Leitgeb and Pettigrew on Updating

Ben Levinstein

Abstract

Leitgeb and Pettigrew (2010a,b) argue that (1) agents should minimize the expected inaccuracy of their beliefs, and (2) inaccuracy should be measured via the Brier score. They show that in certain diachronic cases, these claims require an alternative to Jeffrey-Conditionalization. I claim that this alternative is an irrational updating procedure and that the Brier score, and quadratic scoring rules generally, should be rejected as legitimate measures of inaccuracy.

1   Introduction

In their (2010a; 2010b), Leitgeb and Pettigrew argue for a number of constraints on rational belief through appeal to the following norm:

ACCURACY: An epistemic agent ought to approximate the truth. In other words: she ought to minimize her inaccuracy. (2010a, 202)

Of course, for ACCURACY to be of much use to epistemologists, more must be said, and Leitgeb and Pettigrew are able to make it mathematically precise. In order to show how, I'll first streamline the discussion and assume all agents under consideration obey the following synchronic norm:

PROBABILISM: At any given time, an agent ought to have a probabilistically coherent credence function.

We can treat inaccuracy at a world formally as follows. Let A be a proposition, w a world, and let $\chi_w(A)$ be 1 if A is true at w and be 0 otherwise.[1] The idea is that χ_w—the characteristic function of w—represents the best possible credence to have at w, since it assigns maximal (minimal) credence to all propositions true (false) there. Leitgeb and Pettigrew argue that we should measure the *inaccuracy* of a probability function Cr at w by seeing how far it is from χ_w under the average of the squared Euclidean distance between them.[2] We end up with the *Brier score* as our inaccuracy measure:[3]

$$I(Cr, w) = \frac{1}{|W|} \sum_{v \in W} \left(\chi_w(\{v\}) - Cr(\{v\}) \right)^2$$

The Brier score has a natural generalization to a larger class of inaccuracy measures, known as *quadratic scoring rules*, which are of the following form:

$$I(Cr, w) = \sum_{i=1}^{N} \lambda_i \left(\chi_w(A_i) - Cr(A_i) \right)^2, \text{ where } \sum_{i=1}^{N} \lambda_i = 1, \text{ and each } \lambda_i > 0.$$

Quadratic scoring rules let us thumb the scale and count some propositions as more important than others. If we really care about whether it will rain tomorrow but not so much about whether there are an even number of stars in the universe, we assign a higher weight to the former proposition when measuring our inaccuracy.[4]

---

[1] I assume throughout that the set of possible worlds W is finite. This restriction may seem implausible, but (1) it's fairly standard, and (2) we can understand W to be the set of the most fine-grained possibilities an agent is concerned with. Since real-world agents can only track finitely many distinguished possibilities, I take it the finite case is of primary interest.

[2] We take the average square of the Euclidean distance instead of just the Euclidean distance to guarantee that all probability functions assign themselves lowest expected inaccuracy. I thank a referee for catching earlier sloppiness about this point.

[3] Named after Glenn Brier, who originally used it to measure the inaccuracy of weather forecasts. (See Brier 1950.)

[4] To allow for some propositions to count more than others, quadratic inaccuracy measures may give a score to the agent's credence in each proposition A in the algebra. With the Brier score, since we're counting each proposition equally and all agents under consideration are probabilistically coherent, we only have to consider the agent's credences in propositions of the form $\{w\}$ for w a world.

In addition to Leitgeb and Pettigrew (2010a,b), quadratic scoring rules have a number of defenders and admirers; in fact, it's safe to say that they are the clear front-runners in the debate over how best to measure the inaccuracy of probability functions.[5]

Nonetheless, I think that quadratic scoring rules serve as poor measures of inaccuracy because they naturally lead to an extremely unattractive updating procedure, as can be shown through application of a result in Leitgeb and Pettigrew (2010b). Other measures, such as the logarithmic scoring rule, lead to the more attractive and standard procedure of Jeffrey-Conditionalization. Therefore, quadratic rules should be rejected.[6]

