DIFFERENCE-SPLITTING AND THE EQUAL WEIGHT VIEW

ABSTRACT. Dawid, DeGroot and Mortera showed, a quarter century ago, that any agent who regards a fellow agent as a peer–in particular, defers to the fellow agent's prior credences in the same way that she defers to her own– and updates by split-the-difference is prone (on pain of triviality) to diachronic incoherence. On the other hand one may show that there are special scenarios in which Bayesian updating approximates difference splitting, so it remains an important question whether it remains a viable (approximate) response to "generic" peer update. We look at arguments by two teams of philosophers (Fitelson & Jehle and Nissan-Rozen & Spectre) against difference splitting.

1. ON AN OLD TRIVIALITY RESULT FOR SPLIT-THE-DIFFERENCE

In the literature on peer disagreement, 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). Some philosophers have taken "splitting the difference" (i.e., adoption of the arithmetic mean) between competing peer credences to be constitutive of the Equal Weight 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.[1] Indeed, philosophers who are aware of these incoherence results may dispute that it applies to the situation that interests us, namely that of two peers having identical evidence but different priors. Bradley (2018), for example, rehearses such incoherence results, yet writes:

---

[1]Many philosophers addressing peer update fail to cite this paper. Indeed, being newcomers and working in isolation, we rediscovered the result and submitted it as original research thrice before receiving a report indicating that it was not. (Thanks to two anonymous referees here.)

> ...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.

We are baffled by this; considerations along the lines of Dawid et. al. (1995) clearly *do* "militate" (decisively) against split-the-difference, *considered as an implementation of the Equal Weight View*[2], in cases where two peers hold the same information but do not have almost surely identical prior credences.

We present our original proof of this old result, which is slightly different than extant versions. Suppose 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 that if i's credence in P_0 is x_0 then

---

[2] Splitting the difference with a non-peer can of course 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.

$$E_y(u(x,y)|P = P_0) = x_0,$$

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.) In particular, for almost every x_0 in the essential range of x one has

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

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 $E_x(u(x,y)|y = y_0) = y_0$.

PN cashes out the intuition that if i regards j as a peer then she has the same confidence in j's initial credences that she has in her own. 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 $T(P)$. 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 split-the-difference (and more "general linear pooling") 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 $Prob(x \neq y) = 0$.

Proof. For *reductio*, assume that $Prob(x \neq y) > 0$. The following is routine:

Lemma 1. If X is a random variable and $E(X) = k$ then

$$E(X - k|X > k)Prob(X > k) = E(k - X|X < k)Prob(X < k).$$

Since j is diachronically coherent almost surely, for almost every x_0 in the essential range of x one has

$$E_y((u(x,y)|x = x_0) = E_y(wx + (1 - w)y|x = x_0) = x_0.$$

But obviously $E_y(x|x = x_0) = x_0$, so in fact $E_y(y|x = x_0) = x_0$. By Lemma 1,

$$E_y(y - x|x = x_0 \wedge y > x)Prob(y > x|x = x_0)$$
$$= E_y(x - y|x = x_0 \wedge y < x)Prob(y < x|x = x_0).$$

Multiplying both sides of this equation by x_0, we can move occurrences of x_0 inside the expectations (since these are integrals in the variable y). So for a.e. x_0,

$$E_y(x(y - x)|x = x_0 \wedge y > x)Prob(y > x|x = x_0)$$
$$= E_y(x(x - y)|x = x_0 \wedge y < x)Prob(y < x|x = x_0).$$

Integrating over x_0, we get

$$E(x(y - x)|y > x)Prob(y > x) = E(x(x - y)|y < x)Prob(y < x). \qquad (1)$$

Since i regards j as a peer, for almost every y_0 in the essential range of y one has $E_x(wx + (1 - w)y|y = y_0) = y_0$ (by **PN**). But obviously $E_x(y|y = y_0) = y_0$, so in fact $E_x(x|y = y_0) = y_0$. By Lemma 1, it follows that

$$E_x(x - y|y = y_0 \wedge x > y)Prob(x > y|y = y_0)$$
$$= E_x(y - x|y = y_0 \wedge x < y)Prob(x < y|y = y_0).$$

