JEFFREY POOLING

Richard Pettigrew, Jonathan Weisberg

University of Bristol
University of Toronto

Suppose you are 40% confident that Candidate X will win in the upcoming election. Then you read a column projecting 80%. If you and the columnist are equally well informed and competent on this topic, how should you revise your opinion in light of theirs? Should you perhaps split the difference, arriving at 60%?

Plenty has been written on this topic.[1] Much less studied, however, is the question what comes next. Once you've updated your opinion about Candidate X, how should your other opinions change to accommodate this new view? For example, how should you revise your expectations about other candidates running for other seats? Or your confidence that your preferred party will win a majority?

A natural response is: by Jeffrey conditionalizing (Jeffrey, 1965).[2] When you change your probability for E from $P(E)$ to $P'(E) = x$, Jeffrey conditionalization adjusts your other opinions as follows:

$$P'(H) = P(H \mid E) \cdot x + P(H \mid \overline{E}) \cdot (1 - x).[3]$$

In our example, E is the proposition that Candidate X will win their election, and H is any other proposition, *e.g.* that your party will win a majority. If you split the difference with the columnist, then $x = .6$. So you plug this number into Jeffrey's equation and, together with your existing opinions about H given E and given $\overline{E}$, it determines your new probability $P'(H)$ that your party will win a majority.

Now suppose you read a different column, about another candidate running for a different seat. In light of the opinion expressed there, you update your confidence in the relevant proposition F to some new probability $P''(F) = y$. Then you apply Jeffrey conditionalization again,

1. For some background see Christensen (2007, 2009), Elga (2007), Kelly (2010), Dietrich and List (2016), and Easwaran et al. (2016).
2. See Wagner (2011) and Easwaran et al. (2016) for some prior discussion of this proposal. See also Roussos (2021) for a related model.
3. That is, you retain your credences in H conditional on E and on $\overline{E}$, and you use your new unconditional credences in E and $\overline{E}$, together with the Law of Total Probability, to calculate your new credence in H.

to update your opinions on other matters accordingly:

$$P''(H) = P'(H \mid F) \cdot y + P'(H \mid \overline{F}) \cdot (1 - y).$$

A natural thought now is that the order shouldn't matter here. Which column you read first is irrelevant. Either way, you have the same total information in the end, so your ultimate opinions should be the same.

This requirement is known as *commutativity*, and we will show that it strongly favours one particular way of merging your 40% with the columnist's 80%. Rather than splitting the difference to give 60%, you should use another formula: "upco", also known as "multiplicative pooling." Given some neutral assumptions, this is the only way of combining probabilities that ensures Jeffrey conditionalization delivers the same final result, no matter which opinion you encounter first. And the difference between upco and difference-splitting can be striking: upco combines 40% and 80% to give a new credence of about 73%, rather than 60%.

But let's first address the elephant in the room: why not simply conditionalize? You've learned that the columnist is 80% confident X will win, so shouldn't you just conditionalize on the fact that they hold that opinion? Well, you should, if you can. But the "just conditionalize" answer still isn't fully satisfactory, for two reasons.

First, it's incomplete. After all, you may not have the prior credences, conditional and unconditional, that conditionalizing requires.[4] Perhaps you just haven't given the columnist's opinion and its evidential weight much thought until now. Second, even if you have the relevant priors, the computations needed to conditionalize can be very demanding, especially if you are using Bayes' Theorem for a large partition. It's much easier to apply a simple formula like splitting the difference, and then Jeffrey conditionalize on the result. Indeed, this corresponds to a

---

4. Bayesian writers often assume priors for all propositions an agent might learn. But here we are addressing the part of the Bayesian tradition where this assumption is relaxed; see *e.g.* Jeffrey (1983) and Easwaran *et al* (2016).

natural and intuitive way to break the problem up into two pieces: (i) how should I revise my opinion about Candidate X's prospects, and (ii) how should my other views change in light of the first change?

What's more, this two step analysis is actually equivalent to conditionalization in many cases. Suppose the columnist's opinion about Candidate X is only relevant to other matters insofar as it's relevant to whether X wins or not. More precisely, suppose that conditional on X winning, other matters are independent of the columnist's opinion (and likewise conditional on X not winning). Then, revising all your opinions by conditionalization is equivalent to the two step process of first revising your opinion about *E* by conditionalization, and then revising your remaining opinions by Jeffrey conditionalization.[5]

For multiple reasons then, we would like to know how your opinion about Candidate X might be combined with the columnist's, such that the result can be sensibly plugged into Jeffrey conditionalization. We'll show that one way of performing this combination is uniquely privileged.

---

5. Formally, if the partition $\{E, \overline{E}\}$ screens off H from E^*, and we let $x = P(E \mid E^*)$, then

$$P(H \mid E^*) = P(H \mid E) \cdot x + P(H \mid \overline{E}) \cdot (1 - x).$$

To see why, first recall what it means for $\{E, \overline{E}\}$ to screen off H from E^*:

$$P(H \mid EE^*) = P(H \mid E),$$
$$P(H \mid \overline{E}E^*) = P(H \mid \overline{E}).$$

Then apply the law of total probability to $P(H \mid E^*)$, and substitute x for $P(E \mid E^*)$:

$$P(H \mid E^*) = P(H \mid EE^*)P(E \mid E^*) + P(H \mid \overline{E}E^*)P(\overline{E} \mid E^*)$$
$$= P(H \mid E) \cdot x + P(H \mid \overline{E}) \cdot (1 - x).$$

1. Upco Ensures Jeffrey Pooling Commutes

Splitting the difference between two opinions is known as *linear pooling*. The formula is just the familiar arithmetic mean:

$$P'(E) = \frac{P(E) + Q(E)}{2},$$

where $P(E)$ is your prior opinion about E, before reading any columns, and $Q(E)$ is the columnist's probability. In our example $P(E) = 0.4$ and $Q(E) = 0.8$, so $P'(E) = 0.6$.

But we'll see that commutativity instead favours *upco*, also known as *multiplicative pooling*:

$$P'(E) = \frac{P(E)Q(E)}{P(E)Q(E) + P(\overline{E})Q(\overline{E})}. \tag{1}$$

If $P(E) = 0.4$ and $Q(E) = 0.8$, then $P'(E) \approx 0.73$, significantly larger than the 0.6 recommended by linear pooling.

These two formulas are examples of *pooling rules*, functions that take two probabilities $P(E)$ and $Q(E)$ and return a new probability $P'(E)$. Two more examples come from the other notions of 'mean' included in the classical trio of Pythagorean means: the geometric and harmonic means (Genest and Zidek, 1986). And there are many more, too many to name.

Our question is how these various rules behave when coupled with Jeffrey conditionalization. Suppose we begin with P, fix some pooling rule f, and use the following two-step procedure for responding to Q's opinion about E.

Jeffrey Pooling:

Step 1. Apply pooling rule f to $P(E)$ and $Q(E)$ to obtain $P'(E)$:

$$P'(E) = f(P(E), Q(E)).$$

Step 2. Revise all other credences by Jeffrey conditionalization:

$$P'(H) = P(H \mid E) P'(E) + P(H \mid \overline{E}) (1 - P'(E)).$$

We will call this *Jeffrey pooling P with Q on E using f*. But that's a mouthful, so we'll often leave some of these parameters implicit when context permits. We'll say that *f ensures Jeffrey pooling commutes* if, for any P, Q, and R, Jeffrey pooling P with Q on E and then Jeffrey pooling the result with R on F, has the same final result as Jeffrey pooling P with R on F and then Jeffrey pooling the result with Q on E.