2 Inaccuracy and Updating

2.1 Conditionalization

Leitgeb and Pettigrew have four different precisifications of ACCURACY. We'll be interested in how to update our *entire* credence function *across time*, so the version we want is:

ACCURACY (DIACHRONIC EXPECTED GLOBAL): Suppose an agent has learned evidence be-
   tween t and t' that imposes constraints C on her belief function $Cr_{t'}$ at t', or on the set
   E of worlds that are epistemically possible for her at t', or both. Then, *at time t'*, such an
   agent ought to have a belief function that satisfies constraints C and is minimal amongst
   belief functions thus constrained with respect to expected global inaccuracy by the lights
   of her belief function *at time t*, relative to a *legitimate* global inaccuracy measure [i.e.,
   relative to the Brier score for Leitgeb and Pettigrew], and over the set of worlds that are
   epistemically possible for her *at time t'* given the constraints C. (From (2010a, 207) with

---

[5]For a partial list, see: de Finetti (1974); Greaves and Wallace (2006); Joyce (1998, 2009); Leitgeb and Pettigrew (2010a,b); Savage (1971); Selten (1998). For good discussions of what makes quadratic scoring rules so attractive, see Selten (1998); Leitgeb and Pettigrew (2010a); Joyce (2009, §12).

[6]For our purposes below, I'll focus my attack on the Brier score, but the basic problem I'll highlight will apply to quadratic scoring rules in general.

minor changes)

Here's how ADEG leads to conditionalization: Suppose I start with credence function Cr and learn for certain some new information E. I'm now epistemically obligated to pick some probability function in set C of probability functions that assign 1 to E. But which should I pick? If I follow ADEG, I ought to pick the member of C that minimizes expected inaccuracy by the lights of Cr. So, if $I(\cdot,\cdot)$ is my favored measure of inaccuracy, the best new credence function to adopt upon learning E according to Cr is the credence function $b^* \in S$ such that:

$$EV_{Cr}(I,b,E) := \sum_{w \in E} Cr(\{w\})I(b,w)$$

is minimal.

It turns out that for nearly any measure of inaccuracy one might think is reasonable (including the quadratic scoring rules), the function that minimizes expected inaccuracy under these constraints is simply $Cr(\cdot|E)$. In other words, I minimize expected inaccuracy upon learning new information if I update by conditionalization.[7]

2.2 Updating with Uncertain Evidence

I'll assume with Leitgeb and Pettigrew that not all updating should be by conditionalization. Sometimes, I get perceptual input that doesn't raise my credence in any proposition fully to 1. For instance, I might be in a pitch black room with a piece of paper in front of me that I believe to degree .3 is red. A small bit of light comes in under the door which allows me a better but not perfect view of the paper. Even though I didn't learn any new salient proposition for sure, my credence in red changes to .7, and I'm obligated to update the rest of my beliefs. For ease of reference, let's call a situation where we don't update any of our partial beliefs non-inferentially to 0 or 1 an *Uncertain Evidential Situation* (UES).

---

[7]For demonstrations, see Greaves and Wallace (2006); Oddie (1997); Leitgeb and Pettigrew (2010b).

To set up the primary issue of the paper, we can consider only a special class of UES's, which I'll refer to as *Jeffrey-UES*'s. Let W be the set of possible worlds (with $|W| < \infty$), and let $\{E_i\}_{i \in I}$ be a partition of W with $0 < Cr(E_i) < 1$. Suppose we get some new uncertain evidence that forces us (non-inferentially) to adopt new credence q_i in E_i (where $0 < q_i < 1$ and $\sum q_i = 1$).

The classic answer to this updating problem is that our new credence $Cr_{t'}$ should be given by:[8]

JEFFREY-CONDITIONALIZATION: In Jeffrey-UES's, agents should update their credences through Jeffrey-Conditionalization. I.e., for any proposition A:

$$Cr_{t'}(A) = \sum q_i Cr_t(A|E_i)$$

Jeffrey-Conditionalization is the obvious way to extend standard conditionalization to cases with uncertain evidence, and it has long been more or less the only game in town when it comes to updating in Jeffrey-UES's.

