A Theory Of Bayesian Groups

Dietrich, Franz

Paris School of Economics, CNRS

November 2016

A theory of Bayesian groups

Franz Dietrich[1]

Abstract

A group is often construed as a single agent with its own probabilistic beliefs (credences), which are obtained by aggregating those of the individuals, for instance through averaging. In their celebrated contribution "Groupthink", Russell et al. (2015) apply the Bayesian paradigm to groups by requiring group credences to undergo a Bayesian revision whenever new information is learnt, i.e., whenever the individual credences undergo a Bayesian revision based on this information. Bayesians should often strengthen this requirement by extending it to *non-public* or even *private* information (learnt by *not all* or *just one* individual), or to *non-representable* information (not corresponding to an event in the algebra on which credences are held). I propose a taxonomy of six kinds of 'group Bayesianism', which differ in the type of information for which Bayesian revision of group credences is required: public representable information, private representable information, public non-representable information, and so on. Six corresponding theorems establish exactly how individual credences must (not) be aggregated such that the resulting group credences obey group Bayesianism of any given type, respectively. Aggregating individual credences through averaging is never permitted. One of the theorems – the one concerned with public representable information – is essentially Russell et al.'s central result (with minor corrections).

1 Three challenges for Bayesian groups

Bayesianism requires an agent's beliefs to take the form of coherent probability assignments (*probabilism*) and to be revised via Bayes' rule given new information (*conditionalization*). Let us apply these requirements to a group agent: let a group itself hold probabilistic beliefs and revise them via Bayes' rule. Such Bayesianism for groups – or *group Bayesianism* – faces three challenges which distinguish it from ordinary Bayesianism for individuals.

The first challenge comes from the fact that group beliefs are not free-floating, but determined at any point of time by the current beliefs of the group members, as is usually assumed. Formally, there exists a function, the *pooling rule*, which transforms any possible combination of individual credences into group credences. For instance the averaging rule defines the group credence in an event as the average individual credence in it. The question is: which pooling rules guarantee group Bayesianism? To see the problem, imagine new information comes in. According to the pooling rule, the new group beliefs are obtained by pooling the new individual beliefs. Meanwhile by group Bayesianism the new group beliefs are obtained by revising the old group beliefs via Bayes' rule. So pooling the revised individual beliefs should yield the same as revising the old group beliefs. This places a severe mathematical constraint on the choice of pooling rule. The mentioned averaging rule violates this constraint; so it generates non-Bayesian group beliefs. One might try to defend averaging by arguing that Bayesian conditionalization is not always the right revision policy (Joyce 1999, Hájek 2003) and that averaging may suit the different revision policy of 'imaging' (Leitgeb forth.), and besides that averaging is the basis of Lehrer and Wagner's (1981) consensus formation theory. But if we accept the Bayesian paradigm, as in this paper, then the failure of group beliefs to obey conditionalization is a death penalty for the averaging rule, so that we must search for other pooling rules, as done by Russell et al. (2015) and the present paper.

The second challenge pertains to the question of what information learning actually means for a group. Who learns? I propose to distinguish between *public* information (learnt by *all* members), *private* information (learnt by *only one* member), and *partially spread* information (learnt by *some but not all* members). The question is for which type(s) of information to require Bayesian revision of group beliefs.

The third challenge pertains to the fact that some information might not be representable by any event in the domain (algebra) on which credences are defined. The group might learn that the radio forecasts rainy weather, but it might hold credences only relative to 'weather events', not 'weather-forecast events'. In such a case ordinary Bayesian revision is not even defined. Yet a generalized form of Bayesian revision can still be applied, as explained later. The question is whether to require Bayesian revision of group beliefs even for non-representable information. This question is of course not strictly limited to group agents; it could be raised for individual agents too. But the question is far more pressing for group agents, because the domain of group beliefs (the algebra of events to which the group assigns probabilities) tends to be much smaller than the domain of an individual's beliefs, so that information tends to be far less often representable for groups than for individuals. This is true for practical and theoretical reasons.[2] It is thus urgent to account for

---

[2]In practice, it is hard or impossible to form group beliefs on more than a few events via explicit aggregation or voting. So the domain of real-life group beliefs formed via voting is a fortiori small.

non-representable information when properly studying the revision of group beliefs.

The second and third challenges pertain to the notion of information relevant to groups. Instead of definitely opting for some notion of information, I will consider different notions: *public representable* information, *private representable* information, *public non-representable* information, and so on. Each type of information considered will give rise to a specific form of group Bayesianism, requiring Bayesian conditionalization on information of *this* type.

The paper makes a conceptual and a mathematical contribution. The conceptual contribution is to lay out a taxonomy of six kinds of group Bayesianism, as just indicated. The mathematical contribution is to determine those credence pooling rules which guarantee group Bayesianism of each given sort. This is done in six theorems, one for each kind of group Bayesianism. These theorems respond to the first challenge, and do so for different types of group Bayesianism, i.e., different positions one might take relative to the second and third challenges.

Although the Bayesian paradigm has been applied to groups in the literature on probabilistic opinion pooling (e.g., Madansky 1964, Morris 1974, Dietrich 2010), Russell et al.'s (2015) prize-winning contribution[3] seems to be the first to put on the agenda the most basic Bayesian principle, i.e., standard conditionalization, opening up a new research programme. They however take for granted that information is public and representable, thereby restricting attention to one type of group Bayesianism. The present theorem for this type of group Bayesianism is essentially their central result, except from minor variations and corrections. Another type of group Bayesianism – that for public *non*-representable information – already appeared much earlier in the statistics literature under the label "external Bayesianity" (Madansky 1964). The theorem for this type of group Bayesianism seems to be new. All this calls for an explicit theory of group Bayesianism(s), which this paper hopes to deliver. A very different, so-called 'supra-Bayesian' approach to group beliefs goes back to Morris (1974).[4] Probabilistic opinion pooling is reviewed in Genest and Zidek (1986) and Dietrich and List (2016).

Also in theory group beliefs are defined for fewer events than individual beliefs. Indeed, since group credences are obtained by aggregating individual credences, group credences can only exist where individual credences exist, so that the domain of group beliefs must be *at most as large* as the intersection of the (often different) individual domains of beliefs. That intersection might be very small.

[3]It was selected by The Philosopher's Annual as one of the ten best philosophy papers in 2015.

[4]Supra-Bayesianism reduces group beliefs to the posterior beliefs of an external social planner who treats the group members' credences as evidence on which he conditionalizes his own beliefs. I would argue that this reduction violates the very notion of group beliefs, which is not supposed to depend on any external observer. Supra-Bayesianism is not a theory of Bayesian groups.

2 The formal machinery of credence pooling

Consider a group of n individuals. We label them $i = 1, 2, ..., n$. The group size n is any finite number greater than one. The individuals hold probabilistic beliefs (credences) relative to certain events. As usual, the set of these events forms an algebra, so that we can negate and conjoin events. To model this, I introduce a set W of *worlds*, and define *events* as arbitrary sets of worlds $A \subseteq W$. The number of worlds in W is finite and exceeds two; the infinite case is addressed in Appendix A.[5]

A *credence function* is a probability function C on the set of events.[6] The probability $C(A)$ of an event A is called the *credence in A*. The credence in a world $a \in W$ is of course defined as the credence in the corresponding event: $C(a) := C(\{a\})$. Note that $\sum_{a \in W} C(a) = 1$ and that the probabilities of worlds fully determine those of all events.

The beliefs of the various group members are summarized in the 'credence profile'. Formally, a *(credence) profile* is a list $\mathbf{C} = (C_1, ..., C_n)$ of credence functions, where C_i represents the credences of member i. I use bold-face symbols ($\mathbf{C}$, $\mathbf{C}'$, ...) to denote credence profiles as opposed to single credence functions. For any so-denoted profile I denote its members by 'un-bolding' the symbol and adding individual indices. So the profile $\mathbf{C}$ is made up of $C_1, ..., C_n$, the profile $\mathbf{C}'$ of $C_1', ..., C_n'$, and so on. A credence profile $\mathbf{C}$ is *coherent* if at least one world has non-zero probability under each individual credence function in $\mathbf{C}$; otherwise the profile is *incoherent*. Coherence is a plausible feature. For one would expect that at least the *true* world – whichever world it is – receives non-zero probability by everyone. After all no-one should have any (evidential or theoretical) grounds for totally excluding the true world.

Given a credence profile, what should the group as a whole believe? An answer to this question can be formally captured by a a pooling rule, i.e., a function which aggregates the credence profile into group credences. Formally, a *pooling rule* is a function ag mapping any credence profile $\mathbf{C}$ (from the rule's domain of applicability) to a 'group' credence function $\mathrm{ag}(\mathbf{C})$, denoted $\mathrm{ag}\mathbf{C}$ for short. I now give four examples, representing different approaches or theories of how group credences depend on individual credences:

- The *averaging rule* defines the group credence in an event A as the average of individual credences: $\mathrm{ag}\mathbf{C}(A) = \frac{1}{n}C_1(A) + \cdots + \frac{1}{n}C_n(A)$. The rule's domain of applicability is universal, i.e., consists of *all* credence profiles, since averages of probability functions are always well-defined probability functions.

- More generally, the *weighted averaging rule* with weights $w_1, ..., w_n \geq 0$ of sum

---

[5]Some readers might prefer the objects of beliefs to be propositions; they should simply reinterpret events as propositions. Others might not like modelling events (or propositions) as sets of worlds; I work with sets of worlds following common practice, but nothing hinges on this.

[6]Technically, it is a function C mapping events to numbers in $[0, 1]$ such that C is additive (i.e. $C(A \cup B) = C(A) + C(B)$ whenever $A \cap B = \emptyset$) and $C(W) = 1$.

one is the rule which defines the group credence in an event A as the weighted average of individual credences: $\mathrm{ag}\mathbf{C}(A) = w_1 C_1(A) + \cdots + w_n C_n(A)$. The rule again applies to all credence profiles. Setting all weights to $\frac{1}{n}$ yields the ordinary averaging rule.

- The *geometric rule* defines the group credence in a world a as the (re-scaled) geometric average of individual credences: $\mathrm{ag}\mathbf{C}(a) = k[C_1(a)]^{1/n} \cdots [C_n(a)]^{1/n}$, where k is a profile-dependent scaling factor determined such that the total probability of worlds is one (so $k = 1/\{\sum_{b \in W}[C_1(b)]^{1/n}\cdots[C_n(b)]^{1/n}\}$). The rule's domain of applicability is not universal. It includes only the coherent credence profiles, because for incoherent profiles $\mathbf{C}$ the geometric average $[C_1(a)]^{1/n}\cdots[C_n(a)]^{1/n}$ is zero at all worlds a and so cannot be re-scaled to a probability function. The definition focuses on group credences in *worlds*, but group credences in *events* follow automatically by summing across corresponding worlds.