Upco ensures that Jeffrey pooling commutes, as long as the necessary operations are defined. Zeros can gum up the works in two ways. First, if $P(E) = 1$ and $Q(E) = 0$ or vice versa, then Step 1 fails: upco cannot be applied, because its denominator is 0. Second, the conditional probabilities used in Step 2 need to be defined, so $P(E)$ cannot be either 0 or 1. For a subsequent update on F to have defined conditional probabilities as well, we also need the updated probability of F to be non-extreme.

To avoid these difficulties, we will temporarily make the simplifying assumption that P is *regular*, *i.e.* that it assigns positive probability to EF, $E\overline{F}$, $\overline{E}F$, and $\overline{E}\,\overline{F}$. This ensures no problematic zeros arise when Jeffrey pooling on E and then F, or vice versa. In the Appendix we show that this assumption can be dropped; the result we are about to present holds whenever the relevant Jeffrey pooling operations are defined, even if P is not regular.

If P is regular, then upco is sufficient to make Jeffrey pooling commutative. We attribute this result to Field (1978) for reasons that will become clear in Section 3.

Theorem 1 (Field). *Upco ensures that Jeffrey pooling commutes for any regular P, and any Q and R.*

In the Appendix we generalize this result to pooling over countable partitions, *i.e.* to cases where we don't just hear Q's opinion about E, but about every element in a countable partition.

	EF	$E\overline{F}$	$\overline{E}F$	$\overline{E}\,\overline{F}$
$P'(-)$	9/25	2/25	6/25	8/25
$P''(-)$	18/29	4/29	3/29	4/29

As Theorem 1 claimed, the ultimate posterior P'' is the same either way.

This convergence may seem magical, but its inevitability emerges if we look past the denominators to the relative proportions. We began with the proportions $3:1:2:4$. Then we multiplied the first two entries by 8 and the last two by 2, since $Q(E) = 8/10$ and $Q(\overline{E}) = 2/10$. This gave us for P' the relative proportions $24:8:4:8$, although we divided through by the common factor 4 to write this as $6:2:1:2$. Then, because $R(F) = 6/10$ and $R(\overline{F}) = 4/10$, we multiplied the first and third entries by 6 and the second and fourth by 4, to get $36:8:6:8$—although again we divided through by a common factor to write this as $18:4:3:4$.

Updating in the the opposite order, we began again with $3:1:2:4$ but multiplied the first and third entries by 6, the second and fourth by 4, to get $18:4:12:16$, which reduced to $9:2:6:8$. Then we multiplied the first two entries by 8 and the last two by 2, to get $72:16:12:16$, or $18:4:3:4$.

In both cases the final proportions had to be the same because, ultimately, all we did was multiply the values 3, 1, 2, and 4 by the values $8 \cdot 6$, $8 \cdot 4$, $2 \cdot 6$, and $2 \cdot 4$, respectively (then divide the results by a common factor). We can think of this as multiplying by the values 8, 8, 2, 2 first, and by 6, 4, 6, 4 second, or the other way around. The commutativity of Jeffrey pooling with upco follows by the commutativity of multiplication.

2. Only Upco Ensures Jeffrey Pooling Commutes

While upco ensures that Jeffrey pooling commutes, linear pooling doesn't; nor do geometric and harmonic pooling. Indeed, among the pooling rules that boast four plausible properties—properties the rules just named all share—upco is the only one that ensures this. As we will indicate in the course of introducing these properties, we don't think

An example makes clear why Theorem 1 is true. Recall the case we opened with, where $P(E) = 4/10$ and $Q(E) = 8/10$. Let's further suppose that $R(F) = 6/10$, and that P has the following details:

	EF	$E\overline{F}$	$\overline{E}F$	$\overline{E}\,\overline{F}$
$P(-)$	3/10	1/10	2/10	4/10

According to Theorem 1, P's final opinions will be the same whether they Jeffrey pool with Q first and R second, or vice versa, provided they use upco for the first step in Jeffrey pooling.

Begin with the case where P pools with Q first. Step 1 of Jeffrey pooling combines $P(E) = 4/10$ with $Q(E) = 8/10$ via upco, to yield $P'(E) = 8/11$. For Step 2, the key is to observe that the relative proportions of EF and $E\overline{F}$ must be preserved—this is Jeffrey conditionalization's oft-noted "rigidity". So the 8/11 assigned to E must be divided 3-to-1 between EF and $E\overline{F}$. Similarly, the 3/11 assigned to $\overline{E}$ gets divided 1-to-2 between $\overline{E}F$ and $\overline{E}\,\overline{F}$. The posterior P' that results is:

	EF	$E\overline{F}$	$\overline{E}F$	$\overline{E}\,\overline{F}$
$P'(-)$	6/11	2/11	1/11	2/11

Now we pool P' with R using similar reasoning. Applying upco to $P'(F) = 7/11$ and $R(F) = 6/10$ gives $P''(F) = 21/29$. Jeffrey conditionalization then divides this up proportionally to arrive at:

	EF	$E\overline{F}$	$\overline{E}F$	$\overline{E}\,\overline{F}$
$P''(-)$	18/29	4/29	3/29	4/29

In the case where P pools with R first and Q second, parallel calculations give the following sequence instead:

they will be desirable in all situations. But we do claim that they are desirable in a great many important ones. And in those cases, upco is the only rule that delivers.

The first property is *monotonicity*: if we fix $Q(E) = 1/2$, then as $P(E)$ increases, so does $P'(E)$. This is a familiar feature of linear pooling, and upco has it too.[6] Notice that this is also a feature of conditionalization in many cases. For any proposition Q, conditionalization sets $P'(E) = P(E \mid Q)$, which Bayes' theorem renders

$$P'(E) = \frac{P(E)P(Q \mid E)}{P(Q \mid E)P(E) + P(Q \mid \overline{E})P(\overline{E})}.$$

If the likelihood terms $P(Q \mid E)$ and $P(Q \mid \overline{E})$ stay fixed as $P(E)$ changes, then $P'(E)$ increases with $P(E)$.[7]

The second property our argument will rely on is *uniformity preservation*: if $P(E) = Q(E) = 1/2$, then $P'(E) = 1/2$ too. Crudely put, two empty heads are no better than one. A bit less crudely, if neither party has an opinion about the question at hand, then combining their opinions doesn't change this. There are conceivable cases where this feature would be undesirable. For example, the fact that both parties are so far ignorant about a question could indicate a conspiracy to keep everyone in the dark. But such cases are the exception rather than the rule.

Third is *continuity*: in nearly all cases, if we fix $Q(E)$ and let $P(E)$ approach a value c, then the pool of c and $Q(E)$ should be the limit of the pools of $P(E)$ and $Q(E)$ as $P(E)$ approaches c. Nearly all? Yes, because we have to ensure that all of the pools just mentioned are defined. So we restrict to cases in which, as $P(E)$ approaches c, the pool

of $P(E)$ and $Q(E)$ is always defined, and the pool of c and $Q(E)$ is as well.

To illustrate continuity, fix $Q(E) = 1/2$ and consider what happens in linear pooling as $P(E)$ decreases to 0. As $P(E)$ gets smaller, the value of $P'(E)$ gets closer and closer to 1/4. And, indeed, that is the value $P'(E)$ takes when $P(E)$ finally does reach 0. There is no sudden jump in the value of $P'(E)$ when $P(E)$ finally hits 0.