It's surprising, then, that *quadratic scoring rules don't lead to Jeffrey-Conditionalization*! Leitgeb and Pettigrew (2010b) show that under the Brier score, ADEG requires a different updating procedure entirely, which I'll argue has absurd consequences.[9]

As before, we want to update by minimizing expected inaccuracy under constraints. This time, however, we don't restrict the class of worlds under consideration from W to E. Instead, we just consider the same class of worlds W, but we have new constraints imposed, given by the q_i's. That is, we choose the credence function with minimal expected inaccuracy that assigns credence q_i to E_i for all i.

Here's how this updating procedure works: We first take some element E_j of the partition

---

[8]For extensive discussion, see Jeffrey (1983).
[9]There will be different updating procedures for the different quadratic scoring rules, but each will have a similar form, and similar problems, to the one for the Brier score. To be clear, the objection isn't that quadratic scoring rules don't lead to Jeffrey-Conditionalization per se, but that they lead to extremely unattractive alternatives. It's only because of the general success of Jeffrey-Conditionalization that I'll hold it up as the procedure to beat below.

and then add (!) a constant d_j to $Cr_t(\{w\})$ for each world $w \in E_j$. In order to prevent negative credence, our new credence $Cr_{t'}(\{w\})$ is just $\max(Cr_t(\{w\}) + d_j, 0)$. By requiring that $Cr_{t'}(E_j) = q_j$, we end up with only one possible choice of d_j.[10] In other words, we end up with:[11]

LP-CONDITIONALIZATION: In Jeffrey-UES's, agents should update their credences through Leitgeb-Pettigrew-Conditionalization. That is, let d_i be the unique real number such that

$$\sum_{\{w \in E_i | Cr_t(\{w\}) + d_i > 0\}} Cr_t(\{w\}) + d_i = q_i.$$

Then the agent ought to have belief function $Cr_{t'}$ at t' such that for $w \in E_i$:

$$Cr_{t'}(\{w\}) = \begin{cases} Cr_t(\{w\}) + d_i & Cr_t(\{w\}) + d_i > 0 \\ 0 & \text{Otherwise} \end{cases}$$

It's worth emphasizing that choosing this updating procedure is inevitable for adherents of ADEG and the Brier score. Let's review why exactly: An agent starts with some credence function Cr_t. She then finds herself in a Jeffrey-UES, and her new evidence compels her to have credence q_i in the propositions E_i. She's still left with continuum-many credence functions to choose from that meet these constraints, but she wants the one she expects at t is the most accurate of the ones available. If she uses the Brier score, she'll pick the one mandated by LP-Conditionalization.[12] After discussing the problematic cases, I'll argue below that the problems

---

[10] For instance, in the special case where no worlds in E_j end up getting assigned credence 0, we have $d_j = \frac{q_j - Cr_t(E_j)}{|E_j|}$.

[11] From Leitgeb and Pettigrew (2010b, 254).

[12] The proof is rather long, so I omit it here and refer the interested reader to Leitgeb and Pettigrew (2010b). The basic reason is as follows: ADEG combined with the Brier score leads us to pick the Euclidean-closest credence function to the original that meets the constraints. It's then not hard to show that we minimize Euclidean distance by adding the same constant (so long as we can) to each world in a given element of the partition.

with LP-Conditionalization should lead to the rejection of the Brier score even as a measure of synchronic accuracy.

