



Pettigrew, R. (2018). Précis of Accuracy and the Laws of Credence. *Philosophy and Phenomenological Research*, 96(3), 749-754.  
<https://doi.org/10.1111/phpr.12501>

Peer reviewed version

Link to published version (if available):  
[10.1111/phpr.12501](https://doi.org/10.1111/phpr.12501)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at <https://onlinelibrary.wiley.com/doi/10.1111/phpr.12501>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Précis and Replies to Briggs, Joyce, and Kotzen for Book Symposium on *Accuracy and the Laws of Credence* in *Philosophy and Phenomenological Research*

Richard Pettigrew  
Richard.Pettigrew@bris.ac.uk

April 19, 2018

There are laws of rationality that govern our degrees of belief, our levels of confidence, our credences. They describe features that our credences must have if they are to avoid being irrational. For instance, if I'm 80% confident that Hillary Clinton will win the US presidential election, I should be at least 80% confident that a Democrat will win. As November approaches, and my credence that Hillary will win increases, my credence that she won't win should decrease, and by the same amount. If I learn that the objective chance that Hillary will win is between 70% and 90%, I shouldn't be 98% confident that she will win. If I learn only that Hillary has a dog and a cat, one called Blip, the other Blop, I should be just as confident that Blip is feline and Blop canine as I should be of the reverse. If I'm very confident that Hillary will win the presidency, and I'm very confident that the Democrats will take the House, then I shouldn't plan to become sceptical that the Democrats will take the House should I learn that Hillary has won the presidency.

Each of these claims is a particular consequence of a general law of credence. The first two are consequences of *Probabilism*, which says that an agent's credences should obey the axioms of the probability calculus. The next follows from the *Principal Principle*, which, on one formulation, says that our credence in a proposition should be our expectation of its objective chance, and thus must lie in the range spanned by the various possible objective chances. After that, we have a consequence of the *Principle of Indifference*, which says that, in the absence of evidence, we should be equally confident in every possibility. And the final claim follows from *Conditionalization*, which tells us how we should plan to change our credences in response to new evidence: we should plan to set our new credence in a proposition to be the same as our old credence in that proposition conditional on the piece of evidence we learn. Probabilism, the Principal Principle, and Conditionalization are the central tenets of Bayesian epistemology; add the Principle of Indifference and you have objective Bayesianism. Together, they provide a detailed account of the constraints that rationality places on our theoretical reasoning under uncertainty.

*Accuracy and the Laws of Credence* can be read in two ways. On the first, it is an attempt to justify Probabilism, the Principal Principle, Conditionalization, and with more hesitation the Principle of Indifference. Read like this, it begins by assuming a plausible account of epistemic value, and proceeds to show how each of these laws of credence follows from that account. The account it assumes is a credal version of *veritism*. This says that the purely epistemic value or epistemic utility of a credence in a proposition is its *accuracy*, where a cre-

dence in a true proposition increases in accuracy as the credence increases, while a credence in a false proposition increases in accuracy as the credence decreases.

On the second way of reading the book, it doesn't *assume* the credal version of veritism. It is a by-their-fruits-shall-ye-know-them attempt to *establish* that account of epistemic value. One pressing objection to veritism is that it cannot account for certain core norms of beliefs; or, at the very least, it cannot provide a case for those norms that is as compelling as the case provided by its main rival, *evidentialism*, which says that an agent's credence in a proposition should match the support given to that proposition by the agent's total evidence. The tenets of objective Bayesianism — Probabilism, the Principal Principle, the Principle of Indifference, and Conditionalization — are plausibly such norms. Evidentialism can give a plausible *prima facie* argument for each: as the evidential support for a proposition increases, the evidential support for its negation decreases, and by the same amount; the evidential support for a proposition is always at least as great as the evidential support for any proposition that entails it; and so on. But it is not at all clear, at first sight, that veritism can do likewise. Thus, by showing how these four laws of credence follow from veritism, we answer a significant objection to the position.

Let's see, then, how the arguments for these laws of credence work. They all share the same structure. We might call arguments with this structure *epistemic utility arguments* for laws of credence. Each epistemic utility argument has three premises.

The first premise is the same in all four of the main arguments we provide in the book. It is an account of epistemic value or utility — in particular, it is a precise version of credal veritism. Recall: credal veritism says that the epistemic value or utility of an agent's credal state — how good it is from a purely epistemic point of view — is its accuracy. We model an agent's credal state at a time by her credence function at that time. This is a mathematical function that takes each proposition about which she has an opinion and assigns to it her credence in that proposition at that time. Our account of epistemic value is perfectionist. A perfectionist account of the value of a certain sort of item says that its value is given by how close it comes to the ideal or perfect item of that sort. So, for instance, a perfectionist account of societal value determines the value of a given society by measuring how far it lies from the perfect, utopian society. In our case, perfectionism says that the epistemic value of a credence function at a possible world is its proximity to the credence function that is epistemically ideal or perfect at that world; the credence function that it would be best, epistemically speaking, to have at that world. Our account is veritistic because we add to this perfectionism the claim that the perfect or ideal credence function at a world is the one that gets things exactly right at that world — that is, it assigns maximal credence to truths and minimal credence to falsehoods. So epistemic value is proximity to perfection; and perfection is getting things exactly right. To complete our account of epistemic value, we need only now supply the correct measure of proximity to the ideal — that is, we need only say how to measure how far one credence function lies from another. After the introductory material of chapters 1 and 2, chapters 3 and 4 are devoted to seeking out this measure: chapter 3 concerns existing arguments for various classes of measure, and raises problems for them; chapter 4 provides a new argument for the most popular measure, the so-called *squared Euclidean distance* measure. As with the existing accounts surveyed in chapter 3, our approach is axiomatic. We lay down properties that we would like such a measure to have; and we show that only one measure — squared Euclidean distance — has all of these properties. With this, we complete our mathematically precise version of credal veritism: the epistemic value of a credence function at a world is its proximity to the ideal credence

function for that world; the ideal credence function for a world assigns maximal credence to truths and minimal credence to falsehoods; and squared Euclidean distance supplies the measure of how far one credence function lies from another. We call this version of veritism *Brier Alethic Accuracy*. This gives the first premise shared by all epistemic utility arguments in *Accuracy and the Laws of Credence* (though, as I note, some of these arguments would succeed with a weaker premise).

The second premise differs from one epistemic utility argument to another. But each is a very general putative law of rational choice. The idea is this: Normative decision theory is typically used to evaluate the rationality of an agent's actions — should she take an umbrella when she leaves the house, uncertain as she is whether it will rain or not; should she accept or decline such-and-such a bet and so-and-so a price? But the laws of decision theory and rational choice can in fact be applied to evaluate other options between which an agent faces a choice, provided at least that those options can be assigned a value at each possible world. Thus, we might use these principles to evaluate which action an agent chooses to perform; but we might also use them to evaluate the scientific theories that she accepts or the beliefs she adopts or the credences that she assigns, providing that we have an account, for each possible world, of the value of accepting a particular scientific theory, the value of adopting a given belief, or the value of assigning a certain credence at that world.

