WHAT THE "EQUAL WEIGHT VIEW" IS

Abstract. We examine arguments by Dawid, DeGroot and Mortera, Nissan-Rozen and Spectre and finally Fitelson and Jehle against linear opinion pooling implementations of the "Equal Weight View" response to same-evidence cases of peer disagreement. We then give arguments in favor of geometric opinion pooling, touching on a neglected result of Abbas that equal weight geometric pooling minimizes the sum of the information distances (Kullbeck-Leibler divergences) from the common posterior to the agents' respective priors.

1. A triviality result for split-the-difference

In the epistemology of disagreement literature[1] one encounters an *Equal Weight View*, according to which"When you count an advisor as an epistemic peer, you should give her conclusions the same weight as your own" (Elga 2007). Many have taken "splitting the difference" (i.e., adoption of the arithmetic mean) between competing peer credences to be constitutive of this view. Kelly (2010), e.g., writes:

> ...if the agnostic gives credence .5 to the proposition that God exists while the atheist gives credence .1 to the same proposition, the import of The Equal Weight View is clear: upon learning of the other's opinion, each should give credence .3 to the proposition that God exists.

The popularity of Equal Weight difference splitting persists, despite the fact that it was shown, a quarter century ago in Dawid, DeGroot and Mortera (1995), to entail probabilistic incoherence. Indeed, it has been disputed that this coherence result applies at all to the situation that interests us, namely that of two peers having identical evidence but different priors. Bradley (2018), for example, rehearses a version of the result, but goes on to write:

[1]We advocate in these pages for a certain solution to the problem of peer disagreement. (And critique a popular alternative.) There is a common sentiment that there can't be a unique solution to this problem; that ideal amalgamation of quantified belief by peers is a scenario-dependent "art". Scenario dependence implies that there could be two situations in which pairs of agents have equivalent credence functions over equal-cardinality finite partitions, yet properly choose to amalgamate their quantified beliefs differently owing to differences in their respective beliefs not reflected in their raw credences over this partition. Here we will assume that the agents don't have any relevant evidence beyond their own credences and those of their peer, and that, accordingly, the peer update problem well could have a unique solution.

> ...this study leaves open the question of whether linear averaging is the appropriate response to situations in which you find yourself in disagreement with peers who hold the same information as you and are as good at judging its significance. In the philosophical literature, the view that one should respond to such disagreements by taking an equal-weighted average of your opinions has been hotly debated. But nothing presented here militates either for or against this view.

In this section, we make the case that considerations very much along the lines of Dawid et. al. (1995) do "militate" (decisively) against split-the-difference in cases where two peers[2] have the same external evidence but do not have almost surely identical prior credences. We begin by supposing that i and j are agents and P is a proposition-valued random variable, so that i's initial credence x in P and j's initial credence y in P are also random variables. We take P's distribution to be supported on propositions P_0 for which i and j are peers.

Because P is a random variable, there is some ambiguity concerning "credence in P" we need to address. Suppose that a card, c, is drawn from a standard deck and P is either *Card c is a face card* or *Card c is an Ace* (with equal probabilities). In this case, P is in fact *Card c is an Ace*, but i doesn't know this. In such a case there are two readings of "i's credence in P". On the first reading, it refers to i's credence in *Card c is an Ace*, namely $\frac{1}{13}$. On the second reading it refers to i's credence in the proposition *either P is "Card c is a face card" and Card c is a face card or P is "Card c is an Ace" and Card c is an Ace*, namely $\frac{2}{13}$. We'll write $Cr_i(P)$ when we intend the first reading and $Cr_i(T(P))$ when we intend the second. (It may help to think of $T(P)$ as "P is true", read *de dicto*.) Notice that $Cr_i(P) = x$ regardless of whether i knows the value of P, whereas it will typically be the case that $Cr_i(T(P)) = x$ only after i learns P or at least $Cr_i(P)$.

When we say that i is "diachronically coherent almost surely" or "Reflection[3] obeying almost surely", we mean in particular that if i's credence in P_0 is x_0 then

$$E(u(x, y)|P = P_0) = x_0, \tag{1}$$

[2]Splitting the difference with a non-peer can, by contrast, be coherent; if my credence in P is one-half and I believe that your credence in whichever of P, $\neg P$ is true is 1 with probability .75 and zero with probability .25 then I surely don't consider you a peer (it seems that I think you are more sensitive to which of P, $\neg P$ is true, yet wildly overconfident), but should intend to split the difference with you when you tell me your credence.

[3]"Reflection" was coined by van Fraassen (1984). Roughly, an agent satisfies it when her current credence in a proposition P is equal to the expectation of her credence in P at a future time t, where t is typically an almost-surely future, possibly random time satisfying certain technical criteria (a so-called "stopping time"–see Schervish et. al. 2004). In the current application t is the time immediately after j's credence in P is revealed.

where $u(x,y)$ denotes the posterior credence in P adopted by i upon learning the values of P, x and y. (Or in $T(P)$ upon learning just x and y.)[4] In particular, for almost every x_0 in the essential range of x one has

$$E(u(x,y)|x = x_0) = x_0.$$

That is, if she were to learn her own initial credence $x = x_0$ (without learning P) then she would both come to have credence x_0 in $T(P)$ and expected posterior (posterior to learning y, that is) credence x_0 in $T(P)$.

Here is a natural necessary condition on peerhood.

> **PN:** If i is diachronically coherent almost surely and regards j as a peer then for almost every y_0 in the essential range of y one has $E(u(x,y)|y = y_0) = y_0$.

PN is nothing more or less than a mathematical translation of the truism that if i regards j as a peer then she has the same confidence in j's initial credences as she has in her own. (There is therefore no reasonable cause for objecting to the principle, though we will entertain one such objection in the final section anyway.) As observed above, if i were to learn that, and only that, $x = x_0$, then she would come to have credence x_0 in $T(P)$. So if she regards j as a peer and were to learn that, and only that, $y = y_0$, then she ought, similarly, to come to have credence y_0 in $T(P)$.[5] But if she learns the value of x, say $x = x_0$, after learning that $y = y_0$, her posterior credence in $T(P)$ will be $u(x_0, y_0)$. By Reflection, then, her current (i.e. after learning $y = y_0$ but before learning $x = x_0$) credence, namely y_0, should be the expectation of this posterior, that is $E_x(u(x,y)|y = y_0)$.

We can now establish incoherence (on pain of triviality) of linear opinion pools (including "split-the-difference") for Equal Weighters.

Theorem 1. If i is diachronically coherent almost surely, regards j as a peer and updates by linear pooling (i.e. $u(x,y) = wx + (1-w)y$, where $0 < w < 1$) then $Prob(x \neq y) = 0$.

---

[4]If there are only countably many propositions in the algebra from which P then (1) requires no elaboration, but this assumption is not necessary. In the continuous case, we can interpret $E(u(x,y)|x = x_0)$ as, say, $\lim_n E(u(x,y)|x \in i_n(x_0))$, where $i_n(x_0)$ denotes the unique interval $[\frac{k}{2^n}, \frac{k+1}{2^n})$ containing x_0. (And similarly for other expressions; here k is an integer.) Then all claims are to be interpreted as holding for almost every x in the essential range of x, etc. (Such issues are familiar and not difficult to navigate, so we'll not dwell on them.)

[5]We see here one reason for bothering about $T(P)$. If in the first instance we were to learn the identity of the proposition P (rather than just our credence in P), we would have to imagine learning the identity of the proposition P (rather than just our peer's credence in P) in the second instance as well. But that would automatically inform us of our our own credence in P. So in order for the cases to be analogous, we can't learn the identity of P in either scenario.