3 Problems

3.1 Initial Issues

Some apparent problems with LP-conditionalization are discussed and defended in Leitgeb and Pettigrew (2010b). These won't be my primary focus here, but it's worth going over them briefly:

The first is that the probability of some worlds can be lowered from a positive prior all the way to 0 in a UES. For instance, consider the following case:

	w_1	w_2	w_3
Cr_t	.3	.2	.5
$Cr_{t'}^{LP}$	.8	0	.2

$Cr_{t'}^{LP}$ is the result of raising the probability of $\{w_1\}$ to .8 and lowering that of $\{w_2, w_3\}$ to .2 under LP-conditionalization. Oddly, this results in assigning w_2 credence 0.[13]

A related issue is that LP-Conditionalization, unlike Jeffrey-Conditionalization, doesn't have standard conditionalization as a limiting case. In other words: Raising the probability of one element E of the partition to 1 and LP-Conditionalizing isn't the same as simply conditionalizing on E. To see this, consider the following case, in which the posterior of $\{w_1, w_2\}$ is raised to 1:

---

[13]It's worth working out this example to get a better sense of LP-Conditionalization. Since $\{w_1\}$ is a single-membered element of the partition, we'll just look at what happens to the elements of $E_2 := \{w_2, w_3\}$. We're given the constraint that $Cr_{t'}^{LP}(\{w_2\}) + Cr_{t'}^{LP}(\{w_3\}) = .2$. LP-conditionalization tells us to meet this constraint by finding some constant d_{E_2} such that $\max(Cr_t(\{w_2\}) + d_{E_2}, 0) + \max(Cr_t(\{w_3\}) + d_{E_2}, 0) = .2$. Now, since probabilities can't be negative, to meet the constraints we need $Cr_{t'}^{LP}(\{w_3\}) \leq .2$, so we have $d_{E_2} \leq -.3$. Since $.2 - .3 < 0$, we know that $Cr_{t'}^{LP}(\{w_2\}) = 0$ and $d_{E_2} = -.3$.

	w_1	w_2	w_3	w_4
Cr_t	.3	.2	.25	.25
$Cr_t(\cdot \mid \{w_1, w_2\})$	.6	.4	0	0
$Cr_{t'}^{LP}$	.55	.45	0	0

Unlike Leitgeb and Pettigrew (2010b), I do take both of these facts to be problematic, though I don't think they're necessarily fatal. They argue, perhaps rightly, that there's a stark difference between UES's and situations that call for full conditionalization. The latter involve ruling out worlds from the set of epistemic possibilities. The former involve the same set of epistemic possibilities, just new constraints on the attitudes toward those possibilities. Thus, credence 0 in a proposition, even when there are only finitely many worlds, can represent two very different doxastic attitudes toward it. One of those attitudes rules the proposition out and precludes it from every being assigned positive probability in the future, whereas the other maintains its epistemic possibility, but assigns it maximally low credence. Though these issues call for further discussion, I'll bracket them here.

3.2   Main Problem

Regardless of what one thinks of these results, a more important problem—and a fatal one, in my view—is the potentially dramatic effect LP-Conditionalization can have on the likelihood ratios between different propositions. Since LP-Conditionalization *adds* a constant to the prior credence in a world, important evidential relationships reflected in the prior can then be violated. For illustration, we turn to the following two cases.

Case 1

There's a car behind an opaque door, which you're almost sure is blue, but which you know might be red. You're almost certain of materialism, but you admit that there's some minute

8

possibility that ghosts exist. For ease of reference, we introduce the following abbreviations for propositions:

- B: The car is blue

- G: There are ghosts

To make the case precise, we stipulate you have the following prior:

	w_1 $B \wedge G$	w_2 $B \wedge \neg G$	w_3 $\neg B \wedge G$	w_4 $\neg B \wedge \neg G$
Cr_t	$\sim .000476$	.95	$\sim .000025$	.0495

Now the opaque door is opened, and the lighting is fairly good. You're quite surprised at your sensory input: Your new credence that the car is red is .99! Let's look at your posterior credence under the two updating procedures:

	w_1	w_2	w_3	w_4
$Cr_{t'}^{LP}$	0	.01	$\sim .470262$	$\sim .519738$
$Cr_{t'}^{J}$	$\sim .000005$	$\sim .009995$	$\sim .000495$	$\sim .989505$