Thus, there are perfectly general principles of rational choice that say how an agent should choose between options, whatever those options are. Here is one, which we call Dominance: if one option is guaranteed to be better than another — that is, if it has higher value in every single possible world — then choosing the latter option is irrational. For instance, if I prefer carrying an umbrella to not carrying one when it is wet (for reasons of comfort), and if I prefer carrying an umbrella to not carrying one when it is dry (for reasons of sartorial elegance), then Dominance rules it irrational to not carry an umbrella. After all, however the world turns out, taking an umbrella will be better for me. Notice that this is a completely general principle of rational choice. It applies to choices between actions, as we have just seen; but it applies equally to choices between other sorts of options, such as scientific theories, beliefs, and credence functions. When applied to the choice between credence functions, the relevant notion of value is epistemic value. In that case, it says the following: if one credence function is guaranteed to be better, epistemically speaking, than another — that is, if it has higher epistemic value in every single possible world — then adopting the latter credence function is irrational. Now, it turns out that, when combined with the account of epistemic value given above — that is, the precise veritistic account of epistemic value codified in *Brier Alethic Accuracy* — the principle entails Probabilism. The reason is that, for every credence function that violates the probability axioms, there is a credence function that satisfies them that is closer to the ideal credence function at every single possible world (when we use squared Euclidean distance to measure how far one credence function lies from another). Thus, for instance, if my credence that Hillary will win the presidency is greater than my credence that a Democrat will win, there is some other pair of credences in these propositions that is more accurate than my pair of credences at every possible world. This mathematical fact was first noted by Bruno de Finetti (1974) and then generalized by James M. Joyce (1998), who used it to give one of the first epistemic utility arguments for Probabilism. As well as setting up the framework of epistemic utility arguments, Part I of *Accuracy and the Laws of Credence* is devoted to expounding the best possible version of that argument. Chapter 2 tweaks Dominance to avoid various problems and provide a weaker and more plausible second premise. As we have seen, chapters 3 and 4

build to an argument for Brier Alethic Accuracy, which constitutes the first premise of the argument. Chapters 5 and 6 answer two objections to this epistemic utility argument for Probabilism. And Appendix I explains the mathematical theorem that constitutes the third premise of this epistemic utility argument — it shows that Probabilism does indeed follow from the first two premises.

Each of the remaining parts of the book is devoted to an epistemic utility argument for a different law of credence: Part II treats the Principal Principle and its variants; Part III concerns the Principle of Indifference and other laws concerning credences in the absence of evidence; and Part IV concerns Conditionalization and related laws that purport to govern how we update our credences in response to evidence. As mentioned above, in each case, the first premise of the argument for the law of credence in question is the same — it is Brier Alethic Accuracy. But the second premise is different — though always a putative law of rational choice. As we saw above, in the case of Probabilism, the second premise is a variant of Dominance. In the case of the Principal Principle, it is a variant of a law of rational choice that we might call Chance Dominance. Dominance says that an option is irrational if there is another that is guaranteed to have higher value than it. Chance Dominance says that an option is irrational if there is another that is guaranteed to have higher objective expected value than it. That is, if every possible objective chance function agrees that one option has higher expected value than another, then Chance Dominance declares the latter irrational. Thus, for instance, suppose you know that the coin I hold is biased at least 60% towards heads. I offer you a bet that pays £2 if the coin lands heads on the next toss and £0 if it lands tails. My price is £1. Should you buy the bet? Chance Dominance demands that you do. You don't know the objective chance of the coin landing heads. But you know that, from the point of view of any of the objective chances that you consider possible, taking the bet has greater expected utility than refusing it. So you should take it. In the case of credence functions, Chance Dominance says the following: if each possible chance function expects one credence function to be more accurate than it expects another to be, the latter is irrational. We formulate the Principal Principle in chapter 8 and consider an alternative epistemic utility argument in its favour in chapter 9. But, in chapter 10, we show that the Principal Principle follows from Chance Dominance combined with Brier Alethic Accuracy. That is, we show that, for any credence function that violates the Principal Principle, there is an alternative credence function that every possible chance function expects to be more accurate. Thus, for instance, if I know that the objective chance of Hillary winning is between 70% and 90%, yet I am 98% confident that she will win, there is an alternative credence I might assign that every possible objective chance function will expect to be more accurate than it expects my credence to be. Thus, the Principal Principle, like Probabilism, follows from veritism.

Next, Part III. The laws of rational choice that provide the second premises of the epistemic utility arguments that we give in this part of the book are *risk sensitive* norms. An agent is *risk averse* if, when he evaluates a particular option, he pays greater attention to possible worlds at which that option has its *lowest* value, and accords those worlds greater weight in his decision. An agent is *risk seeking* if she does the opposite: she pays greater attention to possible worlds at which the option has *greatest* value, and accords those worlds greater weight in her deliberation. Thus, a maximally risk averse agent will pay attention only to worst-case scenarios and will choose an option whose worst-case scenario is best — if offered a bet that gains him £1 billion if the sun rises tomorrow and loses him £1 if it does not, the maximally risk averse agent won't accept it, because the worst-case scenario,

in which he loses £1, is worse than the worst-case scenario of the status quo, in which he loses nothing. This putative law of rational choice is called Maximin — an agent should order options by their lowest possible value, and choose the one whose lowest possible value is greatest. We begin, in chapter 12, by noting that, if we take this putative law as the second premise in an epistemic utility argument, we can derive the Principle of Indifference as our conclusion. Thus, for instance, if I am more confident that Blip is a cat and Blop a dog than I am that Blip is a dog and Blop a cat, there is an alternative pair of credences whose worst-case scenario is better than mine; indeed, assigning equal credences to each proposition has a greater lowest value than my credences do; indeed, assigning equal credences has greater lower value than *any* other assignment. The Principle of Indifference is perhaps the norm of belief that we would most expect to be justified by evidential considerations, rather than by veritistic ones. The argument here shows that, in fact, it is a natural consequence of veritism when that account of epistemic value is combined with an extreme sort of risk aversion. In chapter 13, we explore the consequences of other risk-sensitive principles of rational choice, such as Maximin Regret — which tells you not to choose the option that maximizes your minimum utility, but to choose the one that minimizes your maximum regret — and the Hurwicz criterion — which allows you to effect a compromise between maximal risk aversion and maximal risk seeking.

In Part IV, we consider rules for updating credences upon receipt of new evidence. Our aim is to establish that an agent should plan to update her credences by Conditionalization. This is a synchronic law — it says how an agent's credences at a time should cohere with the plans she makes *at that same time* for how to update. Thus, it is distinct from the standard diachronic version of Conditionalization, which says that an agent should in fact update in this way after new evidence comes in, regardless of whether doing so was ever her plan. In chapter 15, I consider the diachronic version of Conditionalization and find it wanting. To give an epistemic utility argument for the synchronic version of Conditionalization, we need to extend our veritistic account of epistemic value so that we can measure the accuracy not only of a credence function, but also of an updating plan. It turns out that this is straightforward: the accuracy of an updating plan at a world is just the accuracy of whichever credence function that plan would have you adopt were you to learn whatever piece of evidence you would learn at that world. With this in hand, in chapter 14, we give three epistemic utility arguments for the synchronic version of Conditionalization. The first is due to Hilary Greaves and David Wallace, who note that, amongst all possible updating plans you might adopt, Conditionalization is the one with the highest expected accuracy from the point of view of your current credence function (Greaves & Wallace, 2006). The second argument turns on the following fact: if you plan to update by some rule other than Conditionalization, there is an alternative credence function you might have currently that each of your planned future credence functions expects to be more accurate than it expects your actual current credence function to be. The third is due to joint work with Rachael Briggs: it notes that, if you plan to update other than by Conditionalization, then there is an alternative current credence function and an alternative updating plan whose combined accuracy is guaranteed to be greater than the combined accuracy of your actual current credence function and your actual updating plan (Briggs & Pettigrew, ms). Thus, for instance, if I am very confident that Hillary will win the presidency and that the Democrats will take the House, yet plan to become sceptical of the latter should I learn the former, three problems will arise: there is an alternative plan I could have made that I will now expect to be more accurate; my planned future credences will unanimously expect some alternative credences to be more accurate than my actual cur-

rent credences; and there will be alternative current credences and planned future credences whose combined accuracy will definitely outstrip the combined accuracy of my actual current credences and planned future credences. Thus, in three different ways, we can see that Conditionalization follows from veritism.

Bayesian epistemology, then, our best account of rational reasoning under uncertainty, can be justified by appeal to veritism, an attractive account of epistemic value. Satisfying the central tenets of Bayesian epistemology — Probabilism, the Principal Principle, the Principle of Indifference, and Conditionalization — promotes the veritistic goal of accuracy. *Accuracy and the Laws of Credence* concludes, in chapter 16, by looking to the future to speculate on what else we might glean from this powerful, mathematically precise veritistic framework.

