, K. (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.

Morris, P. (1974). Decision Analysis Expert Use. *Management Science 20*(9), 1233–1241.

Morris, S. (1995). The Common Prior Assumption in Economic Theory. *Economics & Philosophy 11*(2), 227–253.

Reichenbach, H. (1971, originally published in 1949). *The Theory of Probability* (2nd ed.). Berkeley: University of California Press.

Romeijn, J.-W. (2019a). An Interpretation of Weights in Linear Opinion Pooling. ms..

Romeijn, J.-W. (2019b). Extremizing: the Rationality of Audacious Forecasting. ms..

Romeijn, J.-W. and O. Roy (2018). All Agreed: Aumann Meets DeGroot. *Theory and Decision 85*(1), 41–60.

Russell, J. S., J. Hawthorne, and L. Buchak (2015). Groupthink. *Philosophical Studies 172*(5), 1287–1309.

Steele, K. (2012). Testimony as Evidence: More Problems for Linear Pooling. *Journal of Philosophical Logic 41*(6), 983–999.

Tetlock, P. and D. Gardner (2016). *Superforecasting: The Art and Science of Prediction.* New York: Penguin.

van Fraassen, B. (1984). Belief and the Will. *The Journal of Philosophy 81*(5), 235–256.