- More generally, the *weighted geometric rule* with weights $w_1, ..., w_n \geq 0$ defines the group credence in a world a by a (re-scaled) weighted geometric expression: $\mathrm{ag}\mathbf{C}(a) = k[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}$, where k is again a scaling factor ensuring a total probability of one (so $k = 1/\{\sum_{b \in W}[C_1(b)]^{w_1}\cdots[C(b)]^{w_n}\}$). The rule applies only to coherent credence profiles to ensure well-definedness. The weights $w_1, ..., w_n$ might or might not sum to one. Setting all weights to $\frac{1}{n}$ yields the ordinary geometric rule.

3 Bayesian conditionalization for groups

Bayesianism requires that an agent who learns an event E revises his credence function C by adopting the (conditional) credence function $C' = C(\cdot|E)$ which to any event A assigns the conditional probability $C(A|E) = \frac{C(A \cap E)}{C(E)}$. This assumes that $C(E) \neq 0$ to ensure that conditionalization is defined. Henceforth, expressions like 'conditionalizing the credence function C on E' and 'conditionalization of C on E' will denote that the conditional credence function $C(\cdot|E)$ is being formed, and a fortiori that $C(E) \neq 0$.

Like Russell et al. (2015), I apply the requirement of Bayesian conditionalization to groups: group credences should change by conditionalization whenever a new event E is learnt. So the group's new credences which aggregate the post-information profile $\mathbf{C}'$ must be obtainable by conditionalizing the group's old credences which aggregate the pre-information profile $\mathbf{C}$. Formally: $\mathrm{ag}\mathbf{C}' = \mathrm{ag}\mathbf{C}(\cdot|E)$. In other words, Bayesian revision and aggregation commute, as illustrated in Figure 1. However, what does it mean that E is learnt? Russell et al. take it for granted that information is public: *all* group members learn E, so that the new credence profile is $\mathbf{C}' = (C_1(\cdot|E), ..., C_n(\cdot|E))$. Alternatively, E might be learnt just by individual 1,
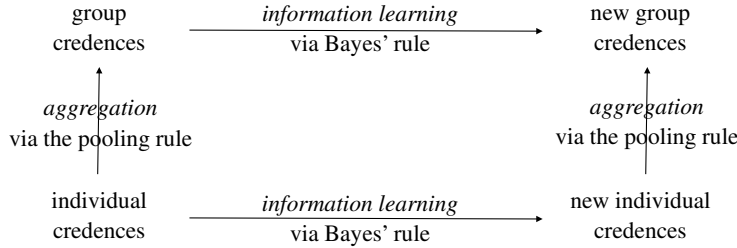
Figure 1: Revising aggregate credences versus aggregating revised credences

so that the new credence profile is $\mathbf{C}' = (C_1(\cdot|E), C_2, ..., C_n)$ in which individuals $2, ..., n$ have kept their old credences. In full generality, E might be learnt by some arbitrary subgroup of one or more individuals, so that only the credences of these individuals change.

These considerations suggest the following group Bayesianism axiom:

Conditionalization on information (Bay): If a credence profile $\mathbf{C}$ changes to another one $\mathbf{C}'$ by conditionalization of *one or more* individual credence functions on an event E (and if the rule applies to $\mathbf{C}$ and $\mathbf{C}'$), then the new group credence function ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on E.

This axiom strengthens a group Bayesianism axiom restricted to public information and introduced by Russell et al.:

Conditionalization on public information (BayPub): If a credence profile $\mathbf{C}$ changes to another one $\mathbf{C}'$ by conditionalization of *all* individual credence functions on an event E (and if the rule applies to $\mathbf{C}$ and $\mathbf{C}'$), then the new group credence function ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on E.

A third group Bayesianism axiom focuses on private information:

Conditionalization on private information (BayPri): If a credence profile $\mathbf{C}$ changes to another one $\mathbf{C}'$ by conditionalization of *exactly one* individual credence function on an event E (and if the rule applies to $\mathbf{C}$ and $\mathbf{C}'$), then the new group credence function ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on E.

All three incarnations of group Bayesianism are prima facie of interest and have their privileged contexts of application, as argued in Section 8. Before exploring each axiom formally, let me give five arguments for why non-public information matters.

First, the Bayesian paradigm requires conditionalization as the universal belief revision policy. There is no principled Bayesian reason for suddenly lifting the requirement if information is not public. Any failure to conditionalize on information is un-Bayesian, regardless of how many or few people have access to the information.

The question of how widely information spreads is epistemically irrelevant, at least to Bayesians. Information matters not in virtue of being widely accessible, but in virtue of being *true*, where truth is ascertained as soon as one individual fully acquires the information. Repeated observation of the exactly same information (by different people) is no better than one-time observation, in vague analogy to the old evidence problem (e.g., Glymour l980, Hartmann and Fitelson forth.)

Second, let us see where radical Bayesianism takes us (without necessarily committing to it). A full-fledged Bayesian has a highly subjective notion of information. He will submit that information is almost never public and hence that the axiom BayPub neglects most instances of information learning in groups. This is because two individuals almost never learn precisely the same event: even when Anne and Peter both see the car arriving, they will have seen the car from slightly different angles and will thus have observed (and conditionalized on) slightly different events. This of course assumes that information is described in full detail, which renders the algebra of events and thus the set of possible worlds W very rich and complex – an unrealistic but standard Bayesian assumption.

Third, groups which fail to conditionalize on information are Dutch-bookable regardless of whether the information is public. Russell et al. put forward the Dutch book argument to defend conditionalization on public information. The argument is easily adapted to non-public information: it suffices to choose the bookie as someone who learns the (non-public) information, possibly even a group member.

Fourth, differences in information across a group constitute a salient real-life phenomenon which is at the heart of theories of group agency, multi-agent systems and distributive cognition. Groups are often said to know more than each of their members. In our framework, this means that group credences incorporate all information held by *at least one* member, which immediately suggests the axiom Bay. By contrast, the weaker axiom BayPub reflects the different idea that a group knows only what *all* (not *some*) members know, so that the group typically knows much less than each of its members.

Fifth, it seems ad hoc to exclude learning of non-public information, i.e., asymmetries in learning across individuals, because on the other hand we do allow asymmetries in status-quo knowledge. Status-quo knowledge can differ across individuals since in a credence profile **C** different individuals can be certain of (i.e., assign probability one to) different events. So the framework is geared towards knowledge asymmetries at any given point of time, i.e., within any given profile. If individuals always learned the same things, one wonders how they could end up knowing different things.

4 The implication of Bayesian conditionalization for groups

What does group Bayesianism in each of the above versions Bay, BayPub or BayPri imply for how group beliefs must be formed, i.e., how the pooling rule must look like? To see how severely group Bayesianism constrains the pooling rule, note that once we have fixed how a given profile $\mathbf{C}$ is aggregated, we are no longer free in how to aggregate any other profile $\mathbf{C}'$ which can arise from $\mathbf{C}$ through information learning: ag$\mathbf{C}'$ must notoriously be given by conditionalization of ag$\mathbf{C}$ on the information.

Before establishing the precise implication of each axiom, I clarify the logical relation *between* the three axioms. Surprisingly, BayPri is only apparently weaker than Bay: groups which conditionalize on *private* information must also conditionalize on *non-private* information (this will no longer be true for non-representable information, as seen later). By contrast, BayPub is a genuinely weaker axiom. The logical gap between BayPub and Bay is filled by a crisp axiom:

Certainty adoption (Cert): Events which are certain to *some* group member are certain to the group, i.e., for all credence profiles profiles $\mathbf{C}$ (in the rule's domain) and events E, if $C_i(E) = 1$ for some individuals i, then ag$\mathbf{C}(E) = 1$.

Cert is a plausible axiom in groups of rational agents, because if some group member is fully certain of E, then he presumably has definitive evidence or arguments for E, so that the group has reason to adopt that certainty. The following result summarizes the mentioned logical relationships:

Proposition 1. *A rule for pooling coherent credence profiles satisfies Bay if and only if it satisfies BayPri, and if and only if it satisfies both BayPub and Cert.*

I now consider each of the three Bayesian axioms in turn and study its implication. I shall use two auxiliary axioms which, broadly speaking, force the pooling rule to be non-degenerate or well-behaved. The first auxiliary axiom requires that if every group member is utterly ignorant, i.e., holds the uniform credence function (which deems each world equally likely), then also the group as a whole is utterly ignorant:

Indifference preservation (Indiff): If $\mathbf{C}$ is the credence profile in which the individuals unanimously hold the uniform credence function (and if the rule applies to $\mathbf{C}$), then the group credence function ag$\mathbf{C}$ is also uniform.

The second well-behavedness axiom requires group credences to depend *continuously* on individual credences: small changes in individual credences should never lead to jumps in group credences. Formally, an infinite sequence of credence functions $C^1, C^2, \ldots$ *converges* to a credence function C if for every event A the sequence of probabilities $C^1(A), C^2(A), \ldots$ converges to $C(A)$.

Continuity (Contin): If a sequence of credence profiles $\mathbf{C}^1, \mathbf{C}^2, \dots$ converges in each individual component to a credence profile $\mathbf{C}$ (and if the rule applies to all these profiles), then the sequence of group credence functions $\mathrm{ag}\mathbf{C}^1, \mathrm{ag}\mathbf{C}^2, \dots$ converges to $\mathrm{ag}\mathbf{C}$.

By the first theorem, the full-blown Bayesian axiom Bay (along with the two well-behavedness axioms) forces the pooling rule to be a weighted geometric rule in which every individual has non-zero weight, i.e., 'has a say':

Theorem 1. *The only rules for pooling coherent credence profiles satisfying Bay, Indiff and Contin are the weighted geometric rules giving non-zero weight to each individual.*

So all pooling rules except weighted geometric rules with non-zero weights are un-Bayesian (by violating Bay) or degenerate (by violating Indiff or Contin). For instance, all weighted or unweighted averaging rules and all weighted geometric rules giving zero weight to someone violate Bay; but they satisfy Indiff and Contin. What is the intuition behind the fact that the three axioms are jointly *necessary and sufficient* for the rule to be of this special geometric sort? Sufficiency is hard to prove. As for necessity, one easily checks that a weighted geometric rule is continuous and preserves indifference. Why does it also satisfy Bay, assuming no individual has zero weight? Suppose certain individuals learn an event E, so that the profile changes. For every individual i who has learnt E, his credences in worlds change to zero for worlds outside E and change proportionally for worlds inside E – this is how conditionalization works. As a result, the expression $[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}$ changes to zero for worlds a outside E and changes proportionally for worlds a inside E. This implies that group credences change via conditionalization on E, as required by Bay. It is crucial in this argument that every weight w_i is non-zero: otherwise it can happen that everyone i who learns E has weight $w_i = 0$, so that his belief revision leaves the w_i-th power of his credences in worlds unchanged. For p^0 is always defined as 1, even for $p = 0$.