## 1 Reply to Joyce

Jim Joyce and I agree on many things: we agree on the conclusions of two of the main accuracy-based arguments for epistemic norms for credences; and we agree on two of the three main ingredients in those arguments. But we disagree on the third main ingredient, as well as the method by which the ingredients are combined to produce their conclusions. For me, an accuracy argument for a given epistemic norm has three ingredients: a specification of the legitimate measures of inaccuracy, a decision-theoretic principle, and a mathematical theorem. The mathematical theorem shows that, if we apply the decision-theoretic principle when the options are different sets of doxastic attitudes towards a given set of propositions and the utility function is one of the legitimate measures of accuracy for that sort of doxastic attitude, then the epistemic norm follows. Joyce's favoured versions of a given accuracy argument also has three ingredients: he agrees that it includes a specification of the legitimate inaccuracy measures and he agrees that it includes a mathematical theorem; but instead of a decision-theoretic principle, he takes it to include a principle concerning the notion of evidential support. This principle has two components. The first component connects evidential support and rationality; it says that, if  $c^*$  is better supported than  $c$  by an agent's evidence, it is irrational for that agent to have  $c$ . The second component connects evidential support and accuracy; it is specific to the epistemic norm to be justified. For instance,  $\text{SUPPORT}_{\text{Dom}}$  is the second component of the evidential support principle that appears in Joyce's favoured accuracy argument for Probabilism, whereas  $\text{SUPPORT}_{\text{Ch}}$  appears in his favoured argument for the Principal Principle.

In his excellent remarks, Joyce raises two worries for my accuracy argument for Probabilism, and one worry for my accuracy argument for the Principal Principle. Let me take these in order.

The first two worries target my response to concerns about the sort of dominance reasoning used in Joyce's original accuracy argument for Probabilism (Joyce, 1998). The first concern was raised by Aaron Bronfman (ms.); the second is an extension of a concern originally raised by Alan Hájek (2008). I take these concerns to show that Joyce's dominance reasoning needs to be amended; Joyce maintains that they do not. In response to Bronfman's objection, I strengthened the characterisation of legitimate inaccuracy measures — appealing to Symmetry, I narrowed the field to just one, namely, the Brier score. Joyce rejects the strengthened version, but holds that it's unnecessary for the success of the argument. In response to Hájek's objection, I weakened the dominance norm to which the argument appeals. Joyce accepts the weakened norm since he accepts the original stronger version, but

he thinks the weakening is unnecessary.

Let's consider the Bronfman objection first. According to the Bronfman objection, in order to establish that a particular credence function  $c$  is irrational, it is not sufficient to show that, for any legitimate inaccuracy measure  $\mathfrak{I}$ , there is a credence function  $c_{\mathfrak{I}}$  that  $\mathfrak{I}$ -dominates  $c$ ; rather, you must show that there is a credence function  $c^*$  that  $\mathfrak{I}$ -dominates  $c$  relative to any legitimate inaccuracy measure  $\mathfrak{I}$ . And this Joyce's argument does not do. According to Joyce, the Bronfman objection is plausible only if you accept Dominators, which says that an agent with a dominated credence function is irrational and should rectify her irrationality by moving to one of the dominating credence functions that is not itself dominated. Joyce claims that Dominators is false, and so the Bronfman objection fails and my solution to it is unnecessary. Yet I agree that Dominators is false. Whether or not it was the basis of Bronfman's original version of the objection, it is not the basis of the version I present in *Accuracy and the Laws of Credence* (Pettigrew, 2016a, Chapter 5). Rather, my version is based on a disjunctive syllogism. I start by asking what we mean when we say that the legitimate accuracy measures are precisely the additive and continuous strictly proper ones. I claim that we can answer this in one of three ways, which I label *epistemicism*, *supervaluationism*, and *subjectivism*, adopting terminology from the literature on the semantics of vague predicates. I claim that, whichever we choose, the argument fails. Therefore, if we take the legitimate inaccuracy measures to be the additive and continuous strictly proper ones, the argument fails.

Epistemicism says that there is, in fact, just one true inaccuracy measure, but we cannot know what it is; we can know only that it is additive, continuous, and strictly proper. The problem here is that, once we build the agent's uncertainty about the true inaccuracy measure into the decision problem, there are non-probabilistic credence functions that are not dominated, and so the accuracy argument for Probabilism fails. Consider, for instance, the example of Phil (Pettigrew, 2016a, 69-71):  $c_{\text{Phil}}(X) = 0.9$  and  $c_{\text{Phil}}(\bar{X}) = 0.2$ . What's more, he is unsure whether the additive logarithmic score  $\mathfrak{L}\mathfrak{A}$  or the additive spherical score  $\mathfrak{S}\mathfrak{A}$  is the one true inaccuracy measure. The set of credence functions that  $\mathfrak{L}\mathfrak{A}$ -dominate  $c_{\text{Phil}}$  is disjoint from the set that  $\mathfrak{S}\mathfrak{A}$ -dominates it (Pettigrew, 2016a, Figure 5.1). That is, no credence function both  $\mathfrak{S}\mathfrak{A}$ -dominates and  $\mathfrak{L}\mathfrak{A}$ -dominates  $c_{\text{Phil}}$ . Now, for Phil, there are four epistemically possible worlds:  $X$  might be true or false, and the true inaccuracy measure might be  $\mathfrak{S}\mathfrak{A}$  or  $\mathfrak{L}\mathfrak{A}$ . So the worlds are  $X$  &  $\mathfrak{S}\mathfrak{A}$ ,  $\bar{X}$  &  $\mathfrak{S}\mathfrak{A}$ ,  $X$  &  $\mathfrak{L}\mathfrak{A}$ , and  $\bar{X}$  &  $\mathfrak{L}\mathfrak{A}$ . Now, since no credence function both  $\mathfrak{S}\mathfrak{A}$ -dominates and  $\mathfrak{L}\mathfrak{A}$ -dominates  $c_{\text{Phil}}$ , there is no credence function that is more accurate than  $c_{\text{Phil}}$  at all four of these worlds. Thus,  $c_{\text{Phil}}$  is not dominated. So, if epistemicism about inaccuracy measures is correct, and the set of legitimate inaccuracy measures includes  $\mathfrak{L}\mathfrak{A}$  and  $\mathfrak{S}\mathfrak{A}$ , the accuracy argument fails — many incoherent credence functions are not in fact dominated. Phil's is one of them.

Next, supervaluationism. This says that the concept of inaccuracy is not sufficiently rich to determine a unique numerical measure. Rather, what is determinate is the ordering  $\preceq$  of pairs of credence functions and worlds by their inaccuracy:  $(c, w) \preceq (c', w')$  iff  $c$  is at least as inaccurate at  $w$  as  $c'$  is at  $w'$ . The legitimate inaccuracy measures are then those such that  $(c, w) \preceq (c', w')$  iff, for any legitimate inaccuracy measure  $\mathfrak{I}$ ,  $\mathfrak{I}(c, w) \leq \mathfrak{I}(c', w')$ . But if that's right, consider again Phil's credence function  $c_{\text{Phil}}$  from above. There is no  $c^*$  that  $\mathfrak{I}$ -dominates  $c_{\text{Phil}}$  for each additive and continuous strictly proper scoring rule. So, if those are the legitimate measures of inaccuracy, then there is no  $c^*$  such that  $(c^*, w) \prec (c_{\text{Phil}}, w)$  for all worlds  $w$ . That is, in the sense that matters,  $c_{\text{Phil}}$  isn't dominated, for in this context the dominance principle says that  $c$  is dominated iff there is  $c^*$  such that  $(c^*, w) \preceq (c, w)$



for all  $w$ , and  $c$  is irrational if it is dominated. Thus, if supervenience about inaccuracy measures is true, and if the legitimate inaccuracy measures are the additive and continuous strictly proper ones, then the accuracy argument for Probabilism fails again.

Finally, subjectivism. On this view, the legitimate inaccuracy measures are the permissible ones. Each agent must pick one and use that to measure inaccuracy; but which one she chooses is up to her. Granted this, the accuracy argument succeeds. But it is too much to grant. The demand that agents pick just one inaccuracy measure to represent the way in which they value the accuracy of their credences is too strong. If any additive and continuous strictly proper inaccuracy measure is permissible, surely it is permissible to value accuracy in a way that doesn't tell between them.