Proof. Since j is diachronically coherent almost surely, for almost every x_0 in the essential range of x one has $E((u(x,y)|x=x_0) = E(wx+(1-w)y|x=x_0) = x_0$. But obviously $E(x|x=x_0)=x_0$, so in fact

$$E(y|x=x_0)=x_0 \tag{2}$$

a.e. Meanwhile since i regards j as a peer, for almost every y_0 in the essential range of y one has $E((u(x,y)|y=y_0) = E(wx+(1-w)y|y=y_0) = y_0$ (by **PN**). But obviously $E(y|y=y_0)=y_0$, so in fact

$$E(x|y=y_0)=y_0 \tag{3}$$

a.e. If we now multiply (2) by x_0 and then integrate both sides with respect to x_0 we get $E(xy)=E(x^2)$. Similarly, if we multiply (3) by y_0 and integrate both sides with respect to y_0, we get $E(xy)=E(y^2)$. Therefore

$$E((x-y)^2) = E(x^2) - 2E(xy) + E(y^2) = 0,$$

completing the proof. qed

These considerations look to kill the difference splitting implementation of the Equal Weight View. On the other hand, coherent peer update schemes may *approximate* difference splitting. There is, moreover, a simple way of constructing such schemes. Namely, by considering the parallel case in which the agents had the same original priors, but have since acquired different evidence. Though this isn't the case we are interested in, the existence of these scenarios limits the methods by which one can hope to argue. (In particular, if one hopes to argue against difference splitting in the same evidence case on coherence grounds, one had better be certain that one's argument doesn't overgeneralize.)

We sketch such a scenario. Suppose a point x is chosen uniformly at random on the unit interval. A standard Brownian motion Z is initiated at x and evolves until it exits the interval. P is the event that it exits to the right, i.e. at 1. Neither i nor j know the value x. Suppose next that two independent standard Brownian motions, Z_i and Z_j, are initiated at x and stopped at time $t=10^{-24}$. i is told the value $x_i = Z_i(t)$ and j is told the value $x_j = Z_j(t)$. Since the standard deviation, 10^{-12}, of x_i is so small, and since the expectation of x_i is x, the expectation x_i' of x conditional on x_i (= i's probability for P conditional on x_i) will, with high probability, be extremely close to (distance much less than 10^{-12}) x_i. Similarly, j's credence in P will be, with high probability, extremely close to x_j.

On the other hand, the expected value x' of x conditional on x_i *and* x_j will with high probability be extremely close to (distance much less than $|x_i - x_j|$) the midpoint of x_i and x_j; so when i and j share their credences, they will, with high probability, adopt posterior credence $x' = u(x'_i, x'_j)$ in P extremely close to (relative to $|x'_i - x'_j|$) the midpoint of their shared credences x'_i and x'_j.

To reiterate, this is not a same evidence scenario. However, the update rule $u(\cdot, \cdot)$ that falls out of it will be coherent in any same-evidence scenario in which the joint distribution of the peers agents' priors x'_i and x'_j is the same. So to argue against peer update schemes that approximate difference splitting, one would have either to propose norms directly constraining such joint distributions, or propose indirect constraints. We examine two approaches of the latter sort presently.

2. On a putative peerhood constraint of Nissan-Rozen and Spectre

Ittay Nissan-Rozen and Levi Spectre (2017) present an original argument against difference splitting as an implementation of the Equal Weight View. It fails, as we shall demonstrate. It begins with a novel proposed constraint on peerhood:

> Our main contribution takes the form of a pragmatic constraint on the notion of peerhood: if an agent, j, is your peer, then assuming that j is sympathetic–she wants you to gain as much as possible–you should be willing–in exchange for a certain payoff–to let her decide for you whether to accept a bet with positive expected utility. If you are not willing to accept this exchange even for a sure payoff, you do not seriously regard j as your peer.

Nissan-Rozen and Spectre go on to prove the following theorem, in which P is a proposition for which i has an initial credence, and i is committed to updating by linear pooling upon learning j's initial credence.

Theorem 2. (Nissan-Rozen and Spectre 2017) Let i be an agent for whom j is a fully rational and sympathetic peer. For any credence function of i that assigns a non-trivial probability value to the possibility that j's degree of belief in P is different from i's degree of belief in P, and for any $0 < w < 1$, there always exists a bet with positive expected utility such that i (if she updates by linear pooling with weights w, $1 - w$) will be willing to pay a positive amount of utility in order to avoid passing the choice of whether to accept the bet (on i's behalf) to j.

The bet guaranteed by Theorem 2 violates Nissan-Rozen and Spectre's pragmatic constraint. If the constraint is viable, then, difference splitting Equal Weighters do not regard their fellow agents as peers, which implies in particular that split-the-difference cannot be a workable implementation of the Equal Weight View.[6]

[6] Nissan-Rozen and Spectre also prove that there will be a bet with positive expected utility such that a difference splitting i will be inclined to pay a positive amount of utility in order to pass the bet to j. This violates an apparently endorsed (if only implicitly) variant of their constraint whereby you should be willing–in exchange for a certain payoff–to decide for yourself

Nissan-Rozen and Spectre claim that their constraint (in conjunction with Theorem 2) “makes room for the development of a new Conciliatory view that calls for varying weights” (Nissan-Rozen and Spectre 2017). We take a “variable weight” rule to be one for which $u(x, y)$ lies strictly between x and y, if $x \neq y$, and $u(x, x) = x$.[7] But on this understanding, Theorem 2 *doesn't* make room for such views; if in fact the argument attaches a deficient notion of peerhood to split-the-difference, variable weight Conciliatory views are collateral damage.

To see why, let $u(y)$ be i's posterior credence under such a scheme when i learns $y = Cr_j(P)$. Since we are assuming a “variable weight” rule the function $u(y)$ satisfies $u(y) = x$ for $x = y$, with $u(y)$ strictly between x and y otherwise. Since i regards j as a peer, meanwhile, we can assume that $u(y)$ is strictly increasing. Finally, since i is coherent, she should obey Reflection; in particular her initial credence $x = Cr_i(P)$ is the expectation of her posterior, i.e. $x = E_y(u(y))$.

Under these assumptions, one can always find a bet that i will pay a positive amount to avoid passing to j whenever y isn't, by i's lights, equal to x almost surely. For in such a case i must, by non-triviality and Reflection, assign positive probability to the event $y < u(y) < x$. Let y_0 be the essential infimum of y.[8] Choose k with $y_0 < k < u(y_0) < x$. Bet 1 pays 1 if P is true and pays 0 if P is false, if accepted; one receives a sure k if Bet 1 is rejected. Since i's posterior (after learning y, that is) credence in P is almost surely greater than k, acceptance of Bet 1 has positive expected utility for i. She will be willing, moreover, to pay any amount less than $(u(y_0) - k)Prob(y < k) > 0$ to avoid having this bet passed to j, since j would reject it whenever $y < k$.

We think what Nissan-Rozen and Spectre had in mind was that one should deem a peer as being no worse (in expectation) than oneself when it comes to accepting or rejecting a bet of the form given (1 if the proposition is true and 0 if it is false,

---

whether to accept a bet with positive expected utility in a case where you are otherwise obliged to pass it to j. The details aren't precisely the same, but this variant overgeneralizes as well, and so cannot be used to resuscitate the Nissan-Rozen/Spectre argument. In any event we set this aside, as they don't formally invoke (or even formulate) the variant in question.

[7]Some authors (e.g. Easwaran et. al. 2016) advocate for *synergy*, which implies that in case $y = x \neq \frac{1}{2}$, i's posterior distribution should be more extreme than the common initial credence. Synergy is appropriate to the more common case where disagreeing peers have different evidence and the same priors. Suppose for example that i and j have common prior distribution that is uniform on $[0, 1]$ for the bias of a coin and are each allowed to toss the coin once, privately. If they reconvene and simultaneously announce credence of $\frac{1}{3}$ in the next toss of the coin landing *heads*, they will update not to $\frac{1}{3}$ but to $\frac{1}{4}$ (Laplace rule of succession). That is because their disclosures effectively allow for a pooling of evidence. Something like this is going on, for example, when so-called *meta-analyses* obtain “statistically significant” results (i.e. sufficiently extreme p values) by pooling studies that individually were unable to derive such results. In the same-evidence case we are interested in, however, the practice is plainly unjustified, indeed a bit like concluding that, because a certain balloon looks orange to everyone in the room, it must therefore be red.

[8]That is, $Prob(y < y_0) = 0$, but $Prob(y > y_0 + \epsilon) > 0$ for every $\epsilon > 0$.

if accepted; a certain amount c if rejected), *prior to learning one's own initial credence x in the proposition in question.* Once one learns the value of x, that might change. If x is very close to c, the agent will recognize that the expected relative utility of her choice is small (non-existent, when $x = c$), and she may want to pass the bet to j. In at least some other cases (cases in which x and c are not close, typically), she will be inclined to want to field the bet herself.

The proposed constraint is therefore implausible–it would, if valid, rule out too much. That is to say, it can't be a requirement of peerhood that for every such offer one should think that one's peer has expected return not less than one's own. We conclude that Nissan-Rozen and Spectre's argument fails.