As with uniformity preservation, there are conceivable cases where this feature would not be appropriate. These might arise if we were to think that some probabilities have a particular significance. For instance, a Lockean might think there is a probabilistic threshold beyond which you count as believing the proposition to which you assign the probability, but below which you don't. And they might think that sudden change in doxastic status should be reflected in our pooling rule—perhaps your probability gains more weight when it suddenly becomes a belief. We'll assume this isn't the case.

Our fourth property is *symmetry*: swapping the values of $P(E)$ and $Q(E)$ makes no difference to $P'(E)$. This is perhaps the most restrictive feature, since exceptions are commonplace. When one party is more competent or better informed than the other, it matters who holds which opinion. Frequently we will want to give more "weight" to $P(E)$ than to $Q(E)$, or vice versa, in which case exchanging their values *should* make a difference.

But our argument only concerns cases where this is not so: cases where the two parties are equally competent and well informed on the topic.[8] When *e.g.* one party has more information, upco may not be appropriate (although in some cases it will be appropriate even then).

There are asymmetrically weighted versions of the various pooling rules we've mentioned, which may be appropriate to such cases. But we won't address these cases here. If we can show that upco is specially

---

6. In fact this property holds for any fixed value of $Q(E)$ other than 0 and 1. But we only need the minimal assumption that it holds for $Q(E) = 1/2$.
7. The derivative with respect to p of $(pq_1)/(pq_1 + (1-p)q_2)$ is positive if $q_1, q_2 > 0$.

8. See Elga (2007) for a defense of the idea that the views of peers should be given "equal weight". See Fitelson and Jehle (2009) for some formal background on articulating the view.

suited when symmetry is appropriate, that will be a significant step forward. Not to mention a strong indicator that a weighted version of upco would be the way to go in some asymmetric cases.

Finally, there's an assumption implicit in the very idea of a pooling rule, which we should pause to examine. Since a pooling rule is a function of $P(E)$ and $Q(E)$ and nothing else, we are assuming from the outset that $P(E)$ and $Q(E)$ are the only factors relevant to $P'(E)$. But other of P's opinions could be relevant, such as their opinion about what evidence $Q(E)$ is based on. Even the fact that it's an opinion about the proposition E, and not some other proposition, could be relevant. Someone might be competent on the topic of E but incompetent on the topic of F. In which case you might apply one formula when faced with their opinion about E, but use another should they opine about F.

So there is a tacit fifth assumption here, which we might call *extensionality*. By assuming extensionality, however, we are not assuming that there is one pooling rule appropriate to all circumstances, regardless of your background beliefs or the content of the question under discussion. On the contrary, different rules will be suited to different circumstances. But the question we are asking is: which rules are suited to circumstances where the above four conditions hold, Jeffrey conditionalization is appropriate, and the order in which sources are consulted should not matter.

In answer to this question, we offer the following result.

Theorem 2. *Among the monotonic, continuous, uniformity preserving, and symmetric pooling rules, only upco ensures that Jeffrey pooling commutes for any regular P, and any Q and R.*

As we noted in connection with Theorem 1, upco ensures Jeffrey pooling commutes even when P is not regular, provided the relevant operations are defined. But Theorem 2 tells us no other pooling rule can claim this feature, even if we restrict our attention to regular P.

It's important to appreciate what this result does not say: it does not tell us that rules like linear pooling *never* commute. It is possible to get lucky with linear pooling and encounter two sources where the

order doesn't matter. For example, suppose Q already agrees with P about E, and R agrees with P about F. Then, linear pooling will keep P's opinion fixed throughout. Whichever order they encounter Q and R in, their opinion at the end will be the same as when they started. But Theorem 2 tells us this can't be counted on to hold generally; only upco is commutative regardless of the particulars of P, Q, and R.

It's also important to recognize that there are cases where the order *should* matter. For example, imagine you're interviewing pundits instead of reading pre-written opinion columns. And pundit Q can be counted on for a serious opinion if you consult them first, but they'll be so insulted if you talk to R first that they'll lose their cool and adopt wild views. Then it really matters what order you hear their opinions in.

But again, we do not mean to argue that upco is always the best rule. Rather, we aim to show that upco is the only rule that will serve in all cases where the assumptions we've laid out are reasonable. And one of those assumptions is that the order shouldn't matter.

That completes our argument for upco. We now turn to locating Theorems 1 and 2 in the context of existing work on Jeffrey conditionalization and commutativity. In Section 3, we show a surprising and illuminating connection with an early result due to Field (1978). Then, in Section 4, we explain how Wagner's (2002) theorems relate.

3. Testimony of the Senses

Field (1978) was the first to identify conditions that make Jeffrey conditionalization commutative. How does his discovery fit with our results, especially Theorem 1?

Field discusses cases where sensory experience, rather than another person's opinion, prompts the shift from $P(E)$ to $P'(E)$. He assumes that each experience has an associated proposition E and number $\beta \geq 0$, where β reflects how strongly the experience speaks in favour of E.[9]

---

9. Field actually uses a log scaled version of β, which he labels α. He then reformulates Jeffrey conditionalization using exponentials, to invert the logs. We've removed these scaling features to make the connection with upco more transparent.

Field's proposal is that we should respond to sense experience by the following two-step procedure.

Field Updating:

Step 1. Update from $P(E)$ to $P'(E)$ using β as follows:

$$P'(E) = \frac{\beta \cdot P(E)}{\beta \cdot P(E) + P(\overline{E})}. \tag{2}$$

Step 2. Update other credences by Jeffrey conditionalization:

$$P'(H) = P(H \mid E) \, P'(E) + P(H \mid \overline{E}) \, (1 - P'(E)).$$

We will call this procedure *Field updating on* (E, β). Field shows that his procedure is commutative: Field updating on (E, β_1) and then (F, β_2) has the same result as Field updating on (F, β_2) followed by (E, β_1).

This may sound familiar. And if you squint, you might see that Field's Equation (2) is actually the same as upco's Equation (1). It's just that β is on the odds scale from 0 to ∞, rather than the probability scale from 0 to 1. To convert from odds to probabilities, we can divide through by $\beta + 1$, in both the numerator and the denominator:

$$P'(E) = \frac{\frac{\beta}{\beta+1} \cdot P(E)}{\frac{\beta}{\beta+1} \cdot P(E) + \frac{1}{\beta+1} \cdot P(\overline{E})}. \tag{3}$$

And this is the same as Equation (1), where Q's probabilities are $Q(E) = \beta/(\beta+1)$ and $Q(\overline{E}) = 1/(\beta+1)$.

So, formally speaking, Field updating is the same thing as Jeffrey pooling with upco. And Theorem 1 is just a restatement of Field's classic result.

This formal parallel suggests two helpful heuristics for thinking about Field's way of responding to sensory experience.

First, we might think of Equation (3) as pooling your prior opinion with a "naive" opinion proposed by your sensory system. Notice that, when $P(E) = P(\overline{E})$, Equation (3) delivers $P'(E) = \beta/(\beta+1)$. So if you have no prior opinion about E, you will defer to your sensory system's proposal, $\beta/(\beta+1)$. We can thus think of β as the odds your sensory system recommends based on the experience alone, absent any prior information.

However, when you do have a prior opinion about E, the naive recommendation has to be merged with it. Field's proposal is to use upco to combine the naive recommendation with your prior opinion, which makes updates commutative under Jeffrey conditionalization. Indeed, Theorem 2 shows that Field's proposal is the only way to do this using a monotonic, continuous, uniformity preserving, and symmetric pooling rule.