This, then, is the version of the Bronfman objection that I favour. Its conclusion is this: if the accuracy argument for Probabilism is to succeed, we must circumscribe the range of legitimate inaccuracy measures narrowly enough that, for any non-probabilistic credence function, there is an alternative credence function that  $\mathfrak{J}$ -dominates the first for every legitimate measure  $\mathfrak{J}$ . Granted this, the Bronfman objection does not arise. I respond by imposing the further constraint of Symmetry on the divergences that generate legitimate inaccuracy measures, and that narrows the field to just one inaccuracy measure, namely, the Brier score.

Before we move to Joyce's second objection, it's worth noting that his own formulation of the accuracy argument for Probabilism may well also be vulnerable to the Bronfman objection understood thus. Consider the particular principle of evidential support to which Joyce appeals in his version of the argument:

SUPPORT<sub>Dom</sub> If  $c^*$  accuracy dominates  $c$ , then  $c^*$  is better supported than  $c$  by every consistent body of evidence.

First, suppose Joyce endorses supervenience about accuracy. Then, since there is no credence function that  $\mathfrak{J}$ -dominates  $c_{\text{Phil}}$  for all legitimate  $\mathfrak{J}$ , there is no credence function that SUPPORT<sub>Dom</sub> tells us is better supported than  $c_{\text{Phil}}$  by every consistent body of evidence. Thus, if we opt for supervenience, Joyce's version of the accuracy argument for Probabilism fails. Next, epistemicism. Here, the situation is less clear. Via SUPPORT<sub>Dom</sub>, we can know that there is *some* credence function that is better supported than  $c_{\text{Phil}}$ , but we cannot know *which* credence functions have this feature. Is this sufficient to show that  $c_{\text{Phil}}$  is irrational? I suspect that it is. So, in contrast with my decision-theoretic version of the accuracy argument for Probabilism, Joyce's version based on principles of evidential support succeeds if epistemicism is true.

As we have seen, Joyce's first objection questions the force of the Bronfman objection and thus the necessity of introducing the stronger claim that the only legitimate inaccuracy measure is the Brier score in place of the weaker claim that all additive and continuous strictly proper inaccuracy measures are legitimate. His second objection also questions the force of an objection, namely, an extension of an objection that Alan Hájek (2008) raised against the dominance principle to which Joyce appealed in his original version of the accuracy argument for Probabilism (Joyce, 1998). In the face of this objection, I weakened that principle. Joyce maintains that the original principle holds.

Hájek's original objection runs as follows: in order to establish Probabilism, it is not sufficient to show that every credence function that violates it is dominated; you must show further that every credence function that satisfies it is not. This is a particular instance of a general condition that Hájek wishes to impose on arguments for a given norm: it is not sufficient to show that there is some Bad Thing that happens if you violate the norm; you

must also show that the Bad Thing does not happen if you satisfy it. As with Bronfman’s objection, Joyce claims that Hájek’s objection must be based on Dominators: you should think that Hájek’s condition on justification of norms is plausible only if you think that the accuracy argument for Probabilism must be advice-giving — that is, only if you think that it should tell you not only that you are irrational but also how to rectify that flaw. But again, I agree with Joyce that Dominators is false. Whatever Hájek’s original motivation, I weakened the dominance principle used in the accuracy argument for Probabilism not because of Dominators, but because of examples from practical decision-making in which Joyce’s dominance principle — you are irrational if you are dominated — seems to fail.

Joyce mentions the Name Your Fortune example that I discuss (Pettigrew, 2016a, 21). In that example, every option is dominated. If Joyce’s dominance principle is correct, every option here is irrational. That is, this is a rational dilemma. But I hold that there can be no rational dilemmas. For me, being rational is doing as well as you can in the situations you encounter. Sometimes the only options available in a situation are all flawed in certain ways. That doesn’t mean that there is no way to do your best in this situation. It may just mean that, as in Name Your Fortune, you do your best whichever option you pick. Joyce disagrees. He bites the bullet and accepts the consequence of his dominance principle: every option in Name Your Fortune is irrational. It’s not clear how to adjudicate this debate. Perhaps Joyce and I are simply working with slightly different conceptions of rationality. However, I think we can agree that it would be rather misleading to say that we had established Probabilism if the accuracy case were like Name Your Fortune. If *every* credence function is dominated, it is surely misleading to say that we’ve established Probabilism, which says that the non-probabilistic credence functions are irrational. While literally correct, doing so would surely violate the Gricean maxim of quantity. We would surely wish our accuracy argument to establish not just that all non-probabilistic credence functions are dominated, but further that no probabilistic credence function is dominated. That is Hájek’s demand.

Name Your Fortune motivates Hájek’s original objection to Joyce’s dominance principle and his alternative dominance principle — you are irrational if you are dominated and there are undominated alternatives. I offer a further example, Name Your Fortune\*, that I take to motivate my further amendment — you are irrational if you are dominated by an alternative that isn’t itself dominated. Let me reproduce the decision table for Name Your Fortune\* here — the possible worlds are  $w_1$  and  $w_2$ ; the options are  $o, o_1, o_2, \dots$ :

**Name Your Fortune\***

	$o$	$o_1$	$o_2$	$o_3$	$o_4$	$\dots$	$o_k$	$\dots$
$w_1$	2	1	2	3	4	$\dots$	$k$	$\dots$
$w_2$	2	1	$2 - \frac{1}{2}$	$2 - \frac{1}{4}$	$2 - \frac{1}{8}$	$\dots$	$2 - \frac{1}{2^{k-1}}$	$\dots$

$o$  is the only undominated option; each  $o_k$  is dominated by  $o_{k+1}$ . Thus, according to Joyce’s original dominance principle and Hájek’s amendment, the only option not ruled irrational is  $o$ . However, it seems that there would be nothing irrational about preferring  $o_{1,000,000}$  to  $o$  — after all, the former will give either 1,000,000 utiles or just very slightly below 2 utiles, while the latter gives 2 utiles for sure. Of course, Joyce might claim that  $o$  is also irrational and that Name Your Fortune\* is a rational dilemma, just like Name Your Fortune. But why? Certainly not for dominance reasons. And, furthermore,  $o$  is rationalizable — that is, there is a probability assignment relative to which  $o$  maximizes expected utility, namely, one that assigns all of its probability to  $w_2$ . And again, even if this is a rational dilemma, it would seem

misleading to say that we had established Probabilism if the accuracy case were like Name Your Fortune\* — that is, if all the non-probabilistic credence functions are dominated, and thus irrational, while all the probabilistic credence functions are not dominated, but nonetheless irrational for other reasons, it would be literally true but disingenuous to say that we had established Probabilism.

Before we leave this point, let me note that Joyce's own favoured formulation of the accuracy argument for Probabilism, which appeals to a principle of evidential support rather than a principle of decision theory, is also vulnerable to Hájek's objection and my extension of it. If, for every credence function, there is another that is better supported by every body of evidence, then Joyce must say that they are all irrational; and again, it would be misleading to say that we have established Probabilism — or at least this is not the sort of justification of Probabilism that we sought.

Let me turn now to Joyce's final objection. This targets not the accuracy argument for Probabilism, but the accuracy argument for the Principal Principle. That argument uses the same account of epistemic value as we use in the accuracy argument for Probabilism — epistemic value is accuracy, and inaccuracy is measured by the Brier score. But it deploys a different, stronger decision-theoretic principle — Chance Dominance instead of Undominated Dominance. Whereas Undominated Dominance rules out an option as irrational if there is an alternative that is guaranteed to be better (and no further alternative that is guaranteed to be better than that), Chance Dominance rules out an option as irrational if there is an alternative that is guaranteed to be *expected to be* better by the objective chance function (and no further alternative that is guaranteed to be *expected to be* better than that). Joyce raises the circularity objection to this argument (Pettigrew, 2016a, Section 10.2). The objection runs as follows: I appeal to Chance Dominance to establish the Principal Principle; but Chance Dominance follows from the Principal Principle and Maximize Subjective Expected Utility; and indeed deriving it from those two more basic principles is the only legitimate way to justify it. Thus, my argument for the Principal Principle begs the question. As Joyce notes, I consider this objection and provide two responses. Joyce considers these but rejects them.