Jeffrey-Conditionalization leads to no change in opinion about ghosts. Under LP-Conditionalization, however, seeing the car makes you about 47% sure there are ghosts. Note that the case was originally set up so that you thought the questions of what color the car is and whether there are ghosts were independent, but somehow, merely acquiring information about car color has drastically changed your opinion about materialism. Note that had you come to know the car was red and conditionalized on ¬B, you would have ended up with credence .0005 that there were ghosts. So, the difference between the near certainty of credence .99 and full knowledge that ¬B is the difference between credence .47 and credence .0005 that G. I think it's

clear that something's gone wrong with LP-Conditionalization here. Becoming more confident in one proposition shouldn't alone raise your credence in another if you initially take them to be independent.

Case 2

Consider the following propositions:

- U: Unemployment will rise before the next election.

- P: The president will be re-elected.

Suppose your prior and posterior credences after LP-conditionalization are given in the chart below:

	w_1	w_2	w_3	w_4
	U $\wedge$ P	U $\wedge$ ¬P	¬U $\wedge$ P	¬U $\wedge$ ¬P
Cr_t	.05	.15	.38	.42
$Cr_{t'}^{LP}$	.44	.54	0	.02

$Cr_{t'}$ is the result of becoming .98 sure that unemployment will rise from an initial credence of .2. In this example, we have $Cr_t(P) = .43$, while $Cr_t(P|U) = .25 < .43$. However, after becoming more confident that unemployment will rise, LP-Conditionalization leads to $Cr_{t'}(P) = .44 > .43$. So, despite the fact that you thought that it was less likely that the president would be re-elected given that unemployment rises, becoming nearly sure that unemployment will rise ends up *raising* your credence that the president will be re-elected.

The general problem in both cases is that LP-Conditionalization does not respect important evidential relationships reflected in the agent's prior. Furthermore, there's nothing particularly odd or unusual about the structure of either case. Therefore, LP-Conditionalizers will often

end up with unreasonable posteriors despite perfectly sensible priors. Below, we consider two potential escapes for defenders of the Brier score, which I'll argue don't work.

4   Rigidity to the Rescue?

For those who accept both ADEG and the Brier score, one way to avoid the unwelcome consequences above is to build in a requirement that certain structural relationships from the prior credence function be preserved. For instance, one might suggest something like the following as an additional constraint:

RIGIDITY: Suppose $\{E_1, \ldots, E_n\}$ is a partition of W, $0 \leq q_1, \ldots, q_n$ and $\sum q_i = 1$. If the agent acquires uncertain evidence that requires $Cr_{t'}(E_i) = q_i$ for all i, then we have that for all $A \subseteq W$, $Cr_{t'}(A|E_i) = Cr_t(A|E_i)$.

With this added requirement, quadratic scoring rules do lead to standard Jeffrey-Conditionalization and indeed do so trivially, since

$$Cr_{New}(A) = \sum Cr_{New}(E_i) \cdot Cr_{New}(A|E_i) = \sum Cr_{New}(E_i) \cdot Cr_{Old}(A|E_i)$$

However, I don't think this move is attractive. Though preservation of these particular conditional probabilities may generally be desirable, it should result from the choice of accuracy measure, not from any added constraint. Why? First, such a move looks like an ad hoc fix unless more motivation can be provided. Second, being accurate is more important than maintaining initial opinions about any conditional relationships. Structural requirements on a credence function should emerge from evidential and alethic requirements. Put differently: Our quest as epistemic agents is for the truth, which we pursue by means of obeying evidential requirements. If we expect to sacrifice some accuracy in order to maintain nice structural relationships between propositions, then we're expecting to go against our primary ends as epistemic agents.

If I'm faced with a choice between two credence functions, one of which preserves my prior probabilities conditional on elements of the partition, but the other of which I think is more accurate, I should prefer the latter. It's then unreasonable to add the constraint that the new probability function not be the one with the highest expected accuracy that's compatible with the new information simpliciter, but instead be the one with the highest expected accuracy that's *both* compatible with the new information *and* that preserves some other features of the prior.