Next we turn to the weaker group Bayesianism axiom BayPub which allows non-Bayesian revision in the face of non-public information. Being weaker, this requirement opens the door to a larger class of pooling rules, namely by allowing geometric rules with some zero weighs:

Theorem 2. *The only rules for pooling coherent credence profiles satisfying BayPub, Indiff and Contin are the weighted geometric rules giving non-zero weight to at least one individual.*

Why does a weighted geometric rule meet BayPub as soon as *one* individual i gets non-zero weight? In short, public information E is then guaranteed to be observed by someone with non-zero weight, which suffices to push the group's credence in worlds outside E to zero.

Theorem 2 is essentially Russell et al.'s central theorem (their 'Fact 4'), to which it however adds three necessary qualifications and one optional amendment. The optional amendment is that I impose indifference preservation instead of Russell et al.'s *neutrality* axiom, since indifference preservation is less demanding and achieves the same.[7] As for the three qualifications, firstly I assume the number of worlds to be finite rather than possibly countably infinite, to ensure that weighted geometric rules are well-defined for any non-negative weights; in Appendix A I show how the countably infinite case can be handled.[8] Secondly, I do not permit all weighted geometric rules, but only those with at least one non-zero weight.[9] Thirdly, I allow only *coherent* credence profiles. The third qualification is already introduced retrospectively by Russell et al. in their proof appendix where they restate their result differently. Some of their readers might come to think that the result is essentially true even without excluding incoherent profiles, i.e., that the result *is* true without domain restriction provided one suitably extends weighted geometric rules to incoherent profiles. This is not the case. Without domain restriction the axioms are *inconsistent* with all weighted geometric rules (however extended) except from the dictatorship-like rules assigning zero weight to all but one individual. I return to the aggregation of possibly incoherent credence profiles in Section 7, where I show that group Bayesianism is essentially impossible in 'incoherent groups'.

Finally, what is the implication of requiring group Bayesianism relative to *private* information? Since BayPri is equivalent to Bay (by Proposition 1), the implication of BayPri is precisely that of Bay. So we can restate Theorem 1 using BayPri instead of Bay:

---

[7]Indifference preservation is a particularly weak sort of unanimity axiom, since it requires preserving not *all* unanimously held credence functions, but only the *uniform* one. The neutrality axiom requires treating all worlds equally. Formally, whenever π is a permutation of the set of worlds (which allows us to transform any credence function C into a new one C^π given by $C^\pi(a) = C(\pi(a))$ for all worlds a), then transforming the aggregate credence function ag$\mathbf{C}$ is equivalent to aggregating the profile $\mathbf{C}^\pi$ of transformed credence functions: $(\text{ag}\mathbf{C})^\pi = \text{ag}\mathbf{C}^\pi$. Neutrality implies indifference preservation because transforming the uniform credence function under a permutation yields the same uniform credence function.

[8]The problem with applying the notion of geometric rules naively to a countably infinite set of worlds is that if the weights $w_1, ..., w_n$ sum to a value below one, then for certain coherent credence profiles $\mathbf{C}$ the geometric average $[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}$ has an infinite sum across worlds a and thus fails to be rescalable such that the sum across worlds is one (defining the scaling factor as $k = \frac{1}{\infty} = 0$ does not do). See Appendix A for details.

[9]The statement of Russell et al.'s Fact 4 ("The only rules which obey [the axioms] are Weighted Geometric Rules") allows for two readings: either the rules obeying the axioms are claimed to be *all the* Weighted Geometric Rules (as suggested by the authors' claim to *characterize* weighted geometric pooling), or the rules obeying the axioms are claimed to be *among the* Weighted Geometric Rules (as suggested by the authors' restatement of their Fact 4 in their appendix). Under the first reading Theorem 2 corrects their Fact 4. Under the second reading Theorem 2 strengthens their Fact 4 by turning an implication into an equivalence, i.e., into a characterization result.

Theorem 3. *The only rules for pooling coherent credence profiles satisfying BayPri, Indiff and Contin are the weighted geometric rules giving non-zero weight to each individual.*

5 Bayesian conditionalization for groups facing non-representable information

A key idealization often made by Bayesians is that any information an agent might ever learn is representable as an event within the domain (algebra) where the agent assigns probabilities. This ensures that Bayes' rule in its ordinary form applies. Real-life information need not be representable in this way, in particular in the context of *group* agents which tend to hold beliefs relative to a small event algebra that excludes much of what can be learnt. Taking up an introductory example, the group might hold credences only relative to weather events; so worlds in W describe the weather, nothing else. The information that *the radio forecasts rain* is not representable as an event $E \subseteq W$ since worlds in W do not describe weather forecasts. Yet credences should clearly be revised, presumably by raising the probability of the (representable) event of rain.

How should credences be revised based on non-representable information? There is a well-known Bayesian answer: model such information as a likelihood function rather than an event and apply Bayes' rule in its generalized form. Although all this is known to Bayesian statisticians, a short introduction is due. A *likelihood function* is an arbitrary function L from worlds to numbers in $[0, 1]$. One interprets L as modelling some information and $L(a)$ as being the probability of that information given that the world is a. In the weather example, $L(a)$ is the probability that the radio forecasts rain (the information) given that the world is a. Since weather forecasts are usually right, $L(a)$ is near 1 for rainy worlds a and near 0 for non-rainy worlds a.

Given how likelihood functions are interpreted, it is clear how one should conditionalize on them, i.e., how Bayes' rule in its generalized version works. An agent who learns a likelihood function L should revise his credence function C by adopting the (conditional) credence function $C(\cdot|L)$ which to every world a assigns the probability $C(a|L) = \frac{C(a)L(a)}{\sum_{b \in W} C(b)L(b)}$. One immediately recognizes Bayes' rule, given that $L(a)$ stands for the probability of information conditional on a. The conditional credence function $C(\cdot|L)$ is only defined if the likelihood function L is *coherent* with C, i.e., if there is at least one world where both L and C are non-zero, ensuring that $\sum_{b \in W} C(b)L(b) \neq 0$. Intuitively, coherence of L with C means that the information is not ruled out by the initial credences. Hereafter, expressions like 'conditionalizing C on L' and 'conditionalization of C on L' will denote that the conditional credence function $C(\cdot|L)$ is being formed, and a fortiori that L is coherent with C.

Likelihood functions generalize events as a model of information, and Bayes' rule for likelihood functions generalizes Bayes' rule for events. Indeed, to any event E corresponds a simple likelihood function L for which $L(a)$ can only be 1 or 0, depending on whether a is in E or outside E; and conditionalizing on E is equivalent to conditionalizing on the corresponding likelihood function L, as one easily checks.

It is natural to require groups to follow Bayes' rule not just if an event is learnt, but more generally if a likelihood function is learnt. This requirement can once again be fleshed out in three different ways, depending on whether the likelihood function is learnt by *any subgroup* of individuals, or by *all* individuals (public information), or by *just one* individual (private information). The three resulting axioms are counterparts of the earlier axioms Bay, BayPub and BayPri. They differ from their counterparts *only* in that the learnt information is given by a likelihood function L rather than an event E. Being based on a more general notion of information to be called *L-information*, each new axiom is strictly stronger than its counterpart, as indicated by the '+' in the label of each new axiom.

Conditionalization on L-information (Bay+): If a credence profile $\mathbf{C}$ changes to another one $\mathbf{C}'$ by conditionalization of *one or more* individual credence functions on a likelihood function L (and if the rule applies to $\mathbf{C}$ and $\mathbf{C}'$), then the new group credence function ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on L.

Conditionalization on public L-information (BayPub+): If a credence profile $\mathbf{C}$ changes to another one $\mathbf{C}'$ by conditionalization of *all* individual credence functions on a likelihood function L (and if the rule applies to $\mathbf{C}$ and $\mathbf{C}'$), then the new group credence function ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on L.

Conditionalization on private L-information (BayPri+): If a credence profile $\mathbf{C}$ changes to another one $\mathbf{C}'$ by conditionalization of *exactly one* individual credence function on a likelihood function L (and if the rule applies to $\mathbf{C}$ and $\mathbf{C}'$), then the new group credence function ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on L.

6 The implication of Bayesian conditionalization for groups facing non-representable information

We have seen in Section 5 that a group which obeys *ordinary* Bayesian conditionalization – Bayesian conditionalization on *events* – must form its credences in a particular way that depends on the chosen group Bayesian axiom (Bay, BayPub or BayPri). What happens to the pooling rule if we impose Bayesian revision even for non-representable information, i.e., if we require Bay+, BayPub+ or BayPri+?

As in Section 5, I start the analysis by clarifying the logical relationship between the three axioms at stake. The situation changes dramatically compared to the earlier

axioms Bay, BayPub and BayPri. While the earlier axioms are highly compatible with each other (by Proposition 1), the new axioms are mutually incompatible:

Proposition 2. *No rule for pooling coherent credence profiles satisfies both BayPub+ and BayPri+.*

So, in short, group credences cannot incorporate both public and private L-information in a proper Bayesian way. As an immediate consequence, the full-blown axiom Bay+, which simultaneously strengthens BayPub+ and BayPri+, is internally inconsistent:

Theorem 4. *No rule for pooling coherent credence profiles satisfies Bay+.*

This striking impossibility does not require imposing any of the well-behavedness axioms Indiff and Contin: Bay+ is *by itself* inconsistent, hence untenable as a normative principle for group beliefs. How should we interpret this? On one interpretation, groups simply cannot be 'fully Bayesian': their belief revision policy cannot be as ideally rational as that of single individuals. But there is a more nuanced interpretation. Recall that the need to conditionalize on non-representable information came from a lack of Bayesian rationality in the first place: an inability to assign probabilities to 'everything', so that the set of events under consideration – the credence domain – fails to encompass all relevant information. I gave an example where the credence domain fails to contain an event representing the information of a rainy weather forecast. If by contrast the credence domain is universal, as many Bayesians routinely assume, then all relevant information is by definition representable by an event in the credence domain, and we lose the justification for introducing L-information and imposing Bay+ because the initial axiom Bay already covers all relevant information. In sum, Bay+ becomes normatively mandatory *only when and because* another Bayesian requirement – that of a universal credence domain – is violated. Accordingly, Theorem 4 does not tell us that groups cannot be fully Bayesian, but that they cannot be 'semi-Bayesian' by failing to entertain a universal credence domain while properly conditionalizing on information outside the credence domain.