First, I note that we might equally object to the accuracy argument for Probabilism that Dominance follows from Probabilism and Maximize Subjective Expected Utility. Joyce responds that the cases are different. In the case of Dominance, Joyce claims, it follows from Maximize Subjective Expected Utility alone, without any help from Probabilism. But I'm not sure that's true. After all, Maximize Subjective Expected Utility only holds as a norm for probabilistically coherent agents. If my credence function is not a probability function, it is not clear even how I calculate expectations, let alone whether I should try to maximize the results of those calculations. Suppose I am deciding whether to take an umbrella or not when I go out. The utility of each option is determined entirely by whether it is raining or not — other features of the world don't matter. Suppose I have credence 0.9 that it is raining and 0.1 that it is not; but I also have credence 0.1 that it is raining and Theresa May is Prime Minister, credence 0.1 that it is raining and Theresa May is not Prime Minister, credence 0.5 that it is not raining and Theresa May is Prime Minister, and credence 0.3 that it is not raining and Theresa May is not Prime Minister. Then I can calculate the expected utility of taking an umbrella relative to either of these two partitions; but since they do not cohere as Probabilism demands, they will give different answers and likely recommend different courses of action. Which am I to choose?

Second, I note that Chance Dominance is a more basic principle than the Principal Principle or Maximize Subjective Expected Utility. It applies to agents with no credences at all,

or imprecise credences, or incoherent credences. It is a basic principle of von Neumann and Morgenstern’s original formulation of decision theory, which does not posit credence functions for the agent (von Neumann & Morgenstern, 1947). Joyce says that, “when I know  $o^*$  dominates  $o$  according to your desires, I do not need to know *anything* about your beliefs to conclude that  $o$  is not your best choice”. But the same is true for Chance Dominance. Choosing dominated options is wrong because, as Joyce says, “it commits one to incurring sure losses or passing up sure gains”. Similarly, choosing chance dominated options is wrong because it commits one to incurring objective expected losses or passing up objective expected gains.

I concede, of course, that there is nothing I can say to a sceptic who asks why we should care what the objective chance function thinks of her actions; the sceptic who wonders what is so special about objective chance that it gets to dictate to her in this way. I could try to point out that, if you choose against the recommendations of Chance Dominance, then in the long run you will end up worse off with objective chance 1. But of course she might simply reply that she doesn’t see why she should care what will happen in the long run with chance 1. And there is nothing I can then say to that. But this shouldn’t undermine Chance Dominance as a principle of rationality. It is just one that sits at normative bedrock.

## 2 Reply to Briggs

R. A. Briggs’ illuminating comments raise interesting problems for my formulation of dominance reasoning and its application in the epistemic realm to the choice between alternative credence functions. While I do not wish to adopt Briggs’ final proposal, I do agree that my original formulation of this reasoning was mistaken and the problem Briggs raises demands a solution. I try to give one here.

Briggs focusses particularly on a principle that I endorse, which they formulate as follows (Pettigrew, 2016a, 23):

### **Rational Dominance** If

- (i) some  $o^*$  strongly  $\mathcal{U}$ -dominates  $o$ ; and
  - (ii)  $o^*$  is not ruled out as irrational;
- then
- (iii)  $o$  is irrational for any agent with utility function  $\mathcal{U}$ .

They raise two problems for this principle. First: When I come to give the accuracy dominance argument for Probabilism, I appeal not to Rational Dominance, but to an alternative principle, which I call Immodest Dominance (Pettigrew, 2016a, 24):

### **Immodest Dominance** Suppose $\mathcal{J}$ is a legitimate measure of inaccuracy. Then if

- (i) some  $c^*$  strongly  $\mathcal{J}$ -dominates  $c$ ; and
  - (ii)  $c^*$  is not extremely  $\mathcal{J}$ -modest;
- then
- (iii)  $c$  is irrational.

And yet Immodest Dominance does not follow from Rational Dominance if you accept other claims I make in the book. Put differently: given other commitments in the book, Rational Dominance is not sufficient to establish Probabilism.

The problem is this: in Chapter 10, I argue that an agent is irrational if she violates (a version of) the Principal Principle. But it is possible to violate that principle while still being extremely  $\mathcal{J}$ -modest — every probabilistic credence function is extremely  $\mathcal{J}$ -modest, but not all obey the Principal Principle. So the following seems possible: there is a non-probabilistic credence function  $c$  such that

- (i)  $c$  is strongly  $\mathcal{J}$ -dominated by at least one probabilistic credence function;
- (ii) all of the non-probabilistic credence functions that strongly  $\mathcal{J}$ -dominate  $c$  are themselves strongly  $\mathcal{J}$ -dominated;
- (iii) all of the probabilistic credence functions that strongly  $\mathcal{J}$ -dominate  $c$  are not extremely  $\mathcal{J}$ -modest;
- (iv) all of the probabilistic credence functions that strongly  $\mathcal{J}$ -dominate  $c$  violate the Principal Principle.

Now, in this situation, Immodest Dominance would rule  $c$  irrational, but Rational Dominance would not. So, Rational Dominance would be insufficient to establish Probabilism. And indeed Briggs identifies just such an example, namely, their Sophisticated Taj example — Taj's only probabilistic dominators violate the Principal Principle, and thus are irrational. Figure 1 gives a closely related example that will prove useful below.

Brigg's second objection to Rational Dominance turns on an ingenious parallel between the Name Your Fortune case and Yablo's paradox (Yablo, 1993). This parallel allows us to see that, for certain decision problems — Name Your Fortune amongst them — Rational Dominance is paradoxical. That is, there is no way to categorise the options in that decision problem as rational and irrational that is consistent with Rational Dominance (just as, in Yablo's paradox, there is no way to categorise the sentences as true and false that is consistent).

How to respond to these objections? I will consider two possibilities and settle finally on one that agrees with Briggs that Rational Dominance is incorrectly formulated. The first response to Briggs' objection bites the bullet in an effort to retain Rational Dominance. There are two possibilities, we might say: we might accept that there is an accuracy-first, veritist-friendly argument for the Principal Principle, perhaps of the sort outlined in Chapter 10, or we might not.

First: Suppose that we don't. Then, according to the veritist, there is nothing irrational about violating the Principal Principle — for the veritist, facts about irrationality must be grounded in and determined by facts about accuracy. So the probabilistic credence functions that dominate Sophisticated Taj are not irrational, and so Sophisticated Taj is irrational by the lights of Rational Dominance and we can establish Probabilism. And similarly for the credence function  $c_2$  in Figure 1.

Second: Suppose we do think that there is an accuracy-first, veritist-friendly argument for the Principal Principle. Then, while Sophisticated Taj is not ruled irrational by Rational Dominance, she is ruled irrational by that argument, whatever it is. After all, the Principal Principle presupposes Probabilism — though we often state it simply as a constraint on conditional probabilities (or expectations), those formulations only make sense and say

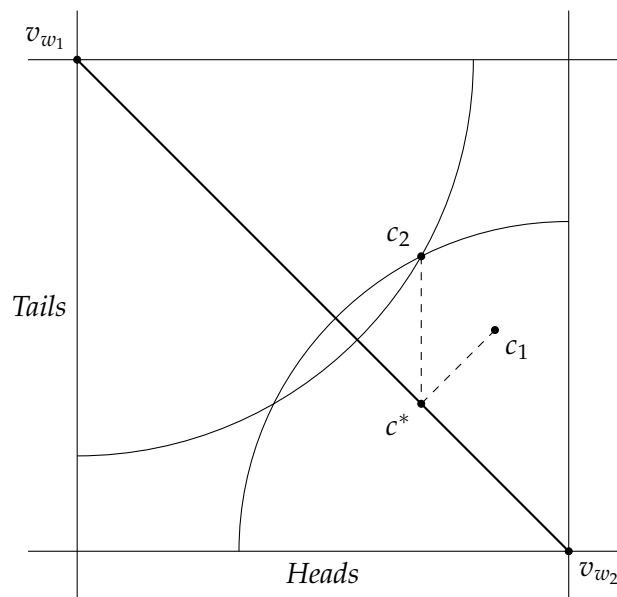


Figure 1: Suppose the only possible chance function is  $c^*$ , where  $c^*(Heads) = 0.7$  and  $c^*(Tails) = 0.3$ . Thus, the only credence function that satisfies the Principal Principle is  $c^*$ . Now consider credence functions  $c_1$  and  $c_2$ , where  $c_1(Heads) = 0.85$  and  $c_1(Tails) = 0.45$  and  $c_2(Heads) = 0.7$  and  $c_2(Tails) = 0.6$ . Both are non-probabilistic. None of the dominators of  $c_2$  satisfies the Principal Principle. But one of the dominators of  $c_1$  does, namely,  $c^*$ . Notice that  $c_1$  is nearer to its nearest rational credence function (namely,  $c^*$ ) than  $c_2$  is to its nearest rational credence function (also,  $c^*$ ).