3. On a would-be desideratum of Fitelson and Jehle

Fitelson and Jehle (2009) attempt to discredit difference splitting *simpliciter* (i.e. their argument does not invoke peerhood) on the grounds that it fails to commute with conditionalization. Such an argument, it's probably worth mentioning, cannot counsel against difference splitting for two cell partitions, for the simple reason that if one conditionalizes on a non-trivial event from a two cell partition, the resulting space is trivial and there is only one candidate credence function over it. So Theorem 1 is more general, even if this alternative argument has merit.

The argument in its current form is seriously gapped, however, owing to the fact that Fitelson and Jehle believed the matter to be much simpler than it is. Indeed, they regarded it as transparent enough to relegate to a footnote:

> Some Bayesian defenders of EWV require that (ideally) the result of an EWV update should be equivalent to a (classical) conditionalization, which conditionalizes "on whatever you...have learned about the circumstances of the disagreement" (Elga 2007, 490). If that's right, then [commutativity] will follow from the definition of (classical) Bayesian conditionalization, since pairs of (classical) conditionalizations must commute. (Fitelson and Jehle 2009, footnote 12.)

That isn't quite right. What's true is that if i conditions on A and then conditions on B, she should arrive at the same posterior as if she were to have first conditioned on B and then on A. But that's not what's going on here.

Suppose for example that i's original prior on (A, B, C) is $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ and j's is $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. It is certainly true that i should arrive at the same posterior if she first learned j's credence function then learned $\neg C$ as she should if she first learned $\neg C$ and then learned j's original prior. And these are just the propositions i will learn in a case where the agents first perform a peer update and then condition on $\neg C$. It is not, however, what i will learn if the agents first condition on $\neg C$ and then perform a peer update. In the latter case, i will learn only that j's original prior was of the form $(x, 2x, 1-3x)$ for some $0 < x \leq \frac{1}{3}$. There is no (obvious) reason to think, then, that her posterior here must be the same.

In fact, one can easily construct coherent peer update rules that fail to commute with conditionalization: rules that approximate split-the-difference, for example, will violate this commutativity. The two-cell scheme presented at the end of Section 1 can be adapted to three cells to this end, as we now show.

Consider an equilateral triangle and a point generated uniformly at random in its interior, having barycentric coordinates (x, y, z) (x, y and z denote the distances from the point to the sides of the triangle; we assume $x + y + z = 1$). A two dimensional Brownian motion will be initiated at this point. When it hits a side (i.e. when one of the coordinates becomes zero), the Brownian motion will become 1-dimensional on that side until it terminates at a vertex. Let P_X be the event that the motion terminates at the vertex X having barycentric coordinates $(1, 0, 0)$; P_Y and P_Z are similarly defined.

Neither i nor j knows the initial point (x, y, z). However, they each learn the identity of a nearby point–points (x_i, y_i, z_i) and (x_j, y_j, z_j) respectively–chosen from independent bivariate normal distributions having mean at the point with barycentric coordinates (x, y, z) and common, extremely small known variance. Assuming the agents to be rational, their resulting credences in (P_X, P_Y, P_Z) will be $(x'_i, y'_i, z'_i) \approx (x_i, y_i, z_i)$ and $(x'_j, y'_j, z'_j) \approx (x_j, y_j, z_j)$. Upon sharing these credences, they will each come to have posterior credence $(x', y', z') \approx \frac{1}{2}((x'_i, y'_i, z'_i) + (x'_j, y'_j, z'_j))$ in (P_X, P_Y, P_Z). The error in these approximations will be small compared to the distance between (x'_i, y'_i, z'_i) and (x'_j, y'_j, z'_j) with very high probability.

In particular, the error will always be small in those (extremely) rare cases where the points (x'_i, y'_i, z'_i) and (x'_j, y'_j, z'_j) are far from each other (and not too close to the edges). For example, when $(x'_i, y'_i, z'_i) = (.02, .2, .78)$ and $(x'_j, y'_j, z'_j) = (.8, .08, .12)$, peer update will result in a credence $\approx (.41, .14, .45)$. If one then conditions on $\neg P_Z$ one will obtain posterior $\approx (\frac{41}{55}, \frac{14}{55}, 0)$. On the other hand if i and j first condition on $\neg P_Z$ they will come to have credences $(\frac{1}{11}, \frac{10}{11}, 0)$ and $(\frac{10}{11}, \frac{1}{11}, 0)$, respectively. If now they perform a peer update, preservation of zero considerations and symmetry imply a posterior of $(\frac{1}{2}, \frac{1}{2}, 0) \not\approx (\frac{41}{55}, \frac{14}{55}, 0)$. So commutativity of conditionalization and peer update simply doesn't follow from naive Bayesian (i.e. coherence) considerations alone.

Fitelson and Jehle did go on to say (as a hedge, perhaps): "But even if we don't think of EWV-rules as equivalent to some conditionalization, we think [commutativity with conditionalization] should remain a desideratum for EWV-updates. We don't have the space to defend this claim here." The claim requires *some* sort of defense. If coherence isn't available to ground it, what is?

One would have to begin, we believe, with an attempt to explain away examples such as the foregoing one in which something near to difference splitting is rationally mandated. Note that the example favors difference splitting because the Euclidean midpoint of the segment connecting the ordered pairs whose barycentric coordinates correspond to the agents' priors minimizes the sum of the absolute deviations of the approximating bivariate normals (and so is near to the expectation of their mean). One would have to say, then, why the Euclidean metric is the wrong one to be working with in the generic situation in which two agents have identical evidence but different priors.

On the other hand, perhaps one would not have to say much here, for there is absolutely no reason to think that the Euclidean metric *would* be an appropriate metric in this context. When measuring the distance from a probability measure $x = (x_1, \ldots, x_n)$ to another probability measure $y = (y_1, \ldots, y_n)$, the information distance–so called *Kullback-Leibler divergence* $KL(x, y) = \sum_{i=1}^{n} x_i \log \frac{x_i}{y_i}$, is a far more likely default candidate. And, as we shall see below, when i and j's common posterior is chosen so as to minimize the sum of these distances to their respective priors, the resulting update scheme *does* commute with conditionalization.

Alternatively, one can argue that commutativity is appropriate in cases where one doesn't have any reason to suspect it would fail. First one would argue that, in a case where i knows j's prior credence function and knows that $\neg C$ (say) is the case, knowledge of her own current credence function "screens off" the significance of her prior credence in C. The example involving barycentric coordinates shows why one cannot make this assumption on the basis of coherence considerations alone...the joint distribution of the two priors over the (ordered) partition in question–(A, B, C), say–may be such that i's prior credence in C yields information about the relative likelihoods of A and B beyond that provided by j's prior credence function and her own current credence function alone. That example relies heavily on the two agents' different evidence, however. One might argue that the same-evidence situation is different.

As an example, consider again the scenario in which i's original prior on (A, B, C) is $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ and j's is $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. Suppose that if the agents first perform a peer update, their common posterior will be (x, y, z). (In this case $y = z$ follows from some seemingly innocuous permutation considerations, but we don't require this here.) If they next condition on $\neg C$, their common posterior will be $(\frac{x}{x+y}, \frac{y}{x+y}, 0)$. We want to say that if the agents instead: (1) condition on $\neg C$; then (2) perform a peer update, they will arrive at the same common posterior $(\frac{x}{x+y}, \frac{y}{x+y}, 0)$. It's clear that their credences after (1) will be $(\frac{1}{2}, \frac{1}{2}, 0)$ and $(\frac{1}{3}, \frac{2}{3}, 0)$ respectively.