The impossibility disappears once we restrict attention to public or to private L-information. Indeed groups *can* follow Bayesian conditionalization on such information, by using a pooling rule of a quite particular kind. I begin with public L-information:

Theorem 5. *The only rules for pooling coherent credence profiles satisfying Bay-Pub+, Indiff and Contin are the geometric rules whose individual weights sum to one.*

The comparison to Theorem 2 shows that BayPub+ constrains the pooling rule much more than BayPub does: the individual weights must now sum to one. Surprisingly, this result seems to be new, although its central axiom BayPub+ has already

been studied under the label "external Bayesianity", though in a different framework (Madansky 1964).[10]

Finally, how can groups follow Bayesian conditionalization on *private* generalized information? They can do so in precisely one way, namely by using the *multiplicative* pooling rule (Dietrich 2010, Dietrich and List 2016). It determines the group credence in each world as the (re-scaled) *product* of the members' credences in that world. This is a special case of weighted geometric pooling in which each individual gets weight one.

Theorem 6. *The only rule for pooling coherent credence profiles satisfying BayPri+ and Indiff is the multiplicative rule, i.e., the geometric rule giving weight one to each individual.*

The comparison of Theorems 5 and 6 shows that it makes a considerable difference whether the group wishes to properly incorporate public or private L-information. In the former case the weights must sum to one, in the latter they must all equal one. This gives an idea of why the two axioms are mutually inconsistent (see Proposition 2). Theorem 6 does not involve the axiom Contin. It is a version of a result by Dietrich and List (2016) in a different framework.[11]

7 The impossibility of group Bayesianism for incoherent groups

Our analysis has so far been limited to rules that pool *coherent* credence profiles, in which at least one world is assigned non-zero probability by everyone. In short, we have excluded radical disagreement. Incoherent profiles are peculiar in that they violate the idea that some world is 'true' and receives non-zero subjective probability from everyone. Can one design pooling rules that are Bayesian and apply also to incoherent credence profiles? The answer is negative: if we permit incoherent profiles, there do no longer exist any non-degenerate Bayesian rules – regardless of which of our six Bayesian axioms is taken to define group Bayesianism.

---

[10]How could it have escaped the statistics literature that this axiom (jointly with well-behavedness axioms) forces to certain geometric pooling rules? Presumably the reason is that the axiom is usually stated and analysed in a different framework in which credence functions and likelihood functions must take non-zero values at all worlds. This excludes representable information, since a likelihood function corresponding to representable information, i.e., to an event E, takes the value 0 outside E and is thus excluded (unless $E = W$). So the classic axiom of external Bayesianity actually differs from BayPub+ in that it covers *only non-representable* rather than *also non-representable* information. The analogue of Theorem 5 in that classic framework is false, because the (restated) axioms can be met by generalized versions of weighted geometric rules whose weights can depend on the profile in certain systematic ways.

[11]Their framework takes all credence functions and likelihood functions to have non-zero values at all worlds. This difference in framework has no consequence for the result.

To state the result formally, I first define two kinds of degenerate rules for pooling arbitrary credence profiles. A *dictatorship* is a rule such that the group always adopts the credences of a fixed group member. Formally, there is an individual i (the *dictator*) such that $\mathrm{ag}\mathbf{C} = C_i$ for all credence profiles $\mathbf{C}$. A *power dictatorship* is a rule which, like an ordinary dictatorship, makes group credences depend solely on a fixed individual. But the group might not adopt that individual's credences as such: it might adopt a transformed version of his credences, obtained by raising the probabilities of worlds to some power. Formally, a *power dictatorship* is a rule for which there exists an individual i (the *power dictator*) and a number $w > 0$ such that for any credence profile $\mathbf{C}$ the group credences in worlds $a \in W$ are given by $\mathrm{ag}\mathbf{C}(a) = k[C_i(a)]^w$, where $k > 0$ is a scaling factor ensuring that probabilities of worlds sum to one (i.e., $k = \sum_{b \in W}[C_i(b)]^w$). In case $w = 1$ we obtain a regular dictatorship.

Theorem 7. *Among all rules for pooling arbitrary (possibly incoherent) credence profiles,*

(a) *no rules satisfy the axioms stated in Theorem 1, 3, 4, or 6, respectively,*

(b) *only the power dictatorships satisfy the axioms stated in Theorem 2,*

(c) *only the dictatorships satisfy the axioms stated in Theorem 5.*

Let me paraphrase this result. If we seek to aggregate arbitrary credence profiles, then only power dictatorships can properly handle public information, only dictatorships can properly handle public L-information, and no rules whatsoever can properly handle the other four types of information.

8 Conclusion: each group Bayesianism matters

I have argued that there are different types of group Bayesianism, depending on the kind of information on which one requires groups to conditionalize. Each form of group Bayesianism is compatible with certain credence pooling rules, determined in Theorems 1–6. Specifically, group beliefs must be formed via a weighted geometric rule, where the weights must obey certain conditions depending on the type of group Bayesianism in question. Group Bayesianism however becomes impossible if the members can disagree radically, i.e., if the credence profile can be incoherent (Theorem 7).

Which of the six group Bayesian axioms is the right rendition of group Bayesianism? The answer depends on the group or application in question. I propose the following stylized classification. The first dimension of classification concerns how widely information can spread in the group in question:

- *Fully symmetrically informed* groups are idealized groups whose members have exactly the same access to new information (perhaps due to perfect deliberation and information sharing). New information is then by definition public, and the Bayesian axiom need only quantify over public information. This leads to BayPub or BayPub+.

- *Fully asymmetrically informed* groups are idealized groups whose members never learn the same information. New information is then by definition private, and the Bayesian axiom need only quantify over private information. This leads to BayPri or BayPri+.

- Groups with *arbitrary information spread* are groups without any restriction on how widely new information is accessible. New information may thus be acquired by any subgroup, and the Bayesian axiom should quantify over information acquired by any subgroup. This leads to Bay or Bay+.

The second dimension of classification concerns the size of the domain (algebra) of events on which the group in question holds credences:

- Groups with *universal credence domain* are idealized groups in which the domain of credences comprises everything relevant to the group in question, including any information that can be acquired. New information is thus always representable, and the Bayesian axiom need only quantify over representable information. This leads to Bay, BayPub or BayPri.

- Groups with *limited credence domain* are groups in which the credence domain fails to encompass certain information that can be acquired in the group in question. New information can thus be non-representable, and the Bayesian axiom should quantify over generalized information. This leads to Bay+, BayPub+ or BayPri+.

		information access		
		arbitrary	*fully symmetric*	*fully asymmetric*
credence domain	*unrestricted*	axiom: Bay pooling rule: weighted geometric, all weights positive	axiom: BayPub pooling rule: weighted geometric, some weight positive	axiom: BayPri pooling rule: weighted geometric, all weights positive
	restricted	axiom: Bay+ pooling rule: inexistent	axiom: BayPub+ pooling rule: weighted geometric, weights of sum one	axiom: BayPri+ pooling rule: multiplicative

Figure 2: Contexts of application and their corresponding group Bayesianism axioms and pooling rules

Figure 2 summarizes the stylized classification of groups or applications, in each case displaying the relevant group Bayesianism axiom and the corresponding pooling rule(s) according to Theorems 1–6. This shows how strongly the axiomatic rendition of group Bayesianism and the pooling rule should depend on the application.

A   Generalization to infinitely many worlds

The main text took the number of worlds (hence, of events) to be finite. This calls for a generalization. In both appendices let the set of worlds be *countable*, i.e., finite or countably infinite. To extend our formal results to that case, we must do two things: generalize the notion of weighted geometric pooling, and adapt the axiom of Indifference Preservation. I shall do both things in turn. But first let me anticipate what is thereby achieved:

Remark 1. All formal results of the main text (the 'theorems' and 'propositions') hold more generally for countably many worlds if weighted geometric rules are generalized as below and Indifference Preservation is replaced by Weak Indifference Preservation defined below.

Generalizing geometric rules: What can happens if we naively apply our earlier definition of the weighted geometric rule to infinitely many worlds? Given weights $w_1, ..., w_n \geq 0$ and a (coherent) credence profile $\mathbf{C} = (C_1, ..., C_n)$, we first form for each world a the weighted geometric average $[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}$. The trouble arises as we attempt to normalize this expression to a probability mass function: normalization fails when the sum $\sum_{a \in W} [C_1(a)]^{w_1} \cdots [C(a)]^{w_n}$ is infinite. To see that the sum *can* be infinite, let the set of worlds be $W = \{1, 2, 3, ...\}$, let the sum of weights be $w_1 + \cdots + w_n = \frac{1}{2}$, and let each individual i have the same credence function assigning probability ca^{-2} to each world a, where c is a positive constant which ensures that the probabilities of worlds sum to one. The weighted geometric average then takes the form $[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} = (ca^{-2})^{w_1 + \cdots + w_n} = \sqrt{c} a^{-1}$, so that $\sum_{a \in W} [C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} = \sqrt{c} \sum_{a=1}^{\infty} a^{-1} = \sqrt{c} \infty = \infty$. Normalization is thus impossible here. However normalization is guaranteed to be possible for certain choices of the weights:

Proposition 3. *If the number of worlds is (countably) infinite, the following two conditions on weights $w_1, ..., w_n \geq 0$ are equivalent:*

- *The weighted geometric average $[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}$ is normalizable (i.e., has finite sum over worlds a) for each coherent credence profiles $(C_1, ..., C_n)$.*

- *The sum of weights satisfies $w_1 + \cdots + w_n \geq 1$.*

This tells us that for infinitely many worlds weighted geometric pooling is meaningful *if and only if* the sum of weights is at least one. I therefore generalize the

notion of geometric rules as follows to the countable case: a *weighted geometric rule* is defined

- for *arbitrary* weights $w_1, ..., w_n \geq 0$ if the number of worlds is finite,

- for weights $w_1, ..., w_n \geq 0$ *of sum at least one* if the number of worlds is countably infinite,

where for each coherent credence profile the group credence in a world a is determined in the usual way, i.e., as the normalized weighted geometric average credence in a. We can now talk meaningfully about weighted geometric rules for countable W, bearing in mind that the weights by definition have sum at least one if W is infinite. Note that if we were to *require* (rather than *permit*) W to be countably infinite, then we could simplify Theorem 2: we would no longer need to require that at least one individual gets non-zero weight, as this already follows from the sum of weights being at least one.

Adapting Indifference Preservation: The axiom of Indifference Preservation (Indiff) is meaningless for infinitely many worlds, because the uniform distribution does then not exist. Indeed, one cannot assign the same probability x to infinitely many worlds, as the sum of probabilities would not be one, but infinite (if $x > 0$) or zero (if $x = 0$). We can instead use this axiom:

Weak indifference preservation (Indiff*): For all worlds a and b, unanimous indifference between a and b is *at least sometimes* preserved, i.e., there is *at least one* credence profile $\mathbf{C}$ (in the rule's domain) such that every individual i satisfies $C_i(a) = C_i(b) \neq 0$ and the group satisfies $\text{ag}\mathbf{C}(a) = \text{ag}\mathbf{C}(b)$.