5   The Brier Score and Synchronic Accuracy

The argument against the Brier score as a measure of inaccuracy comes from its failings in recommending an updating procedure in Jeffrey-UES's. Now, a number of authors endorse ADEG or an equivalent norm and use it to argue for constraints on updating procedures.[14] The only options they have are rejecting the Brier score, accepting LP-Conditionalization, or endorsing different synchronic and diachronic accuracy measures. The last option appears ad hoc at best. It amounts to claiming that the kind of accuracy I care about my future beliefs having is different from the kind of accuracy I care about my current beliefs having. Though I won't rule this possibility out, it would require a lot of further motivation. Since we've seen that LP-conditionalization is unattractive, supporters of ADEG ought to take the first option of rejecting the Brier score.

   Many defenders of the Brier score, however, have been silent about any accuracy norms for updating beliefs and have considered only its synchronic features. Indeed, there is something prima facie odd about ADEG, since it requires evaluation of potential posterior credence functions through the use of a prior credence function that we know is out-dated. That is, it requires us to use the prior Cr_t that is—ex hypothesi—an inappropriate response to the total evidence at t'. To make the case that the considerations above should lead everyone to reject the Brier score,

---

[14]Aside from Leitgeb and Pettigrew (2010a,b), see Greaves and Wallace (2006), Kierland and Monton (2005), and Oddie (1997).

I'll try to strengthen the argument that a norm like ADEG is at least descriptively adequate.

Let's go back to Case 1. Instead of jumping into the diachronic case, this time we have three agents with the following *prior* credence functions.

	w_1 $B \wedge G$	w_2 $B \wedge \neg G$	w_3 $\neg B \wedge G$	w_4 $\neg B \wedge \neg G$
Alice	$\sim .000476$	$.95$	$\sim .000025$	$.0495$
Leopold	0	$.01$	$\sim .470262$	$\sim .519738$
Jeff	$\sim .000005$	$\sim .009995$	$\sim .000495$	$\sim .989505$

That is, Alice's prior is Cr_t in the original case, Leopold's is $Cr_{t'}^{LP}$, and Jeff's $Cr_{t'}^{J}$. Suppose Alice likes the Brier score as a synchronic measure of accuracy, but she doesn't endorse ADEG and instead updates by Jeffrey-Conditionalization. She's now in an odd situation. She *expects* Leopold to be more accurate than Jeff overall. However, were she to have the same constraints placed on her credence function that we find in Case 1, she'd adopt Jeff's current credence function as her own.

Here's one way to think about this. Using ACCURACY and the Brier score, Alice can create a preference ranking for all possible credence functions in terms of how well she expects them to do.[15] She of course expects her own credence function to have a higher level of accuracy than any other, so it's first on her preference list. Still, she doesn't think all other credence functions are on a par. In particular, she thinks that Jeff's is worse than Leopold's. However, she knows now (or at least can be shown) that were nature to place constraints on her credence function that required her to have credence .99 in ¬B, she would adopt Jeff's credence over Leopold's. We can dramatize the case as follows: Suppose an evil scientist told Alice that he was going to force her to take one of two pills. Pill 1 would change her prior to Leopold's, and Pill 2 would

---

[15] I assume that all participants in the discussion accept something like the simple ACCURACY norm.

change it to Jeff's. Her memory of this event would be wiped clean either way, and she'll get no new evidence, uncertain or otherwise, that should affect her credence in either B or G. In this situation, she would opt for Pill 1. If, however, instead of an evil scientist, she has a sensory experience that arationally raises her credence in ¬B to .99, she'll opt for Jeff's credence.

I think that this oddity should worry even someone who denies that an agent should literally use her prior probability function to justify her posterior. Such a philosopher might object to ADEG or any diachronic norm as follows: At t, you have total evidence E, which you should use to generate a rational credence function Cr_t. At t', you have evidence E'. Throw out your earlier Cr_t. It's outdated, and you now shouldn't care about what you thought in the past. Instead, just look at your total evidence at t' and figure out how best to respond to it. It can turn out that Cr_t and $Cr_{t'}$—if they're good—will exhibit certain formal relationships. It might even look as if you've updated by some form of conditionalization, and it might be useful to do so to save computational time. But that's just a spandrel from a normative perspective. Your updated credence function is justified by the evidence you have at t' not by what you thought at t.