Consider an alternate scenario in which the agents instead; (0) share their prior credences in C; then (1) condition on $\neg C$; and finally (2) perform a peer update. Here the agents will definitely land at common posterior $(\frac{x}{x+y}, \frac{y}{x+y}, 0)$ after Step (2), for in this case they would have acquired exactly the same information as in the original case where they first performed a peer update, then conditioned on $\neg C$. It is arguable, moreover, that j's evidence in Step (0), namely i's prior credence in C, should in the absence of any reason for thinking the contrary be treated as neutral with respect to the relative likelihoods of A and B. Indeed, expert testimony as to probability of C (say) is considered *a* or *the* paradigm case in which so-called *Jeffrey conditionalization* (adopt the expert's credence in C as your own, preserve ratios of other partition cells; Jeffrey 1965) is appropriate.[9]

At the beginning of Step (0) in the alternate scenario, for i to learn j's credence in C is as informative as for i to learn her own posterior in C at the conclusion of Step (0) (as in Section 2, we take posterior credence to be a strictly increasing function of j's prior when i's prior is fixed), and so formally equivalent to taking expert (that of i's future self) testimony as to the likelihood of C. If that's right, however, then the agents' credences after performing Step (0) will have the form (a, a, k) and $(b, 2b, l)$ respectively, so that after Step (1) they will have credences $(\frac{1}{2}, \frac{1}{2}, 0)$ and $(\frac{1}{3}, \frac{2}{3}, 0)$ respectively, exactly as in the actual scenario of the previous paragraph. Since the final posteriors in the alternate scenario are $(\frac{x}{x+y}, \frac{y}{x+y}, 0)$, then, the final posteriors in the scenario from the previous paragraph will be $(\frac{x}{x+y}, \frac{y}{x+y}, 0)$ as well. That is, peer update commutes with conditionalization.

These considerations are convincing to us. Accordingly, we accept Fitelson and Jehle's desideratum; in the absence of known protocols to the contrary (as in some not-same evidence cases), peer update should commute with conditionalization.

4. Peer update and relative entropy minimization

The update scheme we consider to be the clear frontrunner (and only serious extant contender for "default" status) is a case of a rule known by various names, including *logarithmic opinion pool* and *geometric opinion pool*. Although many of its positive features are well known, we will argue its virtues in a couple of new ways. We shall also discuss a virtue that, while not original, has been neglected in the philosophical literature–namely that geometric pooling minimizes the sum of the Kullback-Leibler divergences from the priors to the common posterior.

Note that Easwaran et. al. (2016, Section 10.6) discuss the family of geometric opinion pools, but reject those we find most apt. (In their notation, these are the ones for which the exponents w_i sum to unity. But they write "It is clear that an agent should assign her own exponent to be $w_1 = 1$.") This is in part because they are concerned with the different (in fact independent) evidence situation. Indeed, they motivate their approach via a scenario in which i's distribution for j's initial credence in A, conditional on A, is given by the density function $f_A(q) = 2q$,

[9] Though it's difficult to say non-circularly what makes Jeffrey conditionalization appropriate in a given case, the idea is that it should be valid when the information one gets about which cell of the background partition obtains isn't contaminated by further information about the likelihoods of your target proposition conditional on the various cells. In the case of expert testimony as to the probabilities of the cells, there seems to be minimal chance of contamination.

independently of i's initial credence p in A. That's coherent[10] but it can't be a same-evidence case, for in such a case ideal epistemic credence screens off j's credences from the truth value of A. (Cf. Easwaran et. al. 2018 Section 10.9: "the bias of the coin screens off Q's credences from the outcome of the toss".)

On the other hand in Section 10.2 of their paper they do entertain a same-evidence scenario (as such), and still see things very differently from us. In their example,

$$\frac{pqr}{pqr + 16(1-p)(1-q)(1-r)}$$

is selected as posterior credence in A when three peers have priors p, q and r, respectively, assuming a "default" credence of .8. They observe that if all three peers hold to the default value, there will be no synergy. But if $p = q = r = .9$, the agents will update to $\frac{(.9)^3}{(.9)^3+16(.1)^3} \approx .9785$. That's synergy, so we reject the reasoning leading to this formula for the same-evidence case. (Cf. footnote 4.)

According to the version of geometric opinion pooling appropriate to the Equal Weight View, if A and B are two cells of the partition under consideration and $r_i = \frac{P_i(A)}{P_i(B)}$, $r_j = \frac{P_j(A)}{P_j(B)}$ are the ratios assigned by i and j to these cells' respective probabilities, the corresponding ratio arising from the common posterior ought to be the geometric mean of r_1 and r_2. In the two-cell case $(A, \neg A)$, therefore, one updates to $(u, 1-u)$ given priors $(x, 1-x)$ and $(y, 1-y)$, where

$$\frac{u}{1-u} = \left(\frac{x}{(1-x)}\frac{y}{(1-y)}\right)^{\frac{1}{2}}. \tag{1}$$

More generally, suppose n peers have prior credences $(a_{i1}, a_{i2}, \ldots, a_{ik})$ over partition $(A_1, A_2, \ldots, A_k)$, $1 \le i \le n$. Common posterior is $(b_1, b_2, \ldots, b_n)$, where

$$b_j = \frac{\prod_{i=1}^{n} a_{ij}^{1/n}}{\sum_{t=1}^{k} \prod_{i=1}^{n} a_{it}^{1/n}}$$

Call this scheme *EGOP* (For *Equal-weight Geometric Opinion Pooling.*) We'll give three informal arguments in its favor for the case where i's prior is $(\frac{1}{5}, \frac{4}{5})$ and j's is $(\frac{1}{2}, \frac{1}{2})$. That is, we'll argue in three different ways for common posterior $(\frac{1}{3}, \frac{2}{3})$. In an appendix, we'll present more formal (and more general) versions.

The first argument appears in Abbas (2009), and proceeds by relative entropy. The Kullback-Leibler divergence from $(u, 1-u)$ to $(\frac{1}{5}, \frac{4}{5})$ is given by $u \log \frac{u}{1/5} + (1-u) \log \frac{1-u}{4/5}$; the Kullback-Leibler divergence from $(u, 1-u)$ to $(\frac{1}{2}, \frac{1}{2})$ is given by $u \log \frac{u}{1/2} + (1-u) \log \frac{1-u}{1/2}$. The sum of these quantities is

[10] Imagine q, x, y and z are selected independently uniformly at random on the unit interval. A is the event $x < q$. q is reported to j and becomes her initial credence in A. $p = \max\{y, z\}$ if A is true, otherwise $p = \min\{y, z\}$. p is reported to i and becomes her initial credence in A–we leave it to the reader to verify that the situation is now as described above.

$$H(u) = 2u \log u + 2(1-u)\log(1-u) - u \log \frac{1}{5} - (1-u)\log\frac{4}{5} - \log\frac{1}{2}.$$

The minimum of H occurs where

$$H'(u) = 2\log\frac{u}{1-u} + \log 4 = 0.$$

A quick calculation gives $(u, 1-u) = (\frac{1}{3}, \frac{2}{3})$, in agreement with (1).

For the second argument, imagine that a fair coin will be tossed if and only if $\neg A$ obtains. If i and j expand the algebra to accomodate the toss then of course their expanded priors will be $(\frac{1}{5}, \frac{2}{5}, \frac{2}{5})$ and $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$. Suppose now they were to condition on the disjunction of the first two cells (the event $A \vee (\neg A \wedge heads)$, say) and then perform a peer update. Their credences after conditioning will be $(\frac{1}{3}, \frac{2}{3}, 0)$ and $(\frac{2}{3}, \frac{1}{3}, 0)$. Symmetry and preservation of zero considerations now indicate that their credences will be $(\frac{1}{2}, \frac{1}{2}, 0)$ after the peer update.[11]

When i and j first condition on $A \vee (\neg A \wedge heads)$ then perform a peer update (as in the previous paragraph), they learn, in the second step, each others' priors in $\neg A \wedge tails$. (Since the coin is uncontroversially fair.) It follows that i and j gain no more information if they first perform a peer update and then condition; hence they will arrive at posterior $(\frac{1}{2}, \frac{1}{2}, 0)$ by this path as well. After the initial peer update, then, they must have credences of the form (x, x, y). On the other hand, permutation considerations point to $x = y$. That implies that i's posterior credence in A, when she considers the coin toss, is $\frac{1}{3}$. The final step is then that peer update should commute with marginalization onto the original sub-algebra.

Beware: i mustn't subscribe to the commutativity of peer update and marginalization in general. What justifies it in this case (the proponent will say) is that the ratio of the sizes of the to-be-amalgamated subcells is uncontroversial. Indeed, in the case where i and j first marginalize, then update, i's credences first evolve from $(\frac{1}{5}, \frac{2}{5}, \frac{2}{5})$ to $(\frac{1}{5}, \frac{4}{5})$, whereupon she learns j's post-marginalization credence function, namely $(\frac{1}{2}, \frac{1}{2})$. Since the coin is uncontroversially fair, however, this gives her knowledge of j's pre-marginalization credence function as well, namely $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$. So she acquires the same information she would acquire if the agents were to perform a peer update first, then marginalize. Accordingly, the update and the marginalization ought to commute, *in this special case*.[12]

[11] We consider resistance to this step in the argument below; the other steps look more secure.