This axiom has a double advantage over ordinary Indifference Preservation: (i) it stays meaningful for infinitely many worlds, and (ii) it is weaker for finitely many worlds since the credence profile where everyone holds uniform beliefs automatically has the property required in Indiff*.[12] Our results could use Russell et al.'s 'neutrality' axiom instead of Indiff*; that axiom is however much stronger.

B   Proofs

I now prove all results from the main text and Appendix A. The results from the main text will be proved in their generalized version defined in Appendix A. So throughout the set of worlds W is countable (finite or countably infinite), Indiff* is used instead of Indiff, and the notion of weighted geometric rules is extended to the

---

[12]Strictly speaking, Indiff* is weaker under the minimal assumption that the profile of uniform credence functions belongs to the rule's domain.

infinite case in the above-defined way (so that weights must sum to at least one in the infinite case).

Conventions: The conditionalization of a credence function C on an event E or a likelihood function L will (when existent) be denoted by $C|E$ and $C|L$, respectively. As usual, the *support* of a credence function C is $\text{supp}(C) := \{a \in W : C(a) \neq 0\}$.

B.1 The propositions

Proof of Proposition 1. Consider a rule ag for pooling coherent profiles. Axiom Bay obviously implies BayPub and BayPri. The proof is completed by showing three claims.

Claim 1: BayPri implies Bay.

Assume BayPri and consider coherent profiles $\mathbf{C}$ and $\mathbf{C}'$ such that $\mathbf{C}'$ arises from $\mathbf{C}$ by conditionalization of the credence functions of m individuals on an event E, where $1 \leq m \leq n$. Without loss of generality, suppose these m individuals are the individuals $1, ..., m$. Note that for all $j \in \{0, 1, ..., m\}$ the credence profile $\mathbf{C}^j := (C_1|E, ..., C_j|E, C_{j+1}, ..., C_n)$ is coherent. Moreover, each profile $\mathbf{C}^j$ with $j \neq 0$ arises from $\mathbf{C}^{j-1}$ by conditionalization of exactly one individual credence function on E. So we can apply BayPri repeatedly:

$$
\begin{aligned}
\text{ag}\mathbf{C}^m &= (\text{ag}\mathbf{C}^{m-1})|E \\
&= ((\text{ag}\mathbf{C}^{m-2})|E)|E = (\text{ag}\mathbf{C}^{m-2})|E \\
&= ... \\
&= ((\text{ag}\mathbf{C}^0)|E)|E = (\text{ag}\mathbf{C}^0)|E.
\end{aligned}
$$

Since $\mathbf{C}^0 = \mathbf{C}$ and $\mathbf{C}^m = \mathbf{C}'$, we have shown that $\text{ag}\mathbf{C}' = (\text{ag}\mathbf{C})|E$. This proves Bay.

Claim 2: BayPri implies Cert.

Assume BayPri. Let $\mathbf{C}$ be a coherent profile, E an event and i an individual such that $C_i(E) = 1$. So the profile arising from $\mathbf{C}$ by conditionalization of i's credence function on E is $\mathbf{C}$ itself. Hence by BayPri $\text{ag}\mathbf{C} = (\text{ag}\mathbf{C})|E$. So $\text{ag}\mathbf{C}(E) = 1$, proving Cert.

Claim 3: BayPub and Cert together imply BayPri.

Assume BayPub and Cert. Let a coherent profile $\mathbf{C}'$ arise from another one $\mathbf{C}$ by conditionalization of an individual i's credence function an event E. Let $\mathbf{C}''$ be the profile obtained from $\mathbf{C}$ or equivalently from $\mathbf{C}'$ by conditionalization of *every* credence function on E. Note that $\mathbf{C}''$ is coherent given the way it is obtained from the coherent profile $\mathbf{C}'$ in which an individual assigns probability one to E. Since in $\mathbf{C}'$ individual i assigns probability one to E, by Cert $\text{ag}\mathbf{C}'(E) = 1$. Now

$$
\text{ag}\mathbf{C}' = (\text{ag}\mathbf{C}')|E = \text{ag}\mathbf{C}'' = (\text{ag}\mathbf{C})|E,
$$

where the first equation holds as $\text{ag}\mathbf{C}'(E) = 1$, and the second and third because of BayPub. We have shown that $\text{ag}\mathbf{C}' = (\text{ag}\mathbf{C})|E$, proving BayPri. ∎

Proof of Proposition 2. For a contradiction, let some rule ag for pooling coherent profiles satisfy BayPub+ and BayPri+. Consider a coherent profile $\mathbf{C}$ in which every credence function has full support, and let L be a non-constant likelihood function with full support. For all $j \in \{0, 1, ..., n\}$ define the credence profiles $\mathbf{C}^j := (C_1|L, ..., C_j|L, C_{j+1}, ..., C_n)$. Note that all $\mathbf{C}^j$ are coherent. By BayPub+, $\text{ag}\mathbf{C}^n = (\text{ag}\mathbf{C})|L$. On the other hand, repeated application of BayPri+ yields $\text{ag}\mathbf{C}^n = (\text{ag}\mathbf{C})|L^n$, because

$$
\begin{aligned}
\text{ag}\mathbf{C}^n &= (\text{ag}\mathbf{C}^{n-1})|L \\
&= ((\text{ag}\mathbf{C}^{n-2})|L)|L = (\text{ag}\mathbf{C}^{n-2})|L^2 \\
&= ... \\
&= ((\text{ag}\mathbf{C}^0)|L^{n-1})|L = (\text{ag}\mathbf{C}^0)|L^n = (\text{ag}\mathbf{C})|L^n.
\end{aligned}
$$

As $\text{ag}\mathbf{C}^n = (\text{ag}\mathbf{C})|L$ and $\text{ag}\mathbf{C}^n = (\text{ag}\mathbf{C})|L^n$, we have $(\text{ag}\mathbf{C})|L = (\text{ag}\mathbf{C})|L^n$. It follows that L is proportional to L^n, by definition of conditionalization on a likelihood function (and by the fact that $\text{ag}\mathbf{C}$ has full support, which holds via Lemma 3 below as all C_i have full support). So L must be a constant function, in contradiction to our assumption. ∎

Proof of Proposition 3. Let W be countably infinite, and consider weights $w_1, ..., w_n \geq 0$ whose sum is denoted w.

1. First assume $w < 1$. If $w = 0$, so that $w_1 = \cdots = w_n = 0$, then normalization fails for all profiles $\mathbf{C}$ since $\sum_{a \in W}[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} = \sum_{a \in W} 1 = \infty$. Now let $w > 0$. To show that normalizability can fail, I give a counterexample generalizing that stated in Appendix A. Without loss of generality let worlds be natural numbers: $W = \{1, 2, 3, ...\}$. Consider the credence profile $\mathbf{C}$ in which each C_i assigns probability $ca^{-1/w}$ to world a, where c is a normalization constant ensuring that probabilities of worlds sum to one: $c = 1/\sum_{a=1}^{\infty} a^{-1/w}$. This uses the well-known fact that $\sum_{a=1}^{\infty} a^{-1/w} < \infty$ as $1/w > 1$. So

$$
\sum_{a \in W}[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} = \sum_{a=1}^{\infty} \left(ca^{-1/w}\right)^w = c^w \sum_{a=1}^{\infty} a^{-1} = c^w \infty = \infty.
$$

Here $\sum_{a=1}^{\infty} a^{-1}$ is the so-called harmonic series, which is well-known to have infinite limit.

2. Now assume $w \geq 1$, and consider any coherent profile $\mathbf{C}$. I show normalisability by distinguishing between two cases.

Case 1: $w = 1$. For any world a, we have $[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} \leq w_1 C_1(a) + \cdots + w_n C_n(a)$ by the *inequality between (weighted) geometric and arithmetic means* (e.g., Steele 2004). So

$$\sum_{a \in W} [C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} \leq \sum_{a \in W} [w_1 C_1(a) + \cdots + w_n C_n(a)]$$
$$= w_1 \sum_{a \in W} C_1(a) + \cdots + w_n \sum_{a \in W} C_n(a)$$
$$= w_1 + \cdots + w_n = w = 1 < \infty.$$

Case 2: $w > 1$. I reduce this case to Case 1. For all worlds a and individuals i we have $[C_i(a)]^{w_i} \leq [C_i(a)]^{\frac{w_i}{w}}$ (as $C_i(a) \leq 1$ and $w_i > \frac{w_i}{w}$). So

$$\sum_{a \in W} [C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n} \leq \sum_{a \in W} [C_1(a)]^{\frac{w_1}{w}} \cdots [C_n(a)]^{\frac{w_n}{w}} < \infty,$$

where the last inequality holds by Case 1 applied to the new weights $\frac{w_1}{w}, ..., \frac{w_n}{w}$ of sum one. $\blacksquare$

B.2 Preparing the theorems' necessity proofs

The following two lemmas will later allow us to prove that the axioms in our theorems are *necessary*: each axiom in a theorem is satisfied by each particular (weighted geometric) rule specified in that theorem.

Lemma 1. *A weighted geometric rule satisfies*

(a) *Bay (or equivalently BayPri) if and only if all weights are non-zero,*

(b) *BayPub if and only if at least one weight is non-zero,*

(c) *BayPub+ if and only if the weights sum to one,*

(d) *BayPri+ if and only if all weights are one, i.e., the rule is multiplicative.*

Proof. Consider a weighted geometric rule with weights $w_1, ..., w_n$. The proof will be sketched informally.

(a) The proof that Bay holds if all w_i are non-zero was already given (informally) after Theorem 1. Conversely, if some individual's weight is zero, then conditionalizing his credence function on an event E never affects group credences, so that Bay is violated.

(b) The proof that BayPub holds if some w_i is non-zero was again given informally after Theorem 2. Conversely, if all w_i are zero, which by the way implies that W is finite, then group credences are uniform regardless of the profile, violating BayPub.