Multiplying by y_0 and employing $x > y$ or $y > x$ where applicable,

$$E_x(x(x - y)|y = y_0 \wedge x > y)Prob(x > y|y = y_0)$$
$$\geq E_x(y(x - y)|y = y_0 \wedge x > y)Prob(x > y|y = y_0)$$
$$= E_x(y(y - x)|y = y_0 \wedge x < y)Prob(x < y|y = y_0)$$
$$\geq E_x(x(y - x)|y = y_0 \wedge x < y)Prob(x < y|y = y_0),$$

for a.e. y_0, with strict inequality wherever $Prob(x > y|y = y_0)$ is positive. Integrating over y_0, we therefore get

$$E(x(y - x)|x > y)Prob(x > y) > E(x(x - y)|x < y)Prob(x < y).$$

Multiplying both sides by -1, swapping sides and rearranging some comparisons, we get

$$E(x(y - x)|y > x)Prob(y > x) > E(x(x - y)|y < x)Prob(y < x),$$

contradicting (1). $\hfill$ qed

These considerations clearly 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, an extremely simple way of producing such schemes. Recall, we are interested in the case where our fellow agents have identical evidence and different priors. However, any scheme that is adapted to a specific scenario in which the agents begin with identical priors but then acquire different evidence and then condition on the other's new credence will automatically be coherent for some specific same-evidence case.

We sketch such an example. 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 of these motions is so small, the expectation x'_i of x conditional on x_i (which is precisely i's probability for P upon learning x_i) will, with very high probability, be extremely close to (distance much less than 10^{-12}, say) 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 the midpoint of their shared credences x'_i and x'_j.

To reiterate, though this is not a same evidence scenario, 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 as in the example. This fact motivates a search for additional, non-coherence based constraints on peer update that might tell in favor of or against schemes that approximate to difference splitting and purport to be viable in generic (i.e. representative of the mean) scenarios. We examine two such constraints in the following sections.

2. On an overrestrictive 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 presently demonstrate. It begins with a 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 now prove the following theorem, in which P is a proposition for which i has an initial credence, and i is committed to updating via linear pooling (with weight w) 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, linear pooling Equal Weighters do not regard their fellow agents as peers, which implies in particular that difference splitting cannot be a viable implementation of the Equal Weight View.[4]

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). By a variable weight view, we take them to mean a view on which $u(x, y)$ lies strictly between x and y, if $x \neq y$, and $u(x, x) = x$. But on this understanding, Theorem 2 doesn't, in fact, make room for such views; to what extent the argument attaches a deficient notion of peerhood to split-the-difference, plausible 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[5] the function $u(y)$

---

[4]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 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.

[5]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.

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 $x = Cr_i(P)$ ought to be equal to 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.[6] 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, if accepted; a certain k 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 k, the agent will recognize that the expected relative utility of her choice is small (non-existent, when $x = k$), and she may want to pass the bet to j. In at least some other cases (cases in which x and k are not close, typically), she will be inclined to want to field the bet herself, rather than pass it.

The proposed constraint is implausible and would, if valid, rule out too much. The most natural explanation for this data is that it isn't valid–it isn't a requirement of peerhood that for every such offer one should think that one's peer has expected

---

We believe the intuition to be an artifact of the more common case in which the agents have different evidence and the same priors. Suppose we have common prior distribution for the bias of a coin uniform on $[0, 1]$ and are each allowed to toss it once privately. If we reconvene and simultaneously announce credence of $\frac{1}{3}$ in the next toss of the coin landing *heads*, we will update not to $\frac{1}{3}$ but to $\frac{1}{4}$ (Laplace rule of succession). That is because our credences 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.

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

return not less than one's own.[7] 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 has serious problems, 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 just isn't 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 $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ 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 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,

---

[7]The alternative, it would seem, is that there is something wrong with Conciliatory views more broadly construed. For those who subcribe to the uniqueness of rational priors, that conclusion seems inevitable so far as ideal agents are concerned. Even for those, however, there is still an intuition that there ought to be a viable Equal Weight View for non-ideal agents.

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." It's of course a pity that they do not defend the claim, as it certainly requires defense.

Any such defense 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 agent 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.

Another option would be to employ the strategy of Section 1. 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 and the 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. One might argue that the generic, same-evidence situation is different, though.

Indeed, it seems that nothing would change if i hadn't formulated any original prior in C at all. She formulated her conditional probabilities $Prob(A|\neg C)$ and $Prob(B|\neg C)$, and that was enough for her to compute her current credence function, once $\neg C$ was learned. Should she now think, upon learning j's prior, that she might generate "evidence" for or against A, simply by going back and formulating, *ex post facto*, a prior for C? It's difficult to imagine where this evidence might come from, given that she isn't proposing to gather any new data; what she's proposing is simply an armchair exercise. (This should be contrasted with the triangle thought experiment, in which her prior in C was generated by an external process.) Finally, if i can say that her own prior credence in C is not relevant in a situation where she knows her current credence and j's prior, she can invoke peerhood to conclude that j's prior credence in C isn't relevant in a situation where she knows her own prior and j's current credence; at that point she can claim that conditionalization on $\neg C$ commutes with peer update after all.