Even if we concede that this is the right normative picture, we can maintain an 'as if' picture of updating in accord with ADEG and a good inaccuracy measure. Here's why: Suppose Alice doesn't really update, but instead just follows some rational policy for evaluating evidence, which we'll call P. P is a function from potential sets of total evidence to credence functions. That is, whenever you get some body of total evidence, P tells you what your credence function should be without regard for what you thought in the past. The reply then goes as follows: If P is actually a good policy, then it will generally tend to give you accurate credence functions. Of course, sometimes it will be off, and sometimes you'll be in an uncooperative world that feeds you misleading evidence all the time, but in typical worlds with typical evidence, a good policy should be fairly successful. Now suppose the Brier score is the right way to measure inaccuracy. So, if P is generally good, it will let Alice do a good job at coming up with an accuracy ranking of credence functions in terms of their expected Brier scores. However, if P makes it look as if

Alice updates by a rule seriously different from LP-Conditionalization and the Brier score really is the right way to measure inaccuracy, then Alice will either end up with bad posterior credence functions or her prior rankings of expected accuracy will be systematically and predictably off. That is, either P could do better at figuring out a good posterior or P could generate prior credence functions that do better at evaluating other credence functions. Since P could do better either way, it must not be an ideal policy. Of course, this argument is rough-and-ready and falls short of showing that P is dominated, but I think it nonetheless puts an advocate of the Brier score in an uncomfortable position regardless of her views on updating.

Therefore, a norm like ADEG should be at least descriptively accurate at least in normal circumstances where the agent's prior is reasonable. Consequently, we should reject the Brier score as a measure of synchronic accuracy, so long as better alternatives are available. It turns out other, less popular measures of accuracy do result in a superior updating procedure, so we should opt for one of them. In the Appendix, we show that the logarithmic rule, in particular, leads to Jeffrey-Conditionalization.

6   Conclusion

The argument of the paper went as follows. First, we should strive to minimize our expected inaccuracy under some reasonable measure, by ACCURACY. If we use a quadratic scoring rule and follow ADEG, we end up with LP-Conditionalization. We should keep ADEG (at least usually), but we shouldn't update by LP-Conditionalization. Therefore, quadratic scoring rules aren't reasonable inaccuracy measures.

Appendix

We here sketch a proof of the earlier claim that ADEG combined with the following inaccuracy measure leads to Jeffrey-Conditionalization:

LOGARITHMIC RULE (LR): $I(Cr, w) = -\ln(Cr(w))$.

First, it will be useful to have the following rediscription of Jeffrey-Conditionalization. Suppose Alice's prior is Cr_t and let $\{E_i\}_{i \in I}$ partition the set of possible worlds W. News comes in, and Alice must update her credence function so that she now has credence q_i in each element E_i of the partition. If she Jeffrey-Conditionalizes, she updates her credence to $Cr_{t'}^J$ as follows: for each element E_i she finds some constant c_i such that for all $w \in E_i$, $Cr_{t'}^J(\{w\}) = c_i \cdot Cr_t(\{w\})$. Since $\sum_{w \in E_i} Cr_{t'}^J(\{w\}) = q_i$, exactly one constant will work. Therefore, it will suffice to show that LR together with ADEG require multiplying the prior credence in each member of a given element of the partition by the same constant.