(c) Whenever one coherent credence profile $\mathbf{C}'$ arises from another $\mathbf{C}$ by conditionalization of all credence functions on a given likelihood function L, we have (*)

ag$\mathbf{C}' = (\text{ag}\mathbf{C})|L^{w_1+\cdots+w_n}$, i.e., ag$\mathbf{C}'$ is the conditionalization of ag$\mathbf{C}$ on the likelihood function $L^{w_1+\cdots+w_n}$. This is because, for appropriate normalization constants $k, k', k'' > 0$, we have at all worlds a

$$\begin{aligned}
\text{ag}\mathbf{C}'(a) &= k[(C_1|L)(a)]^{w_1} \cdots [(C_n|L)(a)]^{w_1} \\
&= k'[C_1(a)L(a)]^{w_1} \cdots [C_n(a)L(a)]^{w_n} \\
&= k'[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}[L(a)]^{w_1+\cdots+w_n} \\
&= k''[(\text{ag}\mathbf{C})|L^{w_1+\cdots+w_n}](a).
\end{aligned}$$

Clearly, if $w_1 + \cdots w_n = 1$, then BayPub+ holds, as (*) reduces to ag$\mathbf{C}' = (\text{ag}\mathbf{C})|L$. Conversely, assume BayPub+. Then, with $\mathbf{C}, \mathbf{C}', L$ as before, we have ag$\mathbf{C}' = (\text{ag}\mathbf{C})|L$, and hence by (*) $(\text{ag}\mathbf{C})|L = (\text{ag}\mathbf{C})|L^{w_1+\cdots+w_n}$. So L must be proportional to $L^{w_1+\cdots+w_n}$ (assuming ag$\mathbf{C}$ has full support, which we can ensure by letting all credence functions in $\mathbf{C}$ have full support and applying Lemma 3 below). It follows that $w_1 + \cdots + w_n = 1$ (assuming without loss of generality that L was chosen to be non-constant).

(d) For any given individual i, whenever one coherent credence profile $\mathbf{C}'$ arises from another $\mathbf{C}$ by conditionalization of i's credence function on a likelihood function L, we have (**) ag$\mathbf{C}' = (\text{ag}\mathbf{C})|L^{w_i}$. The reason is analogous to that for (*) in the proof of (c).

Now, if $w_1 = \cdots = w_n = 1$, then BayPri+ holds, as (**) reduces to ag$\mathbf{C}' = (\text{ag}\mathbf{C})|L$. Conversely, suppose BayPri+. With $\mathbf{C}, \mathbf{C}', L$ as before, ag$\mathbf{C}' = (\text{ag}\mathbf{C})|L$ by BayPri+. So, for all individuals i, we have $(\text{ag}\mathbf{C})|L = (\text{ag}\mathbf{C})|L^{w_i}$ by (**), implying that $w_i = 1$ by an argument parallel to that in the proof of (c).∎

Lemma 2. *Every weighted geometric rule satisfies Contin, Indiff (if W is finite), and Indiff*.*

Proof. The elementary argument is left to the reader. ∎

B.3 Preparing the theorems' sufficiency proofs

The next lemmas will help us show that the axioms in any of our theorems are *sufficient*: they require the particular type of pooling rule claimed in each theorem, respectively. Central steps of the argument, including the use of Cauchy's functional equation, correspond directly to steps in Russell et al.'s proof of their "Claim 4". Each lemma of this subsection assumes a rule ag for pooling coherent credence profiles.

Lemma 3. *For all coherent credence profiles $\mathbf{C}$,*

(a) under Bay, $\text{supp}(\text{ag}\mathbf{C}) = \cap_i \text{supp}(C_i)$,

(b) under any of the six Bayesian axioms, $\cap_i \text{supp}(C_i) \subseteq \text{supp}(\text{ag}\mathbf{C}) \subseteq \cup_i \text{supp}(C_i)$.

Proof. Let $\mathbf{C}$ be a coherent credence profile. It suffices to prove three claims.

Claim 1: Under Bay, $\mathrm{supp}(\mathrm{ag}\mathbf{C}) \subseteq \cap_i \mathrm{supp}(C_i)$.

Suppose Bay. Let a be a world *not* in $\cap_i \mathrm{supp}(C_i)$. Pick an individual i such that $a \notin \mathrm{supp}(C_i)$. Since $C_i(W \backslash \{a\}) = 1$ and since pooling is certainty adopting by Proposition 1, we have $\mathrm{ag}\mathbf{C}(W \backslash \{a\}) = 1$. So $a \notin \mathrm{supp}(\mathrm{ag}\mathbf{C})$.

Claim 2: Under BayPub (the weakest Bayesian axiom), $\cap_i \mathrm{supp}(C_i) \subseteq \mathrm{supp}(\mathrm{ag}\mathbf{C})$.

Assume BayPub and let $a \in \cap_i \mathrm{supp}(C_i)$. Since the profile $\mathbf{C}'$ in which every individual assigns probability one to a is coherent and arises from $\mathbf{C}$ by conditionalization of everyone's credence function on the singleton event $\{a\}$, BayPub tells us that $\mathrm{ag}\mathbf{C}'$ arises by conditionalization of $\mathrm{ag}\mathbf{C}$ on $\{a\}$. In particular, $a \in \mathrm{supp}(\mathrm{ag}\mathbf{C})$.

Claim 3: Under BayPub, $\mathrm{supp}(\mathrm{ag}\mathbf{C}) \subseteq \cup_i \mathrm{supp}(C_i)$.

Under BayPub, since $\mathbf{C}$ is unchanged if all credence functions are conditionalized on $E = \cup_i \mathrm{supp}(C_i)$, we have $\mathrm{ag}\mathbf{C} = (\mathrm{ag}\mathbf{C})|E$, and thus $\mathrm{ag}\mathbf{C} \subseteq E$.■

Lemma 4. *Under any of the six Bayesian axioms and Indiff*, for all coherent credence profiles $\mathbf{C}$ and worlds $a, b \in W$, if $C_i(a) = C_i(b) \neq 0$ for each individual i, then $\mathrm{ag}\mathbf{C}(a) = \mathrm{ag}\mathbf{C}(b) \neq 0$.*

Proof. Assume Indiff* and BayPub, the weakest Bayesian axiom by Proposition 1. Consider a coherent profile $\mathbf{C}$ and $a, b \in W$ such that $C_i(a) = C_i(b) \neq 0$ for all individuals i. By Indiff* there is another coherent profile $\mathbf{C}'$ such that $C_i'(a) = C_i'(b) \neq 0$ for all individuals i and $\mathrm{ag}\mathbf{C}'(a) = \mathrm{ag}\mathbf{C}'(b)$. Conditionalizing all members of $\mathbf{C}$ on $E = \{a, b\}$ yields the same (coherent) profile, denoted $\mathbf{C}''$, as conditionalizing all members of $\mathbf{C}'$ on E. So, applying BayPub twice, $(\mathrm{ag}\mathbf{C})|E = \mathrm{ag}\mathbf{C}'' = (\mathrm{ag}\mathbf{C}')|E$. Hence, as $((\mathrm{ag}\mathbf{C}')|E)(a) = ((\mathrm{ag}\mathbf{C}')|E)(b)$, we have $((\mathrm{ag}\mathbf{C})|E)(a) = ((\mathrm{ag}\mathbf{C})|E)(b)$, and thus $\mathrm{ag}\mathbf{C}(a) = \mathrm{ag}\mathbf{C}(b)$. Finally, this value is non-zero, since otherwise $\mathrm{ag}\mathbf{C}$ would assign zero prob and could thus not be conditionalized on E. ■

Lemma 5. *Under any of the six Bayesian axioms and Indiff*,*

(a) *group probability ratios are a function of individual probability ratios, i.e., there exists a unique function f from $(0, \infty)^n$ to $(0, \infty)$ such that $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = f\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right)$ for all worlds $a, b \in W$ and all coherent credence profiles $\mathbf{C}$ in which everyone gives non-zero probability to a and to b,*

(b) *this function satisfies $f(\mathbf{1}) = 1$ and $f(\mathbf{xy}) = f(\mathbf{x})f(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in (0, \infty)^n$ (where '$\mathbf{1}$' stands for '$(1, ..., 1)$' and '$\mathbf{xy}$' stands for '$(x_1 y_1, ..., x_n y_n)$').*

Proof. Assume Indiff* and the by Proposition 1 weakest Bayesian axiom, BayPub. I proceed in several claims (the first two of which do not require Indiff*).

Claim 1: For all $a \neq b$ in W there is a unique function $f_{a,b}$ from $(0, \infty)^n$ to $(0, \infty)$ such that $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = f_{a,b}\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right)$ for all coherent profiles $\mathbf{C}$ in which every individual assigns non-zero probability to a and to b.

Consider $a \neq b$ in W. Uniqueness of such a function $f_{a,b}$ follows from the fact that any $\mathbf{x} \in (0, \infty)^n$ can be written as $\mathbf{x} = \left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)} \right)$ for some coherent profile $\mathbf{C}$. As for existence of the function, consider coherent profiles $\mathbf{C}$ and $\mathbf{C}'$ in which a and b receive non-zero probabilities from everyone and $\frac{C_i(a)}{C_i(b)} = \frac{C_i'(a)}{C_i'(b)}$ for all i. We have to show that $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = \frac{\mathrm{ag}\mathbf{C}'(a)}{\mathrm{ag}\mathbf{C}'(b)}$. Conditionalizing everyone's credence function on $E = \{a, b\}$ transforms $\mathbf{C}$ and $\mathbf{C}'$ into the same (coherent) profile $\mathbf{C}''$, which by BayPub implies that $(\mathrm{ag}\mathbf{C})|E$ and $(\mathrm{ag}\mathbf{C}')|E$ each equal $\mathrm{ag}\mathbf{C}''$. So $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = \frac{\mathrm{ag}\mathbf{C}'(a)}{\mathrm{ag}\mathbf{C}'(b)}$, where these two ratios are well-defined and non-zero because $\mathrm{ag}\mathbf{C}(a), \mathrm{ag}\mathbf{C}(b), \mathrm{ag}\mathbf{C}'(a), \mathrm{ag}\mathbf{C}'(b) \neq 0$ by Lemma 3.

Claim 2: $f_{a,c}(\mathbf{xy}) = f_{a,b}(\mathbf{x}) f_{b,c}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in (0, \infty)^n$ and all pairwise distinct $a, b, c \in W$.

Consider $\mathbf{x}, \mathbf{y} \in (0, \infty)^n$ and pairwise distinct $a, b, c \in W$. The claimed relation follows from the definition of the functions $f_{a,b}, f_{b,c}, f_{a,c}$, because one can construct a (coherent) profile $\mathbf{C}$ for which $\mathbf{x} = \left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)} \right)$, $\mathbf{y} = \left(\frac{C_1(b)}{C_1(c)}, ..., \frac{C_n(b)}{C_n(c)} \right)$, and thus $\mathbf{xy} = \left(\frac{C_1(a)}{C_1(c)}, ..., \frac{C_n(a)}{C_n(c)} \right)$.

Claim 3: All $f_{a,b}$ for $a \neq b$ are the same function, to be denoted f. (This shows part (a) restricted to the case $a \neq b$).