A second way of understanding Field's proposal exploits a formal analogy between upco and Bayes' theorem. Notice that Equation (3) just is Bayes' theorem, if we think of the β terms not as unconditional probabilities, but as *likelihoods*. That is, imagine we are calculating $P'(E) = P(E \mid E^*)$ for some proposition E^*. If the likelihoods are $P(E^* \mid E) = \beta/(\beta+1)$ and $P(E^* \mid \overline{E}) = 1/(\beta+1)$, then Equation (3) is just Bayes' theorem.

What is the proposition E^* here? Let E^* describe all epistemically relevant features of the experience prompting the update. The original motivation for Jeffrey conditionalization was that you may not be able to represent E^* at the doxastic level—or maybe you can, but you don't have any priors involving E^*, because it's too subtle or specific. So you can't conditionalize, because $P(E \mid E^*)$ is undefined.

But we can extend P to a compatible distribution P^+ that does

encompass E^*, by stipulating

$$P^+(E^* \mid E) = \beta/(\beta + 1),$$
$$P^+(E^* \mid \overline{E}) = 1/(\beta + 1).$$

Then Equation (3) becomes conditionalization via Bayes' theorem:

$$P'(E) = \frac{P^+(E)P^+(E^* \mid E)}{P^+(E)P^+(E^* \mid E) + P^+(\overline{E})P^+(E^* \mid \overline{E})}.$$

So this interpretation conceives of Field's proposal as conditionalizing on the ineffable but epistemically essential qualities of sensory experience, by relying on the sensory system to do the effing and the expecting—*i.e.* to represent the experience's epistemically relevant features, and supply the likelihood values Bayes' theorem requires.

4. Wagner's Theorems

There's also an important connection between our Theorem 2 and a classic result about Jeffrey conditionalization due to Wagner (2002).

Wagner analyzes Jeffrey conditionalization in terms of "Bayes factors." When we update a probability distribution from P to P', the *Bayes*

factor of E is the ratio of its new odds to its old odds:[10]

$$\frac{P'(E)/P'(\overline{E})}{P(E)/P(\overline{E})}.$$

Crudely put, Wagner's insight is that Jeffrey conditionalization commutes when, and pretty much only when, the Bayes factors are consistent regardless of the order. This needs some explaining.

Suppose two agents begin with the same prior distribution, P. Then they update as in Figure 1. That is, one does a Jeffrey conditionalization update on E that yields a Bayes factor of B_1^E, followed by another on F that yields a Bayes factor of B_1^F. The second agent starts with a Jeffrey conditionalization update on F that yields the Bayes factor B_2^F, then does a second on E that yields the Bayes factor B_2^E. At the end of this process, we label their posteriors P_1'' and P_2'', respectively.

Wagner's first result is that the two agents will end up with the same ultimate posterior if the Bayes factors for their respective E updates are the same, and likewise for their F updates. As before we will assume regularity to ensure everything is defined.[11]

Theorem 3 (Wagner). *In the schema of Figure 1, if P is regular, then*

---

10. Usually, Bayes factors are used to compare two competing models, M_1 and M_2, in light of some data, D. The Bayes factor is defined as the ratio of likelihoods, $P(D \mid M_1)/P(D \mid M_2)$. Using Bayes' theorem, this can be rewritten

$$\frac{P(M_1 \mid D) / P(M_2 \mid D)}{P(M_1)/P(M_2)}.$$

Wagner is applying the same idea, with E in the role of M_1 and $\overline{E}$ in the role of M_2. Except that there is no data D being conditioned on; instead, the posterior probabilities in the numerator are arrived at by Jeffrey conditionalization.

11. Wagner uses a milder assumption than regularity, but for simplicity we'll continue to assume P is regular.
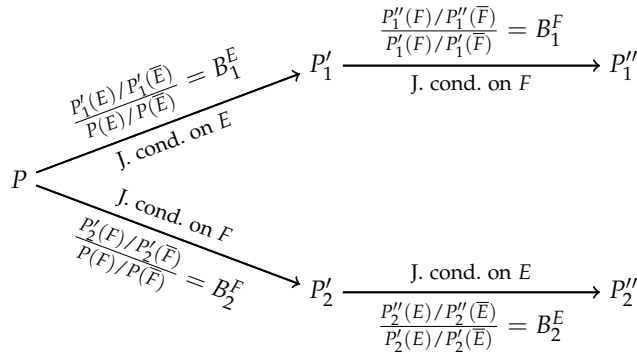
Figure 1: The context for Wagner's Theorems 3 and 4

$B_1^E = B_2^E$ and $B_1^F = B_2^F$ *together imply* $P_1'' = P_2''$.

Loosely speaking, Bayes factor "consistency" is sufficient for Jeffrey conditionalization updates to commute.

Field updating produces exactly this sort of consistency. We can verify with a bit of algebra that a given input value β always yields the same Bayes factor. In fact, solving for β in Equation (2) we find that β *just is* the Bayes factor:

$$\beta = \frac{P'(E)/P'(\overline{E})}{P(E)/P(\overline{E})}.$$

So we can think of Field's Theorem 1 as a corollary of Wagner's Theorem 3.

But, crucially for us here, Wagner also shows that this kind of Bayes factor consistency is necessary for commutativity, in almost every case. Exceptions are possible, for example if E and F are the same proposition. But our regularity assumption precludes this since $E\overline{F}$ can't have positive probability if $E = F$. In fact, regularity suffices to

rule out all exceptions.[12]

Theorem 4 (Wagner). *In the schema of Figure 1, if P is regular then $P_1'' = P_2''$ implies $B_1^E = B_2^E$ and $B_1^F = B_2^F$.*

Does this theorem mean that only Field's Equation (2) can make Jeffrey conditionalization commute? No, other rules can also consistently yield the same Bayes factor for the same value of β.

One silly example is the "stubborn" rule, which just ignores β and always keeps $P'(E) = P(E)$. Substituting this rule into Step 1 of Field updating makes the Bayes factor 1 for all updates. And, trivially, updating this way is commutative: if you never change your mind, the order in which you encounter various sensory experiences won't make any difference to your final opinion.

A less trivial example—call it "upsidedownco"—replaces Field's Equation (2) with

$$P'(E) = \frac{P(E)}{P(E) + \beta \cdot P(\overline{E})}.$$

Doing a bit of algebra to isolate β, we find that this implies

$$\frac{1}{\beta} = \frac{P'(E)/P'(\overline{E})}{P(E)/P(\overline{E})}.$$

So the same value of β always results in the same Bayes factor. By Theorem 3 then, this variation on Field updating is also commutative.

However, both of these alternate rules violate the conditions we laid out in Section 2. Specifically, they violate symmetry. The stubborn rule

---

12. Wagner shows that a weaker assumption will do, but again we'll continue to assume regularity for simplicity.

is plainly not symmetric, since it privileges $P(E)$ and neglects the β proposed by experience entirely. And upsidedownco increases $P'(E)$ as $P(E)$ increases, yet *de*creases $P'(E)$ as β increases.

So Wagner's Theorem 4 is not, by itself, enough to secure Field's proposed Equation (2). Or, returning now to the social interpretation of upco and Equation (1), Wagner's result doesn't secure our Theorem 2. But with the help of further conditions like symmetry, we can rule out alternatives like the stubborn rule and upsidedownco. And this is exactly how our proof of Theorem 2 proceeds. We pick up where Wagner's result leaves off, using the four conditions of Section 2 to rule out any option but upco.

5. Conclusion

No way of combining probabilities is best for all purposes. For some purposes, there are even impossibility results showing that no pooling rule will get you everything you want.[13] But for some purposes, we can identify a single pooling rule that is the only one that will do. If your purpose is to combine your probability with an epistemic peer's and Jeffrey conditionalize on the result, and you want to be assured of commutativity, then upco is the only monotonic, continuous, uniformity preserving, and symmetric game in town.