We don't claim that these considerations are entirely conclusive, but they are compelling. Accordingly, we accept Fitelson and Jehle's desideratum; peer update ought (in the absence of any known protocols suggesting otherwise) to commute with conditionalization.

4. Peer update and relative entropy minimization

In this final section we present an update scheme that we believe should in some sense be considered "standard".[8] The rule isn't new. It is, for example, one member of a family of update rules considered in Easwaran et. al. (2016).[9] On the other hand, we will argue its virtues in a couple of new ways. Among these virtues, one finds that it commutes with conditionalization, preserves independences and minimizes the sum of the Kullback-Leibler divergences from the priors to the common posterior. (Since a general apology for this updating scheme isn't our main purpose, complete proofs of these somewhat trivial properties are omitted.)

According to this rule, 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 ratios assigned by i and j to these cells' respective probabilities, then 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

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

As motivation for (1) we'll provide two arguments (one heuristic) in the case where i's prior is $(1/5, 4/5)$ and j's is $(1/2, 1/2)$. That is, we'll give plausible reasons why, in this case, the common posterior should be $(1/3, 2/3)$.

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

$$H(u) = 2u \log u + 2(1-u) \log(1-u) - u \log 1/5 - (1-u) \log 4/5 - \log 1/2.$$

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).

---

[8]We don't claim that it would be immune from (coherence-based, even) objections. In particular, it would likely be subject to the usual array of criticisms that have plagued entropy maximization solutions in other contexts (see, e.g., Seidenfeld 1986 or Friedman and Shimony 1971). On the other hand, such solutions arguably do have formal merit as approximations to ideal behavior in extreme or limiting cases (we urge caution here, but see Vasudevan 2018). At any rate it isn't possible to criticize a method that one doesn't know about; this scheme, whatever its faults, marks a clear advance on difference splitting, and should be disseminated.

[9]These authors, curiously, do not favor the member of the family we are interested in. This is because they advocate for synergy in updating; we take issue with this in footnote 5.

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 $(1/5, 2/5, 2/5)$ and $(1/2, 1/4, 1/4)$. 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 $(1/3, 2/3, 0)$ and $(2/3, 1/3, 0)$. Symmetry and preservation of zero considerations now indicate that their credences will be $(1/2, 1/2, 0)$ after the peer update.

We next assume that i and j's peer updating commutes with conditionalization. (As we stated at the end of Section 3, we accept Fitelson and Jehle's proposed desideratum.) Thus they will also come to have credence function $(1/2, 1/2, 0)$ should they peer update then condition. After the initial peer update they will 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 $1/3$. 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 updating 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, that the coin is fair was always common knowledge, so i knew in advance that j's ratio would be equal to her own. If they were to have revealed credences over the original two-cell partition first, then these ratios, i wouldn't have obtained any "new information" from the latter reveal affecting her posterior in A.[10]

Unlike split-the-difference, the proposed scheme 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 in agreement with (1). We assume, for simplicity, that i's joint distribution for (x, y) is distributed on eight pairs, with weights as indicated in Table 1. Also, we imagine that learning the value of P provides no more information about the value of y than does learning x. (While not generic, these assumptions are clearly satisfiable, which is all that we require.)

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

---

[10]One may make a fruitful comparison to the "Infomin" solution to the Judy Benjamin problem (van Fraassen 1981) here. (Given that the solution being discussed agrees with that obtained by Kullback-Leibler divergence minimization, this should not be surprising.) When Judy receives a message yielding information about the relative sizes she ought to assign the Red regions, this may (according to 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. In particular (less contentiously, in fact), not in a case where Judy knows this in advance.

<div align="center">

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

</div>

The model moreover represents a plausible 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, then, it is not the case that $x = y$ almost surely, we may conclude that Theorem 1 doesn't overgeneralize in the manner of Theorem 2.

References

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.* 4:263313.

Easwaran, K., Luke Fenton-Glynn, Christopher Hitchcock and Joel D. Velasco. 2016. Updating on the credences of others: disagreement, agreement, and synergy. *Philosopher's Imprint.* Vol. 16. 39pp.

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

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

Friedman, Kenneth and Abner Shimony, Jaynes's maximum entropy prescription and probability theory, *Journal of Statistical Physics* **3** (1971), 381-384.

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, Ittay and Levi Spectre. 2017. A pragmatic argument against equal weighting. *Synthese.* To appear.

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

Seidenfeld, Teddy. 1986. Entropy and Uncertainty. *Philosophy of Science* 53:467-491.

Van Fraassen, Bas 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.

Vasudevan, Anubav. 2018. Entropy and insufficient reason: a note on the Judy Benjamin problem. *The British Journal for the Philosophy of Science.* Forthcoming.