what we wish them to say if we assume Probabilism; the ratio of  $c(A \& B)$  to  $c(B)$  doesn't represent an agent's conditional credence in  $A$  on the supposition of  $B$  unless  $c$  is a probability function. Thus, any veritist argument for the Principal Principle must also establish Probabilism. And indeed that is exactly what the chance dominance argument of Chapter 10 does. Thus, on this response to Briggs' first objection, the accuracy dominance argument doesn't establish Probabilism in full generality. Rather, it shows that some non-probabilistic credence functions are irrational, namely, those that are dominated by probabilistic credence functions that satisfy the Principal Principle — e.g.  $c_1$  in Figure 1. But, for the rest, it does not establish that they are irrational — that is established instead by the veritist argument that establishes the Principal Principle and Probabilism together.

The problem with this response to Briggs, from my point of view, is that I would like accuracy-first, veritist-friendly arguments for the laws of credence not just to establish that it is irrational to violate those laws, but also to tell us what is wrong with doing so. That is, those arguments should also give a reason for not violating the laws. But, on the response proposed here, that would mean that some non-probabilistic credence functions are irrational for one reason, and some are irrational for different reasons. Some are irrational because they are strongly dominated by probabilistic credence functions that satisfy the Principal Principle and Probabilism, while others are irrational because they are chance dominated by such credence functions. And that seems wrong.

What's more, since it is surely worse to be strongly dominated than merely chance dominated, this solution entails that those credence functions that are only chance dominated by a rational credence function are less irrational than those that are also strongly dominated by a rational credence function. Thus, for instance, in Figure 1,  $c_2$  would be less irrational than  $c_1$ . Why? Well,  $c_2$  is strongly dominated only by credence functions that violate the Principal Principle, but chance dominated by one that satisfies it (namely,  $c^*$ ). Whereas  $c_1$  is strongly dominated by a credence function that satisfies the Principal Principle (namely,  $c^*$ ). Now, I do think that irrationality comes in degrees: it is less irrational to have 0.49 in a proposition and 0.5 in its negation than to have 0.99 in the proposition and 1 in the negation, for instance. But it doesn't seem that an ordering based on whether a credence function is only chance dominated or also strongly dominated is going to match our intuitions. After all, it is usual to measure the irrationality of a credence function by the distance between it and the nearest rational credence function. And, as we can see above, on this measure,  $c_2$  is more irrational than  $c_1$ .

So I think this solution won't work. And indeed, independent of how effective it is as a response to Briggs' first objection to Rational Dominance, it has nothing to say to the second, more pressing objection that the principle is in fact paradoxical for certain decision problems.

We move, then, to our second solution. To introduce this, let us reflect on why our decision principles — Rational Dominance, Immodest Dominance, Chance Dominance, etc. — require clauses like (ii). The reason is that, when we rule a credence function irrational, we make a criticism of it. In the strong dominance argument of Part I of the book, we criticise a credence function that violates Probabilism on the grounds that there are alternatives that are guaranteed to be strictly more accurate than it; in the chance dominance argument in Chapter 10, we criticise a credence function that violates the Principal Principle on the grounds that there are alternatives whose objective expected accuracy is guaranteed to be higher than it. Now, as Name Your Fortune shows, in order to criticise an option, it is not sufficient to show that there is some option that is guaranteed to be better than it. If the options that are guaranteed to be better than it are all also criticizable, the criticism seems

unfair. If you pick the integer 99 when God asks, and I point out that 100 would have been better, it is reasonable for you to say that my criticism falls flat because 101 would have been better than that.

That is what I tried to capture in Rational Dominance. But Briggs is right to say that my attempt failed. Here is my diagnosis of the failure and my attempt to cure it. Look again at Name Your Fortune. The reason it is unfair for me to criticise you for picking 99 is not that any option guaranteed to be better than 99 is itself criticizable. It is that any such option is criticizable *in exactly the same way that 99 is criticizable*. That is, any such option has exactly the same flaw that 99 has, namely, that it is strongly dominated. It is this that makes my criticism fall flat. It wouldn't be unfair if those alternative options were also criticizable but on the basis of a different flaw.

To see this, consider yet another variant on the Name Your Fortune case. There are countably many options,  $o_1, o_2, \dots$ . And there are countably many worlds,  $w_1, w_2, \dots$ . Here is the decision table:

**Name Your Fortune<sup>+</sup>**

	$o_1$	$o_2$	$o_3$	$o_4$	$o_5$	$\dots$
$w_1$	1	2	2	2	2	$\dots$
$w_2$	1	2	3	3	3	$\dots$
$w_3$	1	2	3	4	4	$\dots$
$w_4$	1	2	3	4	5	$\dots$

In this case,  $o_1$  is strongly dominated (by  $o_2$ ), but each of the rest is only weakly dominated ( $o_k$  is weakly dominated by  $o_{k+1}$ , for instance, but not strongly dominated by any alternative). Is  $o_1$  irrational? After all, it is strongly dominated, but each of its dominators has a flaw, namely, that it is weakly dominated. This would be sufficient for Rational Dominance not rule it irrational. But it seems to me that  $o_1$  is irrational. And the reason is that, while its dominators are flawed, they have a different flaw from  $o_1$ , namely, being weakly rather than strongly dominated. And indeed their flaw seems less problematic. Being weakly dominated is less irrational than being strongly dominated — after all,  $o_2$  can be probabilistically rationalised by a probability function that places all its credence in  $w_1$ . This suggests that, in order to rule an option irrational, it is sufficient that they are strongly dominated, providing there is a strong dominator that is not itself strongly dominated. This gives the following:

**Strong Undominated Dominance** If

- (i) some  $o^*$  strongly  $\mathcal{U}$ -dominates  $o$ ; and
  - (ii)  $o^*$  is not itself strongly  $\mathcal{U}$ -dominated;
- then
- (iii)  $o$  is irrational for an agent with utility function  $\mathcal{U}$ .

And this is certainly sufficient to establish Probabilism, since it is stronger than Immodest Dominance — that is, it rules out as irrational everything that Immodest Dominance does.

Now, suppose we were only able to show that every non-probabilistic function is strongly dominated by an option that is itself weakly dominated. That is, suppose the accuracy case were like Name Your Fortune<sup>+</sup> and we required something as strong as Strong Undominated Dominance to establish Probabilism. While we would be able to show that violating



Probabilism is irrational — just as picking option  $o_1$  in Name Your Fortune<sup>+</sup> is irrational — I think we would be disappointed with the result. It would seem a weak non-pragmatic vindication of Probabilism to set alongside the pragmatic vindication provided by the Dutch Book argument. Thus, when we learn that, in fact, each non-probabilistic credence function has dominators that are not only not weakly dominated but not even extremely  $\mathcal{I}$ -modest, then we see that the criticism of non-probabilistic credence functions is much stronger.

This all suggests that what is responsible for the failure of Rational Dominance is that it tries to combine two modes of assessment in a single principle. The first is categorical: it is our attempt to categorise some credence functions as irrational and others as rational. The second is graded: it is our attempt to measure how badly flawed certain credence functions are. We do well to keep these two apart. Thus, the principles of rationality should say just this:

- **Strong Undominated Dominance** An option is irrational if it is strongly dominated by an option that is not itself strongly dominated
- **Weak Undominated Dominance** An option is irrational if it is weakly dominated by an option that is not itself weakly dominated
- **Chance Dominance** An option is irrational if it is chance dominated by an option that is not itself chance dominated.
- **Minimax** An option is irrational if it is worst-case dominated by an option that is not itself worst-case dominated.