6. Appendix: Theorems & Proofs

Here we generalize and prove Theorems 1, 2 and 4. We don't prove Theorem 3, proving Theorem 1 directly instead, for simplicity. Readers

---

13. Aczél and Wagner (1980) and McConway (1981) formulated two properties and showed that only linear pooling boasts both: Eventwise Independence says that the pool's probability for a proposition is a function only of the individuals' probabilities for that proposition, while Unanimity Preservation says that, when all the individuals assign the same probability to a proposition, the pool assigns that too. But then Laddaga (1977) and Lehrer and Wagner (1983) noted that linear pooling does not boast the property of Independence Preservation, which says that, when all the individuals take two propositions to be independent, the pool should too. Together, these results provide an impossibility theorem: no pooling rule satisfies Eventwise Independence, Unanimity Preservation, and Independence Preservation.

interested in a proof of Theorem 3 can consult Wagner (2002, Theorem 3.1).

6.1 Pooling Operators

In the main text we discussed pooling rules, which combine $P(E)$ and $Q(E)$ into a new probability for E. Since the probability of $\overline{E}$ is implied by the probability of E, these rules effectively combine probabilities over a two-cell partition, $\{E, \overline{E}\}$. For partitions with more than two elements, we need to extend this definition.

Definition 1 (Pooling operator). *A pooling operator takes a countable partition* **E** *and two probability functions P and Q defined on an agenda that includes* **E**, *and returns a partial probability function $\langle PQ \rangle_{\mathbf{E}}$ defined just on* **E**.

Upco generalizes to countable partitions in the obvious way.

Definition 2 (Upco on countable partitions). *Suppose* $\mathbf{E} = \{E_i\}$ *is a countable partition, and P and Q are probability functions defined on an agenda that includes* **E**. *Suppose further that $P(E_i), Q(E_i) > 0$ for at least one element E_i of* **E**. *Then the upco of P and Q over* **E**, *denoted $\langle PQ \rangle_{\mathbf{E}}^{\mathcal{U}}$, assigns to each E_i*

$$\langle PQ \rangle_{\mathbf{E}}^{\mathcal{U}}(E_i) = \frac{P(E_i)Q(E_i)}{\sum_j P(E_j)Q(E_j)}.$$

Notice that upco is undefined if there is no E_i such that $P(E_i), Q(E_i) > 0$. That is, upco is defined only when P and Q have *overlapping support on* **E**. The *support* of a probability function on a partition is the set of those events from that partition to which it assigns positive probability. In symbols, we write $\mathrm{supp}_{\mathbf{E}}(P) = \{E_i \in \mathbf{E} : P(E_i) > 0\}$. In this notation, $\langle PQ \rangle_{\mathbf{E}}^{\mathcal{U}}$ is defined just in case $\mathrm{supp}_{\mathbf{E}}(P) \cap \mathrm{supp}_{\mathbf{E}}(Q) \neq \emptyset$. What's more, when it is defined, the support of the upco of P and Q is

the intersection of their individual supports. More formally,

$$\text{supp}_{\mathbf{E}}(\langle PQ \rangle^{\mathcal{U}}_{\mathbf{E}}) = \text{supp}_{\mathbf{E}}(P) \cap \text{supp}_{\mathbf{E}}(Q).$$

We now extend the definition of Jeffrey pooling to countable partitions, and introduce more compact notation.

Definition 3 (Jeffrey pooling). *Let* $\mathbf{E}$ *be a countable partition, and let* P *and* Q *be probability functions such that* (i) $\langle PQ \rangle_{\mathbf{E}}$ *is defined and* (ii) $\text{supp}_{\mathbf{E}}(\langle PQ \rangle_{\mathbf{E}}) \subseteq \text{supp}_{\mathbf{E}}(P)$. *The* Jeffrey pool *of* P *and* Q *on* $\mathbf{E}$, *denoted* $\langle\!\langle PQ \rangle\!\rangle_{\mathbf{E}}$, *is the probability function defined by*

$$\langle\!\langle PQ \rangle\!\rangle_{\mathbf{E}}(-) = \sum_{E_i \in \text{supp}_{\mathbf{E}}(P)} P(- \mid E_i) \langle PQ \rangle_{\mathbf{E}}(E_i).$$

Note that the restriction $\text{supp}_{\mathbf{E}}(\langle PQ \rangle_{\mathbf{E}}) \subseteq \text{supp}_{\mathbf{E}}(P)$ is required to ensure $P(- \mid E_i)$ is defined for every E_i where $\langle PQ \rangle_{\mathbf{E}}$ is positive. This ensures that $\langle\!\langle PQ \rangle\!\rangle_{\mathbf{E}}$ is defined and a probability function.

Notice that, since the support of the upco of P and Q is the overlap of their individual supports, this condition is automatically satisfied if upco of P and Q is defined: $\text{supp}_{\mathbf{E}}(\langle PQ \rangle^{\mathcal{U}}_{\mathbf{E}}) = \text{supp}_{\mathbf{E}}(P) \cap \text{supp}_{\mathbf{E}}(Q) \subseteq \text{supp}_{\mathbf{E}}(P)$. So, if $\langle PQ \rangle_{\mathbf{E}}$ is defined, so is $\langle\!\langle PQ \rangle\!\rangle_{\mathbf{E}}$.

6.2   Field's Sufficiency Theorem

We now state and prove the general version of Theorem 1: upco ensures that Jeffrey pooling commutes, given compatible priors.

Theorem 5 (Field). *If* $\langle\!\langle \langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}} R \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}}$ *and* $\langle\!\langle \langle\!\langle PR \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}} Q \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}$ *are defined, then*

$$\langle\!\langle \langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}} R \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}} = \langle\!\langle \langle\!\langle PR \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}} Q \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}.$$

Proof. The proof generalizes the example from page 4. Intuitively, the key idea is that $\langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}$ just multiplies the outcomes within a cell E_i by

$Q(E_i)$, and renormalizes. More formally, if $A \subseteq E_i$ for some i, then

$$\langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}(A) = P(A \mid E_i) \langle PQ \rangle^{\mathcal{U}}_{\mathbf{E}}(E_i)$$

$$= \frac{P(AE_i)}{P(E_i)} \frac{P(E_i)Q(E_i)}{\sum_j P(E_j)Q(E_j)}$$

$$= cP(A)Q(E_i),$$

where c is a normalizing constant identical for all i.

So take an arbitrary proposition H, and consider for each $E_i \in \mathbf{E}$ and $F_j \in \mathbf{F}$ the proposition HE_iF_j. If one of $P(E_iF_j)$, $Q(E_i)$, or $R(F_j)$ is zero, then

$$\langle\!\langle \langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}} R \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}}(HE_iF_j) = \langle\!\langle \langle\!\langle PR \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}} Q \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}(HE_iF_j) = 0.$$

If on the other hand $P(E_iF_j), Q(E_i), R(F_j) > 0$, then

$$\langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}(HE_iF_j) = cP(HE_iF_j)Q(E_i),$$

where c is a normalizing constant independent of i and j, and thus

$$\langle\!\langle \langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}} R \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}}(HE_iF_j) = cc'P(HE_iF_j)Q(E_i)R(F_j),$$

where c' is another normalizing constant independent of i and j. Similarly,

$$\langle\!\langle \langle\!\langle PR \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}} Q \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}(HE_iF_j) = dd'P(HE_iF_j)R(F_j)Q(E_i),$$

where d and d' are again normalizing constants independent of i and j.