Suppose now that Alice follows ADEG and measures inaccuracy with LR. For the sake of simplicity, we first consider the case in which $W = \{w_1, w_2, w_3, w_4\}$ with $E_1 = \{w_1, w_2\}$ and $E_2 = \{w_3, w_4\}$. For readability, we set $Cr_t(\{w_i\}) = \alpha_i$. To follow ADEG, we're now looking for the quadruple $\langle \beta_1^*, \beta_2^*, \beta_3^*, \beta_4^* \rangle$ that minimizes:

$$-\sum_{i=1}^{4} \alpha_i \ln \beta_i \tag{1}$$

where $\beta_1 + \beta_2 = q_1$, $\beta_3 + \beta_4 = q_2$, and $0 \leq \beta_i$ for all i. Since E_1 and E_2 are disjoint, minimizing (1) under these side-constraints is just minimizing both of the following:

$$f(\beta_1) = -\alpha_1 \ln \beta_1 - \alpha_2 \ln(q_1 - \beta_1)$$
$$g(\beta_3) = -\alpha_3 \ln \beta_3 - \alpha_4 \ln(q_2 - \beta_3)$$

To minimize f, we let:

$$f'(\beta_1) = \frac{\alpha_2}{q_1 - \beta_1} - \frac{\alpha_1}{\beta_1} = 0 \tag{2}$$

Note that $\beta_1 = \alpha_1 \cdot c$ and $\beta_2 = \alpha_2 \cdot c^*$ for some constants c and c^*. Substiting these identities in (2) gets us: $\frac{1}{c} - \frac{1}{c^*} = 0$. So, $c = c^*$. To minimize g, we follow the same procedure mutatis mutandis. We then have that $\langle \beta_1^*, \beta_2^*, \beta_3^*, \beta_4^* \rangle = \langle c\alpha_1, c\alpha_2, c'\alpha_3, c'\alpha_4 \rangle$ for constants c and c'. In other words, each element of E_i gets multiplied by the same constant c_i as desired. Thus, in this special case in which each $|E_i| \leq 2$ for all i, we end up with Jeffrey-Conditionalization.

For the more general case, suppose $|E_1| = n$. We then wish to minimize $-\sum_{i=1}^{n} \alpha_i \ln \beta_i$ under the side-constraints. To do so, pick two distinct β_i's to treat as variable and treat the rest as constant. Without loss of generality, we choose β_1 and β_2, and set $\sum_{i=3}^{n} \beta_i = K \leq q_1$ for some constant K. Now, we just need to minimize: $h(\beta_1) = -\alpha_1 \ln \beta_1 - \alpha_2 \ln(q_1 - K - \beta_1)$. Running essentially the same argument as above, we find that $\beta_1^* = c\alpha_1$ and $\beta_2^* = c\alpha_2$ for some constant c. Since the choice of K and the two β_i's was arbitrary, we again end up with Jeffrey-Conditionalization as desired.

References

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review 78*, 1–3.

de Finetti, B. (1974). *Theory of Probability*, Volume 1. John Wiley and Sons.

Greaves, H. and D. Wallace (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind 115*(632), 607–632.

Jeffrey, R. C. (1983). *The Logic of Decision* (2nd ed.). University of Chicago Press.

Joyce, J. M. (1998, December). A nonpragmatic vindication of probabilism. *Philosophy of Science 65*, 575–603.

Joyce, J. M. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber and C. Schmidt-Petri (Eds.), *Degrees of Belief*, Volume 342, pp. 263–297. Springer.

Kierland, B. and B. Monton (2005). Minimizing inaccuracy for self-locating beliefs. *Philosophy and Phenomenological Research 70*(2), 384–395.

Leitgeb, H. and R. Pettigrew (2010a). An objective justification of bayesianism I: Measuring inaccuracy. *Philosophy of Science 77*, 201–235.

Leitgeb, H. and R. Pettigrew (2010b). An objective justification of bayesianism II: The consequences of minimizing inaccuracy. *Philosophy of Science 77*, 236–272.

Oddie, G. (1997). Conditionalization, cogency, and cognitive value. *British Journal for the Philosophy of Science 48*(4), 533–541.

Savage, L. J. (1971). Elicitation of personal probabilities. *Journal of the American Statistical Association 66*, 783–801.

Selten, R. (1998). Axiomatic characterization of the quadratic scoring rule. *Experimental Economics 1*, 43–62.