But we are also interested in how severe the criticism of a certain sort of credence function is. Suppose  $\mathcal{S}$  is a set of credence functions. We learn first that every credence function in class  $\mathcal{S}$  is strongly dominated. So far, we cannot rule out that the situation is like the original Name Your Fortune case in which every option is strongly dominated. So we cannot even conclude that having a credence function in  $\mathcal{S}$  is irrational. Then we discover that all  $\mathcal{S}$ -functions are strongly dominated by credence functions that are not themselves strongly dominated. Now we know that having a  $\mathcal{S}$ -function is irrational, but we cannot rule out that all the dominators are themselves weakly dominated. And if that's the case, it's not so bad to have a  $\mathcal{S}$ -function. Then we learn that all the  $\mathcal{S}$ -functions are dominated by credence functions that are not strongly or weakly dominated. This makes it look worse to have a  $\mathcal{S}$ -function. And finally we learn that each  $\mathcal{S}$ -function is dominated by a probabilistic credence function that is not strongly or weakly dominated and which in fact expects itself to be best out of all credence functions. That makes it even worse to have a  $\mathcal{S}$ -function. But none of these further discoveries changes whether or not having a  $\mathcal{S}$ -function is irrational. They just tell us, once we know that it is irrational, just how flawed the irrational credence functions are. An analogy: Consider an agent with utility function  $\mathcal{U}$ , who is faced with two options,  $A$  and  $B$ . First, she discovers that  $B$  strongly  $\mathcal{U}$ -dominates  $A$ . Since there are only these two options,  $B$  cannot then be strongly  $\mathcal{U}$ -dominated itself, and so she knows that  $A$  is irrational. But it might be that  $B$ 's utility is guaranteed to be around 0.000001 utilities greater than  $A$ ; or it might be that  $B$ 's utility is guaranteed to be at least 1,000,000 utilities greater than  $A$ . And if it's the latter, then choosing  $A$  is a lot worse than if it's the former. Similarly, it isn't so bad to have a credence function that is strongly dominated if all the dominators are themselves weakly dominated; but it's very bad to have a credence function that is strongly dominated if there are dominators that expect themselves to be best.

Thus, I conclude that, in order to establish that violating Probabilism is irrational, it suffices to show that any credence function that does so is strongly dominated by one that isn't itself strongly dominated — that is, Strong Undominated Dominance. But learning that those strong dominators are not only not strongly dominated themselves, but also not weakly dominated and not even extremely  $\mathcal{J}$ -modest shows us just how bad it is, epistemically speaking, to violate Probabilism. And this shows how strong our argument is in favour of Probabilism.

### 3 Reply to Kotzen

Kotzen's comments are rich and fascinating, and I lack the space to do full justice to them here. I have considered versions of some of them to some extent in work carried out since *Accuracy and the Laws of Credence* was published. For instance, in (Pettigrew, 2016b), I consider how we might extend the measures of inaccuracy used in *ALC* so that they allow us to compare credence functions defined over different sets of propositions; in (Pettigrew, ta), I offer my solution to the so-called trade-off objection posed by Greaves (2013); and in (Pettigrew, ms), I explore how we might account for at least one further traditional epistemic virtue, namely, the virtue of justification, in accuracy-only terms, though, as Kotzen points out, there are many more such virtues still to accommodate or reject — knowledge, explanatoriness, simplicity, to name only a few. In light of this, I will focus on only some of Kotzen's concerns; in particular, those that relate to the formal features of our inaccuracy measures. But the others are important and deserve substantial treatment in their own right.

In *ALC*, I assume that the inaccuracy of a credence function at a world is its distance from perfection (Perfectionism); I assume that the perfect credence function at a world is the one that assigns maximal credence (i.e. 1) to all propositions that are true at that world and minimal credence (i.e. 0) to all propositions that are false there (Alethic Vindication); and I assume that distance from one credence function to another is measured by a *divergence*. The class of divergences is a broader class of putative distance measures than the class of metrics. Given a set  $\mathcal{X}$ , a function  $\mathcal{D} : \mathcal{X} \rightarrow [0, \infty]$  is a *divergence* on  $\mathcal{X}$  iff  $\mathcal{D}(x, y) \geq 0$  for all  $x, y$  in  $\mathcal{X}$ , with equality iff  $x = y$ .<sup>1</sup> A *metric* is a divergence that also satisfies two further conditions:

- **Symmetry**  $\mathcal{D}(x, y) = \mathcal{D}(y, x)$ , for all  $x, y$  in  $\mathcal{X}$ ;
- **Triangle Inequality**  $\mathcal{D}(x, y) + \mathcal{D}(y, z) \geq \mathcal{D}(x, z)$ , for all  $x, y, z$  in  $\mathcal{X}$ .

In fact, in response to Bronfman's objection, I do narrow the field of legitimate inaccuracy measures by arguing that the divergences that generate them should satisfy Symmetry. But at no point do I assume the Triangle Inequality, and indeed the other properties of divergences for which I argue — Divergence Additivity, Divergence Continuity, and Decomposition — are incompatible with it. Those properties exactly characterize the additive Bregman divergences (Pettigrew, 2016a, Theorem 4.3.3); and it is possible to show that no Bregman divergence satisfies the Triangle Inequality.<sup>2</sup> However, as Kotzen notes, "the triangle inequality is quite plausible even for lots of non-physical distances; for example, I have a hard

<sup>1</sup>Of course, we are most interested in the case in which  $\mathcal{X}$  is the set of credence functions defined on a given set of propositions.

<sup>2</sup>A divergence  $\mathcal{D} : [0, 1]^n \rightarrow [0, \infty]$  is an additive Bregman divergence iff there is a strictly convex and twice differentiable function  $\varphi : [0, 1] \rightarrow [0, \infty]$  such that  $\mathcal{D}(x, y) = \sum_{i=1}^n \varphi(x_i) - \varphi(y_i) - \varphi'(y_i)(x_i - y_i)$ . Now suppose

time understanding a notion of distance between two colors in color-space, or two organisms in gene-space, or two different companies in financial-valuation-space, if the relevant notion of distance doesn't obey the triangle inequality". And he asks for "some more intuitive guidance here about how to think of an alleged notion of a proximity that violates the triangle inequality". I will try to provide that guidance here.

The first thing to note is that divergences that violate the triangle inequality are used to measure non-physical distances in a diverse range of subjects. The Kullback-Leibler divergence as well as other Bregman divergences are used in information theory (Grünwald & Dawid, 2004; Banerjee et al., 2005), machine learning (Lafferty, 1999; Kivinen & Warmuth, 1999), and coding theory (Kullback, 1959) to measure the distance from one probability distribution to another; and the squared Euclidean distance divergence as well as other Bregman divergences are used in the theory of inequality measurement and mobility measurement in economics to measure the distance from one income or welfare distribution to another (D'Agostino & Dardanoni, 2009; Magdalou & Nock, 2011). However, the fact that others do this is no argument that it is permissible to do it. So let me now try to motivate the idea.