This shows that the probabilities $\langle\!\langle \langle\!\langle PQ \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}} R \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}}$ and $\langle\!\langle \langle\!\langle PR \rangle\!\rangle^{\mathcal{U}}_{\mathbf{F}} Q \rangle\!\rangle^{\mathcal{U}}_{\mathbf{E}}$ assign to the various HE_iF_j have the same proportions. And by a parallel argument, the same is true for the various $\overline{H}E_iF_j$. So the two

distributions have the same proportions over the partition $\{H, \overline{H}\} \times \mathbf{E} \times \mathbf{F}$, hence must be identical on this partition. Since H is a union of elements from this partition, they must assign H the same probability. But H was arbitrary. $\qquad \square$

6.3 Wagner's Necessity Theorem

Wagner identifies an almost necessary condition for Jeffrey conditionalization updates to commute. Note that here we are concerned with Jeffrey conditionalization in general: the shift from $P(E_i)$ to $P'(E_i)$ needn't be driven by a pooling rule, it could be prompted by anything. Wagner's theorem concerns any transition from P to P' that can be described in terms of Jeffrey's formula.

Definition 4 (Jeffrey conditionalization). *We say that P comes from P' by Jeffrey conditionalization on the partition $\mathbf{E}$ if $\mathrm{supp}_{\mathbf{E}}(P') \subseteq \mathrm{supp}_{\mathbf{E}}(P)$ and*

$$P'(-) = \sum_{E_i \in \mathrm{supp}_{\mathbf{E}}(P')} P(- \mid E_i) P'(E_i).$$

We will assume that P is regular on $\mathbf{EF}$; Wagner assumes something weaker, but we only need the result for regular P. Informally, the result says that, for Jeffrey updates of a regular prior to commute, the Bayes factors on each partition must match.

Theorem 6 (Wagner). *Let $\mathbf{E}$ and $\mathbf{F}$ be countable partitions such that P is regular on $\mathbf{EF}$. Let P_1' come from P by Jeffrey conditionalization on $\mathbf{E}$, P_2' from P by Jeffrey conditionalization on $\mathbf{F}$, P_1'' from P_1' by Jeffrey conditionalization*

on $\mathbf{F}$, and P_2'' from P_2' by Jeffrey conditionalization on $\mathbf{E}$. If $P_1'' = P_2''$, then

$$\frac{P_1'(E_{i_1})/P_1'(E_{i_2})}{P(E_{i_1})/P(E_{i_2})} = \frac{P_2''(E_{i_1})/P_2''(E_{i_2})}{P_2'(E_{i_1})/P_2'(E_{i_2})},$$

$$\frac{P_1''(F_{j_1})/P_1''(F_{j_2})}{P_1'(F_{j_1})/P_1'(F_{j_2})} = \frac{P_2'(F_{j_1})/P_2'(F_{j_2})}{P(F_{j_1})/P(F_{j_2})},$$

for all E_{i_1}, E_{i_2} in $\mathrm{supp}_{\mathbf{E}}(P)$, and all F_{j_1}, F_{j_2} in $\mathrm{supp}_{\mathbf{F}}(P)$.

Proof. By the rigidity of Jeffrey conditionalization, for all i, j:

$$P_1'(E_iF_j) = P_1'(E_i)P_1'(F_j \mid E_i) = P_1'(E_i)P(F_j \mid E_i),$$
$$P_1'(E_iF_j) = P_1'(F_j)P_1'(E_i \mid F_j) = P_1'(F_j)P_1''(E_i \mid F_j),$$
$$P_2'(E_iF_j) = P_2'(F_j)P_2'(E_i \mid F_j) = P_2'(F_j)P(E_i \mid F_j),$$
$$P_2'(E_iF_j) = P_2'(E_i)P_2'(F_j \mid E_i) = P_2'(E_i)P_2''(F_j \mid E_i).$$

Coupling the first two equations, and the last two, we get:

$$P_1'(E_i)P(F_j \mid E_i) = P_1'(F_j)P_1''(E_i \mid F_j), \qquad (4)$$
$$P_2'(E_i)P_2''(F_j \mid E_i) = P_2'(F_j)P(E_i \mid F_j). \qquad (5)$$

Now take any E_{i_1}, E_{i_2} in $\mathbf{E}$ and F_j in $\mathbf{F}$. Using Equation (4), we can

analyze our first Bayes factor as follows:

$$
\frac{P_1'(E_{i_1})/P_1'(E_{i_2})}{P(E_{i_1})/P(E_{i_2})} = \frac{P_1'(F_j)P_1''(E_{i_1}\mid F_j)/P(F_j\mid E_{i_1})}{P_1'(F_j)P_1''(E_{i_2}\mid F_j)/P(F_j\mid E_{i_2})} \bigg/ \frac{P(E_{i_1})}{P(E_{i_2})}
$$

$$
= \frac{P_1''(E_{i_1}F_j)}{P_1''(E_{i_2}F_j)}\frac{P(F_j\mid E_{i_2})}{P(F_j\mid E_{i_1})}\frac{P(E_{i_2})}{P(E_{i_2})}
$$

$$
= \frac{P_1''(E_{i_1}F_j)}{P_1''(E_{i_2}F_j)}\frac{P(E_{i_2}F_j)}{P(E_{i_1}F_j)}.
$$

Parallel reasoning with Equation (5) gives:

$$
\frac{P_2''(E_{i_1})/P_2''(E_{i_2})}{P_2'(E_{i_1})/P_2'(E_{i_2})} = \frac{P_2''(E_{i_1}F_j)}{P_2''(E_{i_2}F_j)}\frac{P(E_{i_2}F_j)}{P(E_{i_1}F_j)}.
$$

So the Bayes factors over **E** are identical. The identity of the Bayes factors over **F** follows similarly. □

6.4 Our Theorem

Here we use Wagner's theorem to show the general form of Theorem 2: upco is the only monotonic, uniformity preserving, continuous, symmetric, and extensional pooling operator capable of ensuring that Jeffrey pooling commutes.

Our strategy: first prove that any pooling operator with these features, and which ensures Jeffrey pooling commutes for regular probability functions, must agree with upco when the pooled functions are regular. Then we'll appeal to continuity to show that any pooling operator that agrees with upco on the regular functions agrees with upco everywhere it's defined.

We begin by defining terms:

Definition 5 (Uniform). *A distribution P is* uniform *over* **E** *if* $P(E_{i_1}) = P(E_{i_2})$ *for all* E_{i_1}, E_{i_2} *in* **E**.

Definition 6 (Uniformity preservation). *A pooling operator is* uniformity preserving *if* $\langle PQ\rangle_{\mathbf{E}}$ *is uniform over* **E** *whenever P and Q are uniform over* **E**.

Notice that we must set the infinite case aside now, because uniform distributions don't exist over countably infinite partitions.

Definition 7 (Monotonicity). *A pooling operator is* monotone *if, when P is uniform over* **E**, $Q(E_i) < R(E_i)$ *implies* $\langle PQ\rangle_{\mathbf{E}}(E_i) < \langle PR\rangle_{\mathbf{E}}(E_i)$.

Note that this is a very restricted form of monotonicity, since it only concerns the case where one argument is uniform.

Definition 8 (Symmetry). *A pooling operator is* symmetric *if* $\langle PQ\rangle_{\mathbf{E}} = \langle QP\rangle_{\mathbf{E}}$ *for all P, Q, and* **E**.