Consider worlds a, a', b, b' with $a \neq b$ and $a' \neq b'$, and let $\mathbf{x} \in (0, \infty)^n$. I need to show that $f_{a,b}(\mathbf{x}) = f_{a',b'}(\mathbf{x})$. I distinguish between three cases.

Case 1: $a = a'$. Here I need to show that $f_{a,b}(\mathbf{x}) = f_{a,b'}(\mathbf{x})$. We may pick a coherent profile $\mathbf{C}$ such that $C_i(b) = C_i(b') \neq 0$ for all i and $\mathbf{x} = \left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)} \right) = \left(\frac{C_1(a)}{C_1(b')}, ..., \frac{C_n(a)}{C_n(b')} \right)$. By Lemma 4, $\mathrm{ag}\mathbf{C}(b) = \mathrm{ag}\mathbf{C}(b')$, and so $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = \frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b')}$. Hence, $f_{a,b}(\mathbf{x}) = f_{a,b'}(\mathbf{x})$.

Case 2: $b = b'$. By an argument analogous to that in Case 1, $f_{a,b}(\mathbf{x}) = f_{a',b}(\mathbf{x})$.

Case 3: $a \neq a'$ and $b \neq b'$. I show that $f_{a,b}(\mathbf{x}) = f_{a',b'}(\mathbf{x})$ by distinguishing between three subcases and drawing on Cases 1 and 2:

- If $a \neq b'$, then $f_{a,b}(\mathbf{x}) = f_{a,b'}(\mathbf{x}) = f_{a',b'}(\mathbf{x})$.

- If $a' \neq b$, then $f_{a,b}(\mathbf{x}) = f_{a',b}(\mathbf{x}) = f_{a',b'}(\mathbf{x})$.

- If $a = b'$ and $a' = b$, then, choosing any $c \in W \backslash \{a, b\}$ (by using that $|W| \geq 3$), $f_{a,b}(\mathbf{x}) = f_{a,c}(\mathbf{x}) = f_{b,c}(\mathbf{x}) = f_{b,a}(\mathbf{x})$.

Claim 4: $f(\mathbf{1}) = 1$.

By applying Claims 2–3 with $\mathbf{x} = \mathbf{y} = \mathbf{1}$, one obtains that $f(\mathbf{11}) = f(\mathbf{1}) f(\mathbf{1})$. Since $\mathbf{11} = \mathbf{1}$ it follows that $f(\mathbf{1}) = 1$.

Claim 5: For any possibly identical $a, b \in W$, $f_{a,b} \left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)} \right) = f \left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)} \right)$ for all coherent credence profiles $\mathbf{C}$ in which all C_i assign non-zero probabilities to a and b. (This essentially extends Claim 3 to the case that $a = b$.)

Consider any such $a, b, \mathbf{C}$. By definition of $f_{a,b}$ we have to show that $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = f\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right)$. In case $a \neq b$ this holds already by Claim 3. In case $a = b$ it holds by Claim 4 and the fact that $\frac{\mathrm{ag}\mathbf{C}(a)}{\mathrm{ag}\mathbf{C}(b)} = 1$ and $\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right) = \mathbf{1}$. ∎

Lemma 6. *Given the assumptions and notation of Lemma 5,*

(a) *under BayPri+, $f(x_1, ..., x_n) = x_1 \cdots x_n$ for all $(x_1, ..., x_n) \in (0, \infty)^n$ and the pooling rule is multiplicative pooling,*

(b) *under Contin, there are $w_1, ..., w_n \geq 0$ such that $f(x_1, ..., x_n) = x_1^{w_1} \cdots x_n^{w_n}$ for all $(x_1, ..., x_n) \in (0, \infty)^n$ and the pooling rule is the weighted geometric rule with weights $w_1, ..., w_n$ (in particular, $w_1 + \cdots + w_n \geq 1$ if W is infinite).*

Proof. We use the assumptions and notation of Lemma 5.

Claim 1: Under BayPri+, $f(x_1, ..., x_n) = x_1 \cdots x_n$ for all $(x_1, ..., x_n) \in (0, \infty)^n$.

Assume BayPri+ and let $(x_1, ..., x_n) \in (0, \infty)^n$. I prove by induction that $f(x_1, ..., x_i, 1, ..., 1) = x_1 \cdots x_i$ for all $i = 0, 1, ..., n$. The initial step where $i = 0$ is obvious: $f(1, ..., 1) = 1$ by Lemma 5. Now assume the claim holds for a given $i \in \{0, ..., n-1\}$, i.e., $f(x_1, ..., x_i, 1, ..., 1) = x_1 \cdots x_i$. I have to show that $f(x_1, ..., x_{i+1}, 1, ..., 1) = x_1 \cdots x_{i+1}$. Pick worlds $a \neq b$ and a coherent credence profile $\mathbf{C}$ such that everyone assigns non-zero probabilities to a and b and such that $\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right) = (x_1, ...x_i, 1..., 1)$. Let $\mathbf{C}'$ be the coherent profile arising from $\mathbf{C}$ by conditionalizing the credence function of individual $i+1$ on a likelihood function L for which $L(a), L(b) \neq 0$ and $\frac{L(a)}{L(b)} = x_{i+1}$. Note that $\frac{C'_{i+1}(a)}{C'_{i+1}(b)} = \frac{C_{i+1}(a)L(a)}{C_{i+1}(b)L(b)} = 1 \cdot x_{i+1} = x_{i+1}$. So $\left(\frac{C'_1(a)}{C'_1(b)}, ..., \frac{C'_n(a)}{C'_n(b)}\right) = (x_1, ...x_{i+1}, 1..., 1)$. Now

$$\begin{aligned}
f(x_1, ..., x_{i+1}, 1, ..., 1) &= f\left(\frac{C'_1(a)}{C'_1(b)}, ..., \frac{C'_n(a)}{C'_n(b)}\right) = \frac{\mathrm{ag}\mathbf{C}'(a)}{\mathrm{ag}\mathbf{C}'(b)} \\
&= \frac{((\mathrm{ag}\mathbf{C})|L)(a)}{((\mathrm{ag}\mathbf{C})|L)(b)} = \frac{\mathrm{ag}\mathbf{C}(a)L(a)}{\mathrm{ag}\mathbf{C}(b)L(b)} \\
&= f\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right)\frac{L(a)}{L(b)} = f(x_1, ...x_i, 1..., 1)x_{i+1} \\
&= (x_1 \cdots x_i)x_{i+1} = x_1 \cdots x_{i+1},
\end{aligned}$$

where the first equation on the second line applies BayPri+.

Claim 2: Under BayPri+, the pooling rule is the multiplicative rule.

Assume BayPri+. Let ag^* be the multiplicative rule. I show that $\mathrm{ag} = \mathrm{ag}^*$. Consider a coherent profile $\mathbf{C}$. Since BayPri+ implies BayPri and thus Bay (see Proposition 1), the group credence function $\mathrm{ag}\mathbf{C}$ assigns zero probability to worlds outside $\cap_i \mathrm{supp}(C_i)$ by Lemma 3(a). So does clearly the multiplicative group credence function $\mathrm{ag}^*\mathbf{C}$. It thus remains to show that $\mathrm{ag}\mathbf{C}$ and $\mathrm{ag}^*\mathbf{C}$ coincide on words in $\cap_i \mathrm{supp}(C_i)$, i.e., worlds to which everyone assigns non-zero probability. It suffices

25

to show that for any two such worlds a and b the probability ratio is the same both times: $\frac{\text{ag}\mathbf{C}(a)}{\text{ag}\mathbf{C}(b)} = \frac{\text{ag}^*\mathbf{C}(a)}{\text{ag}^*\mathbf{C}(b)}$. This equation holds because each side equals $\frac{C_1(a)}{C_1(b)} \cdots \frac{C_n(a)}{C_n(b)}$. Indeed, $\frac{\text{ag}\mathbf{C}(a)}{\text{ag}\mathbf{C}(b)} = \frac{C_1(a)}{C_1(b)} \cdots \frac{C_n(a)}{C_n(b)}$ by Claim 1, and $\frac{\text{ag}^*\mathbf{C}(a)}{\text{ag}^*\mathbf{C}(b)} = \frac{C_1(a)}{C_1(b)} \cdots \frac{C_n(a)}{C_n(b)}$ by definition of the multiplicative rule.

Claim 3: Under Contin, there are n numbers, henceforth denoted $w_1, ..., w_n \in \mathbb{R}$, such that $f(x_1, ..., x_n) = x_1^{w_1} \cdots x_n^{w_n}$ for all $(x_1, ..., x_n) \in (0, \infty)^n$.

Assume Contin. Define the function $g : \mathbb{R}^n \to \mathbb{R}$ by $g(\mathbf{x}) = \log(f(e^{x_1}, ..., e^{x_n}))$ for all $\mathbf{x} \in \mathbb{R}^n$. By Lemma 5(b) and the properties of the logarithm and the exponential function, it follows that $g(\mathbf{x} + \mathbf{y}) = g(\mathbf{x}) + g(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. So g obeys Cauchy's functional equation. Further, g is continuous, since f is continuous by Contin. So g is linear, i.e., there are weights $w_1, ..., w_n \in \mathbb{R}$ such that

$$g(\mathbf{x}) = w_1 x_1 + \cdots + w_n x_n \text{ for all } \mathbf{x} \in \mathbb{R}^n$$

by a fundamental theorem on functional equations (see Aczél 1966). It follows that

$$f(\mathbf{x}) = e^{g(\log x_1, ..., \log x_n)} = e^{\log(x_1^{w_1} \cdots x_n^{w_n})} = x_1^{w_1} \cdots x_n^{w_n} \text{ for all } \mathbf{x} \in (0, \infty)^n.$$

Claim 4: Under Contin, for each full-support profile $\mathbf{C}$ (i.e., each profile in which everyone assigns non-zero probability to all worlds) there is a constant $k > 0$ such that $\text{ag}\mathbf{C}(a) = k[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n}$ for all worlds a. (This 'almost' shows that ag is a weighted geometric rule, except that we only quantify over full-support profiles and have not proved that $w_1, ..., w_n$ are non-negative.)

Assume Contin. Consider a full-support profile $\mathbf{C}$. Fix a world $b \in W$. Define the constants $k' = \text{ag}\mathbf{C}(b)$ and $k'' = [C_1(b)]^{w_1} \cdots [C_n(b)]^{w_n}$. Note that $k', k'' > 0$ (using that $\text{ag}\mathbf{C}$ has full support by Lemma 3(b)). For all worlds a,

$$\text{ag}\mathbf{C}(a) = k' \frac{\text{ag}\mathbf{C}(a)}{\text{ag}\mathbf{C}(b)} = k' f\left(\frac{C_1(a)}{C_1(b)}, ..., \frac{C_n(a)}{C_n(b)}\right)$$
$$= k' \left(\frac{C_1(a)}{C_1(b)}\right)^{w_1} \cdots \left(\frac{C_n(a)}{C_n(b)}\right)^{w_n} = \frac{k'}{k''}[C_1(a)]^{w_1} \cdots [C_n(a)]^{w_n},$$

where the first equation on the second line holds by Claim 3. This show Claim 4 with $k = \frac{k'}{k''}$.