[12] One may make a fruitful comparison to van Fraassen's (1981) “Infomin” solution (which, tellingly, is also that given by Kullback-Leibler divergence minimization) to the Judy Benjamin problem here. When Judy receives a message yielding information about the relative sizes she ought to assign the Red regions, this may (says Infomin) influence her credence in *Blue*–but not in a case where the message fails to alter Judy's relative credences in the Red regions (*a fortiori*, in a case where Judy knows this in advance).

For the third argument, observe that when an update rule commutes with conditionalization, the posterior ratio $P(A) : P(B)$ is a function of the two prior ratios $P_i(A) : P_i(B)$ and $P_j(B) : P_j(B)$. For example (the general case is not more subtle), consider two scenarios in which these prior ratios are equal, say

Scenario 1: $(P_i(A), P_i(B), P_i(C)) = (\frac{1}{10}, \frac{3}{10}, \frac{6}{10})$, $(P_j(A), P_j(B), P_j(C)) = (\frac{2}{7}, \frac{1}{7}, \frac{4}{7})$;

Scenario 2: $(P_i(A), P_i(B), P_i(C)) = (\frac{1}{6}, \frac{3}{6}, \frac{2}{6})$, $(P_j(A), P_j(B), P_j(C)) = (\frac{2}{5}, \frac{1}{5}, \frac{2}{5})$.

Denote posterior credence upon peer update in Scenario k by (a_k, b_k, c_k), $k = 1, 2$. The claim is that the ratios $a_1 : a_2$ and $a_2 : b_2$ coincide. Note that if the agents first conditionalize on $\neg C$, they will arrive at

$$(P_i(A), P_i(B), P_i(C)) = (\frac{1}{4}, \frac{3}{4}, 0),\ (P_j(A), P_j(B), P_j(C)) = (\frac{2}{3}, \frac{1}{3}, 0)$$

in either scenario. So if they do this and go on to a peer update, they will acquire the same posterior $(a, b, 0)$ in either scenario. Assuming that peer update commutes with conditionalization, it follows that if they first perform a peer update and then conditionalize on $\neg C$ they will again acquire posterior $(a, b, 0)$ in either scenario. So the ratios $a_k : b_k$ and $a : b$ coincide, $k = 1, 2$.

For $x, y \geq 0$, write $g(x, y) = z$ if it is the case that whenever $P_i(B) = xP_i(A)$ and $P_j(B) = yP_j(A)$, posterior credence upon performing a peer update has the form (a, az, c). We assume an insensitivity to permutations of (A, B, C) whereby also $g(x, y) = z$ if it is the case that whenever, e.g., $P_i(C) = xP_i(B)$ and $P_j(C) = yP_j(B)$, posterior credence upon performing a peer update has the form (a, b, bz).

We claim that $g(tx, sy) = g(t, s)g(x, y)$. To see this, suppose that

$$P_i(C) = tP_i(B) = txP_i(A) \text{ and } P_j(C) = sP_j(B) = syP_j(A).$$

Posterior credence will on the one hand have the form $(a, ag(x, y), c)$, and on the other hand it will have the form $(a, b, bg(t, s))$. Putting these together, posterior credence will have the form $(a, ag(x, y), ag(x, y)g(t, s))$. Finally, since

$$P_i(C) = txP_i(A) \text{ and } P_j(C) = syP_j(A),$$

posterior credence will have the form $(a, b, ag(tx, sy))$ (more permutation invariance is assumed here), yielding $g(tx, sy) = g(t, s)g(x, y)$.

This is enough, assuming continuity of g, to conclude that $g(x, y) = x^n y^m$ for some real exponents n and m. Add now the premise that i and j should "hold fast" when their priors agree (so that $g(x, x) = x$, in particular), one gets the restriction $n + m = 1$. Finally, if one wants updating to conform to the Equal Weight View, one should have $n = m$, i.e. $n = \frac{1}{2} = m$, resulting in *EGOP*. (Easwaran et. al. 2016 advocate for $n = m = 1$, which produces synergy.)

Unlike split-the-difference, *EGOP* can be coherently implemented with a peer. For imagine a proposition-valued random variable P. Denote i's initial credence in P by x and j's initial credence in P by y. Suppose that i regards j as a peer and updates by *EGOP*. Assume for simplicity that i's distribution for (x, y) is distributed on eight pairs, with weights as indicated in Table 1.

It is now easy to see that i is Reflection-obeying. For example, if P_0 is such that $x = \frac{1}{5}$ then, upon learning that $P = P_0$, i's posterior distribution for u will be $(\frac{2}{5}, \frac{3}{5})$ on $(0, \frac{1}{3})$. In particular, $E(u|P = P_0) = \frac{2}{5} \cdot 0 + \frac{3}{5} \cdot \frac{1}{3} = \frac{1}{5}$. The remaining cases are similar, so i's behavior under this model exhibits diachronic coherence.

TABLE 1

x	y	u	Prob
0	1/5	0	1/10
1/5	0	0	1/10
1/5	1/2	1/3	3/20
1/2	1/5	1/3	3/20
1/2	4/5	2/3	3/20
4/5	1/2	2/3	3/20
4/5	1	1	1/10
1	4/5	1	1/10

Though simple, this model's features render it a plausible (toy) implementation of the Equal Weight View: apart from employing the EWV-friendly (1), i's joint distribution for (x, y) and update function $u(x, y)$ are symmetric in the variables x and y, implying that, from i's perspective, her own credences and those of j are treated interchangeably. Since also $x \neq y$ with positive probability, we may conclude that Theorem 1 doesn't overgeneralize in the manner of Theorem 2.

5. OBJECTIONS ENTERTAINED

Having seen some of *EGOP*'s virtues, it is natural to ask about its potential shortcomings. The first we shall consider is that the scheme does not preserve (all) independences.[13] For example, if (A, B, C, D, E) is a partition over which agents i and j have credence functions $\mu_i = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0, \frac{1}{4})$ and $\mu_j = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0)$ respectively, it is easily checked that they will update to $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$ upon mutual disclosure of their priors. What might seem odd about that is the fact that both i and j initially considered the sets $F = A \cup B$ and $G = B \cup C$ to be probabilistically independent, whereas they do not so judge after updating.

But since diachronic coherence and peerhood (as we have defined these) are non-negotiable, one should almost surely not raise a given zero probability, and so too should update a cell's probability to zero when one's peer has zero credence in

[13] It does preserve independence of joined partitions; cf. Easwaran et. al. (2016), Theorem 9.2 (proof of which easily generalizes to our situation).

it. By symmetry, then, $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$ *is* the correct posterior, so the intuition to preserve independences at this level of generality must simply be misguided.[14]

The second objection is related to the first, and concerns the fact that *EGOP* doesn't commute with marginalization. Consider two agents who have credence functions (a_i, b_i, c_i) and (a_j, b_j, c_j) over a measurable partition (A, B, C). Suppose they share only their credences in A, learning that $a_i = a_j = \frac{1}{2}$. How should they update in response? A naive intuition is that they should hold fast.

However, suppose that the agents intend presently to share their full priors. By Reflection, their current credences in A should be the expectations of these credences after the ensuing peer update. But if the agents subscribe to *EGOP* then their joint credence in A after updating will definitely not be less than $\frac{1}{2}$, and may well be greater than $\frac{1}{2}$. Indeed, it will be greater than $\frac{1}{2}$ precisely when $b_i \neq b_j$, so unless the agents assign full credence to the event $b_i = b_j$, their current credences in A ought now to be greater than $\frac{1}{2}$.

To bring the seeming oddness of this into relief, let us pretend that these agents only had the partition (A, B, C) described to them *after* sharing their priors in A. (Up that that time, if they had been contemplating any partition, it was $(A, \neg A)$.) Clearly they would retain their initial credences in A up to the time that the events B and C were described to them. Are we to take it, then, that the mere act of calling attention to a refinement of the initial partition is what led to their increased credences in A? If that's right then we could just as easily have *decreased* their credences in A, by calling attention to a partition $(D, E, \neg A)$.