Definition 9 (Continuity). *A pooling operator is* continuous *if*

$$
\lim_{n\to\infty}\langle P_nQ\rangle_{\mathbf{E}} = \langle(\lim_{n\to\infty}P_n)Q\rangle_{\mathbf{E}},
$$

whenever $\langle P_nQ\rangle_{\mathbf{E}}$ *is defined for each n and* $\langle(\lim_{n\to\infty}P_n)Q\rangle_{\mathbf{E}}$ *is defined.*

The restriction avoids ruling out operators like geometric pooling and upco from the get go, since there are sequences $P_1, P_2, \ldots$ such that $\langle P_i, Q\rangle_{\mathbf{E}}^{\mathcal{U}}$ is defined for each i, but $\langle(\lim_{n\to\infty}P_n)Q\rangle_{\mathbf{E}}^{\mathcal{U}}$ is not defined.

Definition 10 (Extensionality). *A pooling operator is* extensional *if, given partitions* **E** *and* **F** *of equal size,* $P(E_i) = R(F_i)$ *and* $Q(E_i) = S(F_i)$ *for all i imply* $\langle PQ\rangle_{\mathbf{E}}(E_i) = \langle RS\rangle_{\mathbf{F}}(F_i)$ *for all i.*

The main work in establishing the theorem of this section is showing that the pooling operator must treat uniform distributions as "neutral." That is, pooling any distribution with a uniform distribution just returns the original distribution. We now use the conditions just defined,

together with commutativity for Jeffrey pooling, to derive this feature in the case in which the function pooled with the uniform one is regular.

Lemma 7. *Suppose that $\langle\!\langle\langle\!\langle PQ\rangle\!\rangle_{\mathbf{E}}R\rangle\!\rangle_{\mathbf{F}} = \langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}$ for any finite partitions $\mathbf{E}$ and $\mathbf{F}$ such that P, Q, and R are regular. Then, if the pooling operator is uniformity preserving, monotonic, symmetric, continuous, and extensional, it must treat uniform distributions as neutral. That is, for P uniform over $\mathbf{E}$ and Q regular on $\mathbf{E}$, $\langle PQ\rangle_{\mathbf{E}}(E_i) = Q(E_i)$.*

Proof. Let $\mathbf{E} = \{E_i\}$ and $\mathbf{F} = \{F_j\}$ be finite partitions of size n, let Q be uniform over $\mathbf{E}$, and let R be positive for every element of $\mathbf{F}$. Define P as follows, where $0 < \epsilon < 1/(n-1)$:

$$P(E_i F_j) = \begin{cases} \frac{1}{n} - \frac{n-1}{n}\epsilon & \text{if } i = j, \\ \frac{1}{n}\epsilon & \text{if } i \neq j. \end{cases}$$

Observe that $P(E_i) = 1/n = P(F_j)$, so P is uniform over $\mathbf{E}$ and over $\mathbf{F}$. Note for later that

$$P(E_i \mid F_j) = \begin{cases} 1 - (n-1)\epsilon & \text{if } i = j, \\ \epsilon & \text{if } i \neq j. \end{cases}$$

P, Q, and R are regular, so Theorem 6 gives the following Bayes factor identity for all E_{i_1}, E_{i_2}:

$$\frac{\langle\!\langle PQ\rangle\!\rangle_{\mathbf{E}}(E_{i_1})}{\langle\!\langle PQ\rangle\!\rangle_{\mathbf{E}}(E_{i_2})} \Big/ \frac{P(E_{i_1})}{P(E_{i_2})} = \frac{\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_{i_1})}{\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_{i_2})} \Big/ \frac{\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_{i_1})}{\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_{i_2})} . \tag{6}$$

Since P is uniform over $\mathbf{E}$, the denominator on the left is 1. And since

Q is also uniform over $\mathbf{E}$, uniformity preservation implies that the numerator is also 1. Also, $\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_i) = \langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_i)$ for all i by the definition of Jeffrey pooling. So Equation (6) reduces to

$$\frac{\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_{i_1})}{\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_{i_2})} = \frac{\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_{i_1})}{\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_{i_2})}.$$

Since this holds for all E_{i_1}, E_{i_2}, the distributions $\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}$ and $\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}$ have the same relative proportions over $\mathbf{E}$, hence must actually be the same distribution. That is, for all i:

$$\langle\!\langle\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}Q\rangle\!\rangle_{\mathbf{E}}(E_i) = \langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_i).$$

Using symmetry to move Q to the left, and then substituting P for Q on grounds of extensionality, this becomes:

$$\langle P\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}\rangle_{\mathbf{E}}(E_i) = \langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_i). \tag{7}$$

Now, by definition the right hand side, is:

$$\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}(E_i) = \sum_j P(E_i \mid F_j)\langle PR\rangle_{\mathbf{F}}(F_j)$$

$$= (1 - (n-1)\epsilon)\langle PR\rangle_{\mathbf{F}}(F_i) + \epsilon \sum_{j \neq i}\langle PR\rangle_{\mathbf{F}}(F_j).$$

So in the limit as ϵ goes to 0, $\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}$ assigns over $\mathbf{E}$ the same values $\langle PR\rangle_{\mathbf{F}}$ assigns over $\mathbf{F}$. Let S be this distribution that $\langle\!\langle PR\rangle\!\rangle_{\mathbf{F}}$ approaches, *i.e.* S is a copy over $\mathbf{E}$ of the assignments $\langle PR\rangle_{\mathbf{F}}$ makes over $\mathbf{F}$:

$$S(E_i) = \langle PR\rangle_{\mathbf{F}}(F_i),$$

for all i. By continuity we have for all i:

$$\lim_{\epsilon \to 0} \langle P \langle\langle PR \rangle\rangle_{\mathbf{F}} \rangle_{\mathbf{E}}(E_i) = \langle P \lim_{\epsilon \to 0} \langle\langle PR \rangle\rangle_{\mathbf{F}} \rangle_{\mathbf{E}}(E_i) = \langle PS \rangle_{\mathbf{E}}(E_i) = \langle PR \rangle_{\mathbf{F}}(F_i).$$

The last identity here is the one we need.

Now suppose for a contradiction that $\langle PR \rangle_{\mathbf{F}}(F_j) \neq R(F_j)$ for some F_j. Then there must be an F_k for which $\langle PR \rangle_{\mathbf{F}}(F_k) < R(F_k)$. Since S copies $\langle PR \rangle_{\mathbf{F}}$, this implies $S(E_k) < R(F_k)$. Thus we have:

$$S(E_k) < R(F_k),$$
$$\langle PS \rangle_{\mathbf{E}}(E_k) = \langle PR \rangle_{\mathbf{F}}(F_k).$$

And this contradicts monotonicity. By extensionality, the partition doesn't matter, since P is uniform over both $\mathbf{E}$ and $\mathbf{F}$. So increasing the kth value of the non-uniform input should increase the corresponding output.

This shows that $\langle PR \rangle_{\mathbf{F}}(F_j) = R(F_j)$ for all F_j. Since P was uniform over $\mathbf{F}$ and R regular, extensionality then implies that for any P uniform over $\mathbf{E}$ and Q regular on $\mathbf{E}$, $\langle PQ \rangle_{\mathbf{E}}(E_i) = Q(E_i)$ for all E_i, as desired. $\square$

We now show that only upco has the five features defined above, and makes Jeffrey pooling commutative.