Claim 5: Under Contin, $w_1, ..., w_n \geq 0$.

Assume Contin. Suppose for a contradiction that i is an individual such that $w_i < 0$. Consider a world $a \in W$, and a sequence of full-support profiles $\mathbf{C}^k$ ($k = 1, 2, ...$) converging to a credence profile $\mathbf{C}$ in which C_i has support $W \setminus \{a\}$ and each C_j with $j \neq i$ has full support W. By the fact that $w_i < 0$ and Claim 4, $\text{ag}\mathbf{C}^k$ converges to the probability measure assigning probability one to a. This is because $[C_1^k(a)]^{w_1} \cdots [C_n^k(a)]^{w_n}$ tends to infinity (the term $[C_i(a)]^{w_i}$ tends to infinity) while for all other worlds $b \neq a$ $[C_1^k(b)]^{w_1} \cdots [C_n^k(b)]^{w_n}$ tends to a finite value. Meanwhile by Contin $\text{ag}\mathbf{C}^k$ also converges to $\text{ag}\mathbf{C}$. It follows that $\text{ag}\mathbf{C}(a) = 1$. So the support

26

of ag**C** is $\{a\}$. This contradicts the fact that the support of ag**C** must include the intersection of supports $\cap_m \text{supp}(C_m) = W \backslash \{a\}$ by Lemma 3.

Claim 6: Under Contin, ag is the weighted geometric rule with weights $w_1, ..., w_n$.

Assume Contin. By Claim 4, ag coincides with this weighted geometric rule on the subdomain of full-support profiles. This subdomain is *dense* in the full domain of coherent profiles: every coherent profile is the limit of some sequence of full-support profiles, as readers can easily check. Since ag and the weighted geometric rule with weights $w_1, ..., w_n$ are two *continuous* rules on the domain of coherent profiles which coincide on a dense subdomain, the two rules coincide globally. ∎

B.4 Completing the theorems' proofs

Proof of Theorem 1. *First*, any weighted geometric rule whose weights are all positive satisfies Bay by Lemma 1(a) and satisfies Contin and Indiff* (and under finite W Indiff) by Lemma 2. *Conversely*, if a rule for aggregating coherent profiles satisfies Bay, Contin and Indiff*, then by Lemma 6(b) it is a weighted geometric rule, where by Lemma 1(a) the weights are all positive. ∎

Proof of Theorem 2. *First*, any weighted geometric rule with at least one positive weight satisfies BayPub by Lemma 1(b), as well as Contin and Indiff* (and under finite W Indiff) by Lemma 2. *Conversely*, if a rule for aggregating coherent profiles satisfies BayPub, Contin and Indiff*, then by Lemma 6(b) it is a weighted geometric rule, where by Lemma 1(b) some weight is positive. ∎

Proof of Theorem 3. This result follows from Theorem 1 via Proposition 1. ∎

Proof of Theorem 4. This result follows from Proposition 2, as Bay+ implies BayPub+ and BayPri+. ∎

Proof of Theorem 5. *First*, each weighted geometric rule whose weights sum to one satisfies BayPub+ by Lemma 1(c), and also Contin and Indiff* (and under finite W Indiff) by Lemma 2. *Conversely*, if a rule for aggregating coherent profiles satisfies BayPub+, Contin and Indiff*, then by Lemma 6(b) it is a weighted geometric rule, where by Lemma 1(c) the weights sum to one. ∎

Proof of Theorem 6. *First*, the multiplicative rule satisfies BayPri+ by Lemma 1(d), and satisfies Indiff* (and under finite W Indiff) by Lemma 2. *Conversely*, if a rule for pooling coherent profiles satisfies BayPri+ and Indiff*, then by Lemma 6(a) it is the multiplicative rule. ∎

Proof of Theorem 7. Let $\mathcal{D}$ be the domain of all coherence profiles, and $\mathcal{D}'$ the subdomain of all coherent credence profiles. I prove the three claims in a different order.

(b) *First*, any power dictatorship satisfies BayPub, Contin and Indiff* (and Indiff if W is finite). The argument is similar to that given for weighted geometric rules; it suffices to adapt Lemmas 1 and 2. *Conversely*, consider a rule ag defined on $\mathcal{D}$ and satisfying BayPub, Contin and Indiff*. Let ag′ be its restriction to $\mathcal{D}'$. Check that ag′ still satisfies the three axioms. So it must by Theorem 2 be a weighted geometric rule whose weights $w_1, ..., w_n$ are not all zero. I consider two cases.

Case 1: only one individual, say individuals i, has non-zero weight w_i. Then ag is the power dictatorship with power dictator i and power w_i, because (i) ag coincides with this power dictatorship on the subdomain $\mathcal{D}'$ which (as one may check) is dense in $\mathcal{D}$, and (ii) ag and the power dictatorship are continuous rules.

Case 2: at least two individuals, say individuals i and j, have non-zero weights. I derive a contradiction. Fix two worlds $a \neq b$, and consider profiles $\mathbf{C}^k$ ($k = 1, 2, ...$) in which i's credences are given by $C_i^k(a) = 2^{-k}$ and $C_i^k(b) = 1 - 2^{-k}$, j's credences are given by $C_j^k(a) = 1 - 2^{-k^2}$ and $C_j^k(b) = 2^{-k^2}$, and any other member m's credences are given by $C_m^k(a) = C_m^k(b) = \frac{1}{2}$. As $\mathbf{C}^k$ is coherent, ag$\mathbf{C}^k$ is given by weighted geometric pooling, so that ag$\mathbf{C}^k(c) = 0$ for worlds $c \neq a, b$ and

$$\frac{\text{ag}\mathbf{C}^k(a)}{\text{ag}\mathbf{C}^k(b)} = \frac{[C_i^k(a)]^{w_i}[C_j^k(a)]^{w_j}}{[C_i^k(b)]^{w_i}[C_j^k(b)]^{w_j}} = \frac{2^{-kw_i}(1 - 2^{-k^2})^{w_j}}{(1 - 2^{-k})^{w_i}2^{-k^2 w_j}} = 2^{k^2 w_j - kw_i}\frac{(1 - 2^{-k^2})^{w_j}}{(1 - 2^{-k})^{w_i}},$$

which converges to ∞. So ag$\mathbf{C}^k$ converges to the credence function assigning probability one to a.

Now construct another sequence of profiles $\mathbf{D}^k$ ($k = 1, 2, ...$), in which $\mathbf{D}^k$ is defined like $\mathbf{C}^k$ except that the roles of k and k^2 are interchanged: so $D_i^k(a) = 2^{-k^2}$, $D_i^k(b) = 1 - 2^{-k^2}$, $D_j^k(a) = 1 - 2^{-k}$, $D_j^k(b) = 2^{-k}$, and $D_m^k(a) = D_m^k(b) = \frac{1}{2}$ for all members $m \neq i, j$. Applying the weighted geometric formula again, we find that ag$\mathbf{D}^k(c) = 0$ for worlds $c \neq a, b$ and that $\frac{\text{ag}\mathbf{D}^k(a)}{\text{ag}\mathbf{D}^k(b)}$ converges to 0 rather than ∞. So ag$\mathbf{D}^k$ converges to the credence function assigning probability one to b rather than a.

Meanwhile, as one easily checks, the profiles $\mathbf{C}^k$ and $\mathbf{D}^k$ both converge to a same limiting profile $\mathbf{C}$ (in which $C_i(b) = 1$, $C_j(a) = 1$, and $C_m(a) = C_m(b) = \frac{1}{2}$ for members $m \neq i, j$). So ag$\mathbf{C}^k$ and ag$\mathbf{D}^k$ both converge to ag$\mathbf{C}$ by Contin. This contradicts the fact that ag$\mathbf{C}^k$ and ag$\mathbf{D}^k$ converge to different credence functions.

(c) *First*, any dictatorship satisfies BayPub+, Contin and Indiff* (and Indiff if W is finite). The argument is again similar to that for weighted geometric rules. *Conversely*, consider a rule ag on $\mathcal{D}$ satisfying BayPub+, Contin and Indiff*. Its restriction to $\mathcal{D}'$, denoted ag′, still satisfies these axioms. So it must by Theorem 5 be a weighted geometric rule whose weights $w_1, ..., w_n$ sum to one. There are two cases.

Case 1: only one individual i has non-zero weight, hence weight one. Then ag is the dictatorship by individual i, by the same continuity argument as under Case 1

above.

Case 2: more than one individual has non-zero weight. Then a contradiction can be derived by an argument parallel to that under Case 2 above.

(a) Consider a rule ag on $\mathcal{D}$ satisfying the axioms in Theorem 1, 3, 4 or 6. Its restriction to $\mathcal{D}'$, denoted ag$'$, still satisfies these axioms. In the case of the axioms of Theorem 4 this already is a contradiction. In the case of the axioms of Theorem 1, 3 or 6, it follows by the theorem that ag$'$ is a weighted geometric rule whose weights $w_1, ..., w_n$ are all non-zero. This implies a contradiction, just as under Case 2 in the proofs of (b) and (c). $\blacksquare$

References

[1] Aczél, J. (1966) *Lectures on Functional Equations and their Applications*, New York and London: Academic Press

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

[3] Dietrich, F., List, C. (2016) Probabilistic Opinion Pooling. In: C. Hitchcock & A. Hájek (eds.), *Oxford Handbook of Probability and Philosophy*, Oxford University Press

[4] Genest, C., Zidek, J. V. (1986) Combining Probability Distributions: A Critique and Annotated Bibliography, *Statistical Science* 1: 114-135

[5] Glymour, C. (l980) *Theory and Evidence*, Princeton: Princeton University Press

[6] Hájek, A. (2003) What conditional probability could not be, *Synthese* 137: 273–323

[7] Hartmann, S., Fitelson, B. (forth.) A new Garber-style solution to the problem of old evidence, *Philosophy of Science*

[8] Joyce, J. M. (1999) *The Foundation of Causal Decision Theory*, Cambridge: Cambridge University Press

[9] Leitgeb, H. (forth.) Imaging all the people, *Episteme*

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

[11] Madansky, Albert (1964) Externally Bayesian Groups, Technical Report RM-4141-PR, RAND Corporation

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

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

[14] Steele, J. M. (2004) *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*, MAA Problem Books Series, Cambridge University Press