Recall, though, that indifference principles typically *are* partition relative. If your attention is called to a 2 cell partition (A, D) and you have no reason to favor either cell of the partition, your credences ought surely to be $(\frac{1}{2}, \frac{1}{2})$ over the partition. But if next you are told that D is the disjoint union of events B and C, what will be your new credence function? It seems uncontentious that this should depend on which partition(s) you take the indifference principle to apply to. Why should the current case be different? If application of *EGOP* is appropriate over some three cell partition (A, B, C), then one ought not expect it to be appropriate to apply the principle over $(A, B \cup C)$. (And vice versa.)

Note that partition relativity fails to undermine our second argument for *EGOP*. Call a partition over which it is appropriate to apply *EGOP admissible*. In the second argument, one starts with an admissible partition $P_1 = (A, \neg A)$ then tosses a fair coin if and only if $\neg A$, thus refining the partition to $P_2 = (A, (\neg A) \wedge heads, (\neg A) \wedge tails)$. One then conditions on the union of the first two cells of

[14] A referee worried that "you and I might agree...that C does not cause E and then after (updating) agree...that C does cause E." Things are perhaps clearer in the different evidence case. Returning to the example in the text, if I've learned "not D" and you've learned "not E" then after pooling our data it's not contentious that our priors are as described and that we update to $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, 0)$. So that independences aren't preserved by peer update is a brute fact.

P_2, obtaining the partition $P_3 = (A, (\neg A) \wedge heads)$. It is implicitly assumed that P_3 is admissible. But why? Well, we have also been implicitly assuming that any partition obtained from an admissible partition via conditionalization is again admissible. (It would make little sense to speak of commutativity of *EGOP* with conditionalization otherwise.) So if P_2 is admissible then P_3 is. The question, however, is why P_2 should be admissible. As it happens, there is a completely satisfactory answer to this question. Namely, P_2 is admissible because P_1 is and the coin is *uncontroversially* fair. More generally, if a finite partition P is admissible and Q is a finite refinement of P then Q will be admissible provided that a certain "rigidity" condition holds over P. That is, provided that, for every cell A of P and every cell B of Q, the probability of B conditional on A is public knowledge. (One may check that when this condition is met, applying *EGOP* over Q is consistent with applying *EGOP* over P.)

A third objection relates to the rejection of linear pooling, and was formulated by a referee thus: "A more cautious conclusion would be that this update rule comes with conditions of applicability, for instance that you do not also accept the principle **PN**." The referee suggested that one adopting this line might think that "epistemological problems cannot be settled by mathematical argument alone". The latter point may be a misunderstanding: we do not even assume here that we know the proposition P under debate. All we know is our credence in P, and our peer's credence in P. We needn't even know who the peer is, and indeed we needn't even be a party to the compromise–we could be arbitrating on behalf of two unknown agents assumed to be peers. All we have to go on are the credences; there isn't any data on which to base a "non-mathematical" analysis. As for rejecting **PN**, it's difficult to see how one could justify having credence .4 in $T(P)$ upon learning that "peer" credence in P is .3 (say). To see why, again step outside and pretend you are an arbitrator. If you defer to the credences of one of the agents, but correct the other's for ostensible systematic error, you cannot reasonably be said to consider them peers.[15]

6. Appendix

The following occurs as Proposition 4 in Abbas (2009), but has been largely ignored by philosophers. This may be in part because Abbas also proves that if one substitutes "inverse relative entropy" for "relative entropy", the resulting optimization problem is solved by *linear* pooling. That might suggest to some readers that entropy methods can be used to motivate either approach. But the inverse relative entropy result strikes us as philosophically insignificant: when one measures the Kullback-Leibler divergence from P to Q, one is treating P as actual. It's nonsense to treat even one prior as actual in the act of updating to a

[15]If you attribute greater *random* error to the first agent you might view them as "equally competent", but this violates our agreement that we know nothing beyond the credences.

posterior considered as non-actual–the inverse relative entropy application in fact treats disparate priors as actual at different locations within a single equation.[16]

Theorem 3. Let $(A_1, A_2, \dots, A_n)$ be an ordered measurable partition. Suppose that i and j have priors

$$\mu_i = (a_1, a_2, \dots, a_n) \quad \text{and} \quad \mu_j = (b_1, b_2, \dots, b_n)$$

over $(A_1, A_2, \dots, A_n)$. Let $H(\mu) = KL(\mu, \mu_i) + KL(\mu, \mu_j)$ take on its minimum value (as μ ranges over probability measures on $(A_1, A_2, \dots, A_n)$) at $\mu_0 = (c_1, c_2, \dots, c_n)$. Then for any fixed indices l, m, $1 \leq l \neq m \leq n$,

$$c_l^2 a_m b_m = c_m^2 a_l b_l. \tag{1}$$

In particular, if $a_m b_m \neq 0$ then $\frac{c_l}{c_m}$ is the geometric mean of $\frac{a_l}{a_m}$ and $\frac{b_l}{b_m}$.

Proof. Permuting indices if necessary, we may assume that $l = 1$ and $m = n$. Writing $0 \cdot \log 0 = 0$, $H(\mu)$ is continuous on a compact domain and so attains its minimum value at some $\mu_0 = (c_1, c_2, \dots, c_n)$. Plainly $c_1 = 0$ if $a_1 b_1 = 0$ and $c_n = 0$ if $a_n b_n = 0$; in either case, (1) follows.

We may therefore assume that $a_1 b_1 a_n b_n > 0$. Writing $\mu = (x_1, x_2, \dots, x_n)$,

$$H(\mu) = H(x_1, \dots, x_{n-1}) = \sum_{t=1}^{n} (2x_t \log x_t - x_t \log a_t b_t).$$

Since $x_n = 1 - x_1 - x_2 - \cdots - x_{n-1}$, the first partial derivative of H is

$$H_{x_1}(x_1, \dots, x_{n-1}) = 2 \log \frac{x_1}{x_n} + \log \frac{a_n b_n}{a_1 b_1}.$$

[16] Abbas offers no discussion about the "aptness" of the inverse relative entropy method. Douven and Romeijn (2011), however, who defend the one-half solution to the Judy Benjamin problem via inverse relative entropy, do, writing of it and the classical method (Infomin; see van Fraassen 1981): "they are not just formally but also conceptually very similar: where INFOMIN has you select the probability function that is RE-closest to your present probability function *as seen from your current perspective*, IRE minimization has you select the probability function that is RE-closest to your present probability function *as seen from the perspective you will have after adopting the probability function to be selected.*" (Emphases in original.) But first, Douven and Romeijn have it backwards–it is Infomin that has you select the probability function that is RE-closest to your present probability function as seen from the posterior's perspective. And second, in the Judy Benjamin problem one constrains one's posterior to a convex set not containing one's prior. To accept such a constraint is to already have admitted, in particular, that one's prior isn't actual, so to accept the constraint *and* view things "from you current perspective" is inconsistent. The methods aren't, then, "close cousins" at all–in fact inverse relative entropy minimization is, in this context, an unmotivated and specious patchwork.

Since H takes on its minimum value at μ_0 one must have $H_{x_1}(\mu_0) = 0$. Therefore, $\log(\frac{c_1}{c_n})^2 = \log \frac{a_n b_n}{a_1 b_1}$. Clearing logarithms, we obtain (1). qed

The following is a special case of the known fact that under geometric opinion pools, "finding the consensus distribution commutes with the process of revising distributions using a commonly agreed likelihood" (Genest and Zidek 1986).

Theorem 4. (Cf. Easwaran et. al. 2016, Theorem 9.3.) Let $(A_1, A_2, \ldots, A_n)$ be an ordered measurable partition. Let u be the update function having the property that

$$u((a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n)) = (c_1, c_2, \ldots, c_n),$$

where $c_l^2 a_m b_m = c_m^2 a_l b_l$, $1 \leq l \neq m \leq n$. Then u commutes with conditionalization on events in the algebra generated by $A_1, A_2, \ldots, A_n$.