Theorem 8. *Suppose that $\langle\langle \langle\langle PQ \rangle\rangle_{\mathbf{E}} R \rangle\rangle_{\mathbf{F}} = \langle\langle \langle\langle PR \rangle\rangle_{\mathbf{F}} Q \rangle\rangle_{\mathbf{E}}$ for any finite partitions $\mathbf{E}$ and $\mathbf{F}$ and any compatible P, Q, and R. Then, if the pooling operator is uniformity preserving, monotonic, symmetric, continuous, and extensional, it must be upco.*

Proof. We begin by proving that, if $\langle\langle \langle\langle PQ \rangle\rangle_{\mathbf{E}} R \rangle\rangle_{\mathbf{F}} = \langle\langle \langle\langle PR \rangle\rangle_{\mathbf{F}} Q \rangle\rangle_{\mathbf{E}}$ for all regular P, Q, and R, then the pooling operator must agree with upco on regular functions. Then we show that any continuous operator that agrees with upco on the regular functions must be upco.

Let $\mathbf{E}$ and $\mathbf{F}$ be finite partitions of size n, and define P as in the proof of Lemma 7. Let Q and R be positive everywhere on $\mathbf{E}$, and let R' mimic on $\mathbf{F}$ the distribution of R on $\mathbf{E}$, i.e. $R'(F_i) = R(E_i)$ for all i.

P, Q, and R' are regular, so by Theorem 6 Equation (6) holds, with R' in place of R. By Lemma 7, $\langle PQ \rangle_{\mathbf{E}}(E_i) = Q(E_i)$ for all i, so in this case Equation (6) reduces to

$$\frac{\langle Q \langle\langle PR' \rangle\rangle_{\mathbf{F}} \rangle_{\mathbf{E}}(E_{i_1})}{\langle Q \langle\langle PR' \rangle\rangle_{\mathbf{F}} \rangle_{\mathbf{E}}(E_{i_2})} = \frac{Q(E_{i_1})}{Q(E_{i_2})} \frac{\langle\langle PR' \rangle\rangle_{\mathbf{F}}(E_{i_1})}{\langle\langle PR' \rangle\rangle_{\mathbf{F}}(E_{i_2})}. \tag{8}$$

But

$$\langle\langle PR' \rangle\rangle_{\mathbf{F}}(E_i) = \sum_j P(E_i \mid F_j) \langle PR' \rangle_{\mathbf{F}}(F_j)$$

$$= (1 - (n-1)\epsilon) \langle PR' \rangle_{\mathbf{F}}(F_i) + \epsilon \sum_{j \neq i} \langle PR' \rangle_{\mathbf{F}}(F_j).$$

So

$$\lim_{\epsilon \to 0} \langle\langle PR' \rangle\rangle_{\mathbf{F}}(E_i) = \langle PR' \rangle_{\mathbf{F}}(F_i) = R'(F_i) = R(E_i).$$

Thus, by continuity and Equation (8):

$$\frac{\langle QR \rangle_{\mathbf{E}}(E_{i_1})}{\langle QR \rangle_{\mathbf{E}}(E_{i_2})} = \frac{Q(E_{i_1})}{Q(E_{i_2})} \frac{R(E_{i_1})}{R(E_{i_2})}.$$

Now observe that this is the same ratio delivered by upco:

$$\frac{\langle QR \rangle_{\mathbf{E}}^{\mathcal{U}}(E_{i_1})}{\langle QR \rangle_{\mathbf{E}}^{\mathcal{U}}(E_{i_2})} = \frac{Q(E_{i_1})R(E_{i_1}) / \sum_k Q(E_k)R(E_k)}{Q(E_{i_2})R(E_{i_2}) / \sum_k Q(E_k)R(E_k)} = \frac{Q(E_{i_1})R(E_{i_1})}{Q(E_{i_2})R(E_{i_2})}.$$

So $\langle QR \rangle_{\mathbf{E}}$ and $\langle QR \rangle_{\mathbf{E}}^{\mathcal{U}}$ have the same relative proportions, hence must

be the same distribution.

This shows $\langle PQ \rangle_{\mathbf{E}} = \langle PQ \rangle_{\mathbf{E}}^{\mathcal{U}}$ if P and Q are both regular on $\mathbf{E}$.

Finally, suppose one or the other or both of P and Q is not regular on $\mathbf{E}$, but $\langle PQ \rangle_{\mathbf{E}}^{\mathcal{U}}$ is defined. Then there are sequences $P_1, P_2, \ldots$ and $Q_1, Q_2, \ldots$ of regular probability functions such that $\lim_{n \to \infty} P_n = P$ and $\lim_{n \to \infty} Q_n = Q$. And so, by continuity,

$$\langle PQ \rangle_{\mathbf{E}} = \langle \lim_{n \to \infty} P_n, \lim_{n \to \infty} Q_n \rangle_{\mathbf{E}} = \lim_{n \to \infty} \langle P_n Q_n \rangle_{\mathbf{E}}$$
$$= \lim_{n \to \infty} \langle P_n Q_n \rangle_{\mathbf{E}}^{\mathcal{U}} = \langle \lim_{n \to \infty} P_n, \lim_{n \to \infty} Q_n \rangle_{\mathbf{E}}^{\mathcal{U}} = \langle PQ \rangle_{\mathbf{E}}^{\mathcal{U}}.$$

This completes the proof. ☐

References

Aczél, J. and Carl G. Wagner. 1980. "A Characterization of Weighted Arithmetic Means." *SIAM Journal on Algebraic Discrete Methods* 1(3):259–260.

Christensen, David. 2007. "Epistemology of Disagreement: The Good News." *The Philosophical Review* 116(2):187–217.

Christensen, David. 2009. "Disagreement as Evidence: The Epistemology of Controversy." *Philosophy Compass* 4(5):756–67.

Dietrich, Franz and Christian List. 2016. Probabilistic Opinion Pooling. In *Oxford Handbook of Philosophy and Probability*, ed. Alan Hàjek and Christopher Hitchcock. Oxford University Press pp. 519–42.

Easwaran, Kenny, Luke Fenton-Glynn, Christopher Hitchcock and Joel D. Velasco. 2016. "Updating on the Credences of Others: Disagreement, Agreement, and Synergy." *Philosophers' Imprint* 6(11):1–39.

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

Field, Hartry. 1978. "A Note on Jeffrey Conditionalization." *Philosophy of Science* 45(3):361–7.

Fitelson, Branden and David Jehle. 2009. "What is the 'Equal Weight View'?" *Episteme* 6(3):280–293.

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

Jeffrey, Richard C. 1965. *The Logic of Decision*. New York: University of Chicago Press.

Jeffrey, Richard C. 1983. Bayesianism With A Human Face. In *Testing Scientific Theories*, ed. John Earman. University of Minnesota Press pp. 133–156.

Kelly, Thomas. 2010. Peer Disagreement and Higher Order Evidence. In *Social Epistemology: Essential Readings*, ed. Alvin I. Goldman and Dennis Whitcomb. Oxford University Press. pp. 183–217.

Laddaga, Robert. 1977. "Lehrer and the Consensus Proposal." *Synthese* 36(4):473–77.

Lehrer, Keith and Carl Wagner. 1983. "Probability Amalgamation and the Independence Issue: A Reply to Laddaga." *Synthese* 55(3):339–346.

McConway, K. J. 1981. "Marginalization and Linear Opinion Pools." *Journal of the American Statistical Association* 76(374):410–414.

Roussos, Joe. 2021. "Expert Deference as a Belief Revision Schema." *Synthese* 199(1–2):3457–84.

Wagner, Carl. 2002. "Probability Kinematics and Commutativity." *Philosophy of Science* 69(2):266–78.

Wagner, Carl. 2011. "Peer Disagreement and Independence Preservation." *Erkenntnis* 74(2):277–88.