Proof. Let E be a union of partition cells. By rearranging indices if necessary we may assume that $E = A_1 \cup A_2 \cup \cdots \cup A_k$ for some k. Suppose we first condition on E, then update. After conditioning on E, we get

$$\mu_i' = \Big(\frac{a_1}{a_1 + a_2 + \cdots + a_k}, \frac{a_2}{a_1 + a_2 + \cdots + a_k}, \ldots, \frac{a_k}{a_1 + a_2 + \cdots + a_k}, 0, \ldots, 0\Big)$$

and

$$\mu_j' = \Big(\frac{b_1}{b_1 + b_2 + \cdots + b_k}, \frac{b_2}{b_1 + b_2 + \cdots + b_k}, \ldots, \frac{b_k}{b_1 + b_2 + \cdots + b_k}, 0, \ldots, 0\Big).$$

Now $u(\mu_i', \mu_j') = (d_1, d_2, \ldots, d_k, 0, \ldots, 0)$, where

$$\begin{aligned} &d_l^2 \frac{a_m b_m}{(a_1 + a_2 + \cdots + a_k)(b_1 + b_2 + \cdots + b_k)} \\ =&d_m^2 \frac{a_l b_l}{(a_1 + a_2 + \cdots + a_k)(b_1 + b_2 + \cdots + b_k)}, \end{aligned}$$

which implies that $d_l^2 a_m b_m = d_m^2 a_l b_l$, $1 \leq l \neq m \leq k$.

Next we first update, then condition on E. We have $u(\mu_i, \mu_j) = (c_1, c_2, \ldots, c_n)$, where $c_l^2 a_m b_m = c_m^2 a_l b_l$, $1 \leq l \neq m \leq n$. Conditioning next on E, we get to $(d_1, d_2, \ldots, d_k, 0, \ldots, 0)$, where $(c_1 + c_2 + \cdots + c_k) d_m = c_m$, $1 \leq m \leq k$. Therefore

$$d_l^2 (c_1 + c_2 + \cdots + c_k)^2 a_m b_m = d_m^2 (c_1 + c_2 + \cdots + c_k)^2 a_l b_l,$$

so that $d_l^2 a_m b_m = d_m^2 a_l b_l$, $1 \leq l \neq m \leq k$. Since these equations clearly determine $d_1, d_2, \ldots, d_k$ subject to the constraint $d_1 + d_2 + \cdots + d_k = 1$, we are done. qed

Theorem 5. *EGOP* is the unique peer update scheme compatible with the following set of premises.

Premise 1. Update commutes with conditionalization. In particular admissibility is closed under conditionalization: if $\{A_1, A_2, \ldots, A_n\}$ is admissible for an agent a, then were a to conditionalize on

$$\bigcup_{i \in I = \{i_1, \ldots, i_t\}} A_i,$$

$(A_{i_1}, \ldots, A_{i_t})$ would then be admissible for a.

Premise 2. Update commutes with transparent (informationless) expansion. In particular admissibility is closed under transparent expansion. Suppose that if $(A_1, A_2, \ldots, A_n)$ is admissible and $\{B_{ij} : 1 \leq j \leq k_i\}$ partitions A_i, $1 \leq i \leq n$. If $Prob(B_{ij}|A_i)$ is public knowledge, $1 \leq i \leq n$, $1 \leq j \leq k_i$, then $(B_{11}, \ldots, B_{1k_1}, B_{21}, \ldots, B_{2k_2}, \ldots, B_{n1}, \ldots, B_{nk_n})$ is admissible.

Premise 3. Equal weighting. If (A, B) is admissible, agent a has credence function $(t, 1-t)$ over (A, B) for some $t \in (0, 1)$ and a's peer has credence function $(1-t, t)$ over (A, B) then a updates to $(\frac{1}{2}, \frac{1}{2})$ upon learning her peer's credences.

Proof. Let $(A_1, \ldots, A_n)$ be an admissible partition. Denote the credence functions of the agent and the peer by $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$, respectively. We must show that

$$u((a_1, a_2, \ldots, a_n), (b_1, b_2, \ldots, b_n)) = (c_1, c_2, \ldots, c_n),$$

where $c_i^2 a_j b_j = c_j^2 a_i b_i$, $1 \leq i < j \leq n$.

Assume, without loss of generality, that $a_i b_i \leq a_j b_j$. (Otherwise swap the roles of i and j in what follows.) Consider now two hypothetical sequences of actions.

First sequence: update, then conditionalize. Here the agents first perform a peer update over the partition $(A_1, \ldots, A_n)$, getting to common posterior $(c_1, \ldots, c_n)$, then conditionalize on $A_i \cup A_j$, getting to the credence function

$$\left(\frac{c_i}{c_i + c_j}, \frac{c_j}{c_i + c_j}\right) \tag{4}$$

over the partition (A_i, A_j), which is admissible by Premise 1.

Second sequence: conditionalize, then update. Here the agents first conditionalize on $A_i \cup A_j$, getting to the credence functions

$$\left(\frac{a_i}{a_i + a_j}, \frac{a_j}{a_i + a_j}\right) \text{ and } \left(\frac{b_i}{b_i + b_j}, \frac{b_j}{b_i + b_j}\right) \tag{5}$$

respectively, then perform a peer update. Since by Premise 1 their common posterior must then agree with (4), this shows that

$$u\left(\left(\frac{a_i}{a_i+a_j}, \frac{a_j}{a_i+a_j}\right), \left(\frac{b_i}{b_i+b_j}, \frac{b_j}{b_i+b_j}\right)\right) = \left(\frac{c_i}{c_i+c_j}, \frac{c_j}{c_i+c_j}\right). \tag{6}$$

Suppose now that the agents have conditionalized on $A_i \cup A_j$, but have not performed a peer update. So their current credence functions are as in (5). Introduce now events B_{j1} and B_{j2} (tossing a coin of appropriate known bias will do) that partition A_j such that the probability of B_{j1} conditional on A_j is public and equal to $\sqrt{\frac{a_i b_i}{a_j b_j}} \leq 1$. Consider now the following two hypothetical sequences of actions.

First sequence: update, then expand. Here the agents first perform a peer update, getting to $\left(\frac{c_i}{c_i+c_j}, \frac{c_j}{c_i+c_j}\right)$ by (6), then expand to the partition (A_i, B_{j1}, B_{j2}), getting to common posterior

$$\left(\frac{c_i}{c_i+c_j}, \frac{c_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{c_i+c_j}, r\right). \tag{7}$$

Second sequence: expand, then update. Here the agents first expand to the partition (A_i, B_{j1}, B_{j2}), getting to credence functions

$$\left(\frac{a_i}{a_i+a_j}, \frac{a_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{a_i+a_j}, r_i\right) \text{ and } \left(\frac{b_i}{b_i+b_j}, \frac{b_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{b_i+b_j}, r_j\right), \tag{8}$$

over the partition (A_i, B_{j1}, B_{j2}), which is admissible by Premise 2. They then perform a peer update. Since by Premise 2 their common posterior must then agree with (7), this shows that

$$u\left(\left(\frac{a_i}{a_i+a_j}, \frac{a_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{a_i+a_j}, r_i\right), \left(\frac{b_i}{b_i+b_j}, \frac{b_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{b_i+b_j}, r_j\right)\right) = \left(\frac{c_i}{c_i+c_j}, \frac{c_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{c_i+c_j}, r\right). \tag{9}$$

Suppose finally that the agents have expanded to (A_1, B_{j1}, B_{j2}), but have not performed a peer update. So their current credence functions are as in (8). Consider now the following two hypothetical sequences of actions.

First sequence: update, then conditionalize. Here the agents first perform a peer update, which gets them to the credence function on the right hand side of (9). They then conditionalize on $A_1 \cup B_{j1}$, which gets them to common posterior

$$\Big(\frac{c_i}{c_i + c_j\sqrt{\frac{a_i b_i}{a_j b_j}}}, \frac{c_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{c_i + c_j\sqrt{\frac{a_i b_i}{a_j b_j}}}\Big). \tag{10}$$

Second sequence: conditionalize, then update. Here the agents first conditionalize on $A_1 \cup B_{j1}$, which gets them to the credence functions

$$\Big(\frac{a_i}{a_i + a_j\sqrt{\frac{a_i b_i}{a_j b_j}}}, \frac{a_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{a_i + a_j\sqrt{\frac{a_i b_i}{a_j b_j}}}\Big) \text{ and } \Big(\frac{b_i}{b_i + b_j\sqrt{\frac{a_i b_i}{a_j b_j}}}, \frac{b_j\sqrt{\frac{a_i b_i}{a_j b_j}}}{b_i + b_j\sqrt{\frac{a_i b_i}{a_j b_j}}}\Big)$$

over (A_i, B_{j1}), which is admissible by Premise 1. Letting

$$t = \frac{a_i}{a_i + a_j\sqrt{\frac{a_i b_i}{a_j b_j}}},$$

the agents' credences are therefore $(t, 1-t)$ and $(1-t, t)$, respectively. So when they next perform a peer update, they get to common posterior $(\frac{1}{2}, \frac{1}{2})$ by Premise 3. On the other hand, this result must agree with that of (10) by Premise 2. So $c_i = c_j\sqrt{\frac{a_i b_i}{a_j b_j}}$, from which it follows that $c_i^2 a_j b_j = c_j^2 a_i b_i$, as desired. qed

Theorem 6. *EGOP* is the unique peer update scheme compatible with the following set of premises.

Premise 1. Update commutes with conditionalization. In particular admissibility is closed under conditionalization: if $\{A_1, A_2, \ldots, A_n\}$ is admissible for a, then were a to condition on

$$\bigcup_{i \in I = \{i_1, \ldots, i_t\}} A_i,$$

$(A_{i_1}, \ldots, A_{i_t})$ would then be admissible for them.

Premise 2. Equal weighting. If (A, B) is admissible, a has credence function $(t, 1-t)$ over (A, B) for some $t \in (0, 1)$ and a's peer has credence function $(1-t, t)$ over (A, B) then the agent updates to $(\frac{1}{2}, \frac{1}{2})$.

Premise 3. Continuous update. The update function $u(p, q)$ is continuous.

Premise 4. Update commutes with permutations of the (ordered) partition.

Premise 5. Holds fast. If the agents' priors agree, these don't change on update.

Proof. This proof is an elaboration on the third argument from Section 4. First:

Lemma A Let n and an ordered partition $(A_1, \ldots, A_n)$ be fixed, and let $1 \leq i \neq j \leq n$. If update commutes with conditionalization, then writing

$$u\Big((a_1,\ldots,a_n),(b_1,\ldots,b_n)\Big)=(c_1,\ldots,c_n),$$

the ratio $c_i : c_j$ is a function of the ratios $a_i : a_j$ and $b_i : b_j$. That is to say, if

$$u\Big((d_1,\ldots,d_n),(e_1,\ldots,e_n)\Big)=(f_1,\ldots,f_n)$$

and we have $d_i : d_j \equiv a_i : a_j$ and $e_i : e_j \equiv b_i : b_j$, then also $f_i : f_j \equiv c_i : c_j$.

Proof. Assume first that $(a_i+a_j)(b_i+b_j)>0$. Imagine two sequences of actions.

First sequence. Conditionalize, then update. After conditionalizing on $A_i \cup A_j$), the agents a, b, d and e have credence functions $(\frac{a_i}{a_i+a_j}, \frac{a_j}{a_i+a_j})$, $(\frac{b_i}{b_i+b_j}, \frac{b_j}{b_i+b_j})$, $(\frac{d_i}{d_i+d_j}, \frac{d_j}{d_i+d_j})$ and $(\frac{e_i}{e_i+e_j}, \frac{e_i}{e_i+e_j})$ respectively. By hypothesis, d's credence are equal to a's and e's are equal to b's. Next a and b perform an update, as do d and e. Obviously a and d wind up with the same credences.

Second sequence. Update, then conditionalize. After updating (a with b and d with e), a and d have credence functions $(c_1,\ldots,c_n)$ and $(f_1,\ldots,f_n)$. After they then conditionalize on $(A_i \cup A_j)$ they have credence functions $(\frac{c_i}{c_i+c_j}, \frac{c_j}{c_i+c_j})$ and $(\frac{f_i}{f_i+f_j}, \frac{f_i}{f_i+f_j})$. By Premise 1 these are the same credences they had at the conclusion of the first sequence, which recall were equal. So $f_i : f_j \equiv c_i : c_j$. qed

So for $x, y > 0$ we can write $g(x,y)=z$ if whenever $1 \le i \ne j \le n$, $(a_i+a_j)(b_i+b_j)>0$, $a_i = xa_j$ and $b_i = yb_j$, one has $c_i = zc_j$. We claim that $g(tx,sy)=g(t,s)g(x,y)$. For suppose that $a_3 = ta_2 = txa_3$ and $b_3 = sb_2 = syb_1$. a's post-update credence will have the form $(c_1, c_1g(x,y), c_3, \ldots)$. On the other hand ,it will have the form $(c_1, c_2, c_2g(t,s), \ldots)$. Putting these together, a's post-update credence will have the form $(c_1, c_1g(x,y), c_1g(x,y)g(t,s))$. Finally, since $a_3 = txa_1$ and $b_3 = syb_1$, posterior credence will have the form $(c_1, c_2, c_1g(tx,sy), \ldots)$, yielding $g(tx,sy)=g(t,s)g(x,y)$.

This brings us to our final lemma:

Lemma B. Let $g : (0,\infty)\times(0,\infty)$ be continuous and satisfy $g(tx,sy)=g(t,s)g(x,y)$. Then there exist real n, m such that for all $x, y > 0$, one has $g(x,y)=x^n y^m$.

Proof. Fix n and m so that $g(2,1)=2^n$ and $g(1,2)=2^m$. It's clear that for $k \in \mathbf{N}$ one has $g(2^{1/k},1)^k = g(2,1) = 2^n$, so that $g(2^{1/k},1)=2^{n/k}$, and similarly $g(1,2^{1/k})=2^{m/k}$. Fixing k, it's now clear that for $l \in \mathbf{N}$ one has $g(2^{l/k},1)=g(2^{1/k},1)^l=(2^{n/k})^l=(2^{l/k})^n$, and similarly $g(1,2^{l/k})=(2^{l/k})^m$. Let x and y be positive reals. Choosing a sequences of rationals $r_i = l_i/k_i$ converging to x and invoking continuity, one then obtains that $g(x,1)=x^n$. Similary $g(1,y)=y^m$, so that $g(x,y)=g(x,1)g(1,y)=x^n y^m$. qed

Premise 5 implies that $g(x, x) = x$, so the m and n found in Lemma B must satisfy $m + n = 1$. Premise 2 now tells us that $g(\frac{1-t}{t}, \frac{t}{1-t}) = 1$ for $t \in (0, 1)$. So $x^{n-m} = g(x, \frac{1}{x}) = 1$ for all positive x yielding $n = m = \frac{1}{2}$. Making a final appeal to continuity in the case where one has a zero ratio, *EGOP* follows. qed

References

Abbas, A. 2009. A Kullback-Leibler view of linear and log-linear pools. *Decision Analysis.* 6:25-37.

Bradley, Richard. 2017. Learning from others: conditioning versus averaging. *Theory and Decision.* 85:5-20.

Dawid, A. P., M. H. DeGroot and J. Mortera. 1995. Coherent combination of experts' opinions. *Test.* 4:263-313.

Douven, I. and J-W Romeijn. 2011. A new resolution of the Judy Benjamin Problem. *Mind.* 120:637-670.

Easwaran, K., L. Fenton-Glynn, C. Hitchcock and J. D. Velasco. 2016. Updating on the credences of others: disagreement, agreement, and synergy. *Philosopher's Imprint.* 16(11):1-39.

Elga, A. 2007. Reflection and disagreement. *Nous.* 41(3):478-502.

Fitelson, B. and D. Jehle. 2009. What is the "Equal Weight View"? *Episteme.* 6:280-293.

Genest, C. and Zidek, J.V. 1986. Combining probability distributions: A critique and an annotated biliography. *Statistical Science* 1:114-135.

Jeffrey, R. 1965. *The logic of decision.* University of Chicago Press.

Kelly, T. 2010. Peer disagreement and higher order evidence. In Alvin I. Goldman & Dennis Whitcomb (eds.), *Social Epistemology: Essential Readings.* Oxford University Press. pp. 183-217.

Kullback, S. and R.A. Leibler. 1951. On information and sufficiency. *Annals of Mathematical Statistics.* 22(1):79-86.

Nissan-Rozen, I. and L. Spectre. 2017. A pragmatic argument against equal weighting. *Synthese.* Forthcoming.

Schervish, M.J., T. Seidenfeld and J.B. Kadane. 2004. Stopping to reflect. *Journal of Philosophy.* 101:315-322.

Van Fraassen, B. C. 1981. A Problem for Relative Information Minimizers in Probability Kinematics. *The British Journal for the Philosophy of Science* 1981:375-379.

Van Fraassen, B.C. 1984. Belief and the will. *Journal of Philosophy.* 81:235-256.

RMCCTCHN@MEMPHIS.EDU