First, let's see why we might not expect the divergences used in information theory to measure the distance from one probability function to another to satisfy the Triangle Inequality. The standard measure is the Kullback-Leibler measure: if  $p, q$  are probability functions defined on a partition  $\{X_1, \dots, X_n\}$ , then

$$\mathfrak{D}_{\text{KL}}(p, q) = \sum_{i=1}^n p(X_i) \log \frac{p(X_i)}{q(X_i)}$$

And it is straightforward to see that this decomposes as follows:

$$\mathfrak{D}_{\text{KL}}(p, q) = \left( - \sum_{i=1}^n p(X_i) \log q(X_i) \right) - \left( - \sum_{i=1}^n p(X_i) \log p(X_i) \right)$$

Now, the first term on the right-hand side is known as the *cross entropy from  $p$  to  $q$* , while the second term is known as the *entropy of  $p$* . Suppose I wish to communicate which event out of  $X_1, \dots, X_n$  occurs. And suppose I develop a coding that is optimised for the distribution  $q$  — let's call this a  $q$ -coding. Then the cross entropy from  $p$  to  $q$  is the expected number of bits that would be required to identify the true element of the partition using that  $q$ -coding if the true probability distribution were in fact  $p$ . The entropy of  $p$  is then the expected number of bits required to identify the true element of the partition using a  $p$ -coding when  $p$  is in fact the true distribution. Thus, the Kullback-Leibler divergence from  $p$  to  $q$  measures the expected difference between the number of bits required to communicate the outcome using a  $q$ -coding and using a  $p$ -coding, when  $p$  is the true distribution. That is, it measures how much less efficient  $p$  takes  $q$ -codings to be than  $p$ -codings.

that  $\mathfrak{D}$  satisfies the Triangle Inequality. Then, for all  $0 < a < b$ ,

$$[\varphi(0) - \varphi(a) - \varphi'(a)(0 - a)] + [\varphi(a) - \varphi(b) - \varphi'(b)(a - b)] \geq [\varphi(0) - \varphi(b) - \varphi'(b)(0 - b)]$$

This entails  $\varphi'(b) \geq \varphi'(c)$ . So  $\varphi'$  is monotone non-increasing, and thus either  $\varphi'$  is constant or  $\varphi'$  is strictly decreasing over some interval in  $[0, 1]$ . If  $\varphi'$  is constant, then  $\varphi$  is linear and not strictly convex. If  $\varphi'$  is decreasing over an interval, then there are points at which  $\varphi''$  is negative. But, since  $\varphi$  is strictly convex,  $\varphi'' \geq 0$ , which gives a contradiction.  $\square$

With this explanation of the Kullback-Leibler divergence in hand, consider a failure of the Triangle Inequality: there are probability distributions  $p, q, r$  such that the expected decrease in efficiency that results from using a  $r$ -coding instead of a  $p$ -coding *by the lights of  $p$*  is greater than the sum of the expected decrease in efficiency that results from using a  $q$ -coding instead of a  $p$ -coding *by the lights of  $p$*  and the expected decrease in efficiency that results from using an  $r$ -coding instead of a  $q$ -coding *by the lights of  $q$* . But, having spelt this out, it doesn't seem so unintuitive. The key point is this: expected decreases in efficiency are measured by the lights of *different probability functions in the three cases of interest*. Thus,  $\mathfrak{D}_{\text{KL}}(p, q)$  and  $\mathfrak{D}_{\text{KL}}(p, r)$  measure the expected decreases in efficiency *by the lights of  $p$* , whereas  $\mathfrak{D}_{\text{KL}}(q, r)$  measures the expected decrease in efficiency *by the lights of  $q$* . An analogy will help to make the point clearer. It is as if we were to measure the distance from one person to another by how much worse the first person expects the second one to be at singing. Khaled expects Lori to be a little bit worse than him, giving a distance of 3, say; and Lori expects Maura to be a little bit worse than her, giving another distance of 3, say; but Khaled expects Maura to be *much* worse, giving a distance of 10, say. Thus, we have a violation of the triangle inequality. And violations of the triangle inequality for Kullback-Leibler divergence arise for the same reason. Indeed, I suspect that our intuitions about the Triangle Inequality arise because we imagine that the standards by which the distance is measured remain fixed from one measurement to another. But in the case of divergences like the Kullback-Leibler divergence, this doesn't happen. Of course, if we were to fix a single privileged probability function  $u$  and were to measure the distance from  $p$  to  $q$  as the expected difference in efficiency that results from using a  $q$ -coding instead of a  $p$ -coding *by the lights of  $u$* , we would recover the Triangle Inequality (and, indeed, Symmetry). But that isn't the measure we want in information theory.

Hopefully, this explains why the appropriate notion of distance between probability distributions in information theory violates the Triangle Inequality. But it doesn't yet explain why the appropriate notion of distance between credence functions in epistemology should violate it. However, the following fact helps us to see the way. Suppose  $\mathfrak{D}$  is an additive Bregman divergence; and let  $\mathfrak{J}$  be the additive and continuous strictly proper inaccuracy measure generated by  $\mathfrak{D}$ : that is,  $\mathfrak{J}(c, w) = \mathfrak{D}(v_w, c)$ . Then the following holds (Pettigrew, 2016a, Theorem I.B.4): if  $c$  is a probabilistic credence function and  $c'$  is a credence function, then

$$\mathfrak{D}(c, c') = \text{Exp}_{\mathfrak{J}}(c'|c) - \text{Exp}_{\mathfrak{J}}(c|c)$$

That is, the divergence from  $c$  to  $c'$  is the expected loss of accuracy in moving from  $c$  to  $c'$  *where the expectation is calculated by the lights of  $c$* . That is,  $\mathfrak{D}(c, c')$  measures how much less accurate  $c$  expects  $c'$  to be than it expects itself to be. So, when we use any Bregman divergence to measure the distance between credence functions it is analogous to when we use Kullback-Leibler to measure distance between probability functions; and both are analogous to when we use the singing-ability measure of distance between people introduced above. Thus, the explanation for the failure of the triangle inequality is the same: the Bregman divergence from one credence function to another measures the expected loss in inaccuracy that occurs when we move from the first to the second; and, while the same inaccuracy measure is used throughout, the standpoint from which the expectation is calculated changes for different pairs of credence function; and so it is little surprise that the triangle inequality fails.

So much for the Triangle Inequality. Kotzen's second question about our inaccuracy

measures relates to their convexity.<sup>3</sup> In his original paper, Joyce (1998) offered an argument that legitimate inaccuracy measures are strictly convex. A little later, Patrick Maher (2002) noted that, if it works at all, Joyce’s argument establishes only the weaker conclusion that they should be weakly convex (Weak Convexity). But, as Kotzen points out, the argument does seem to work, and so Joyce does seem to establish Weak Convexity. However, I don’t assume that anywhere in *ALC*. Kotzen asks what happens if I do.

Of course, while I don’t assume Weak Convexity anywhere, I do argue for Symmetry, and Symmetry entails that the only legitimate inaccuracy measure is the Brier score, which is both weakly and strictly convex (Pettigrew, 2016a, Section 4.4). That is, Symmetry entails Weak Convexity, and my favoured inaccuracy measure satisfies the conclusion that Joyce’s argument does seem to establish. But my endorsement of Symmetry is tentative. I introduce it in part because it is plausible in itself, but also in part because I take my version of the Bronfman objection to show that, if the accuracy dominance argument for Probabilism is to work, we must narrow the field of legitimate inaccuracy measures, and Symmetry does that in a principled way. I take the Bronfman objection to show that the accuracy dominance argument does not work if there are two legitimate inaccuracy measures  $\mathfrak{J}$ ,  $\mathfrak{J}'$  and some non-probabilistic credence function  $c$  such that there is no probabilistic credence function  $c^*$  that both  $\mathfrak{J}$ -dominates and  $\mathfrak{J}'$ -dominates  $c$ . Appealing to Symmetry solves this problem by restricting the legitimate accuracy measures so severely that there cannot be two legitimate inaccuracy measures at all — there is only one, and it is the Brier score. But that might seem like using a sledgehammer to crack a nut.

Perhaps instead we might restrict the legitimate inaccuracy measures by adding Weak Convexity instead of Symmetry to our other constraints? Interestingly, doing so does indeed rule out one of the two inaccuracy measures that I use in my version of the Bronfman objection. In that version, I consider the example of Phil, who has credences  $c_{\text{Phil}}(X) = 0.9$  and  $c_{\text{Phil}}(\bar{X}) = 0.2$ . And I note that there is no credence function at all that both  $\mathfrak{L}\mathfrak{A}$ -dominates and  $\mathfrak{S}\mathfrak{A}$ -dominates  $c_{\text{Phil}}$ , let alone a probabilistic one. But, while  $\mathfrak{L}\mathfrak{A}$  is strictly convex,  $\mathfrak{S}\mathfrak{A}$  is not even weakly convex. And, as you can see from Figure 5.1 in *ALC*, it is this failure of convexity that ensures that there is no overlap between the set of  $\mathfrak{L}\mathfrak{A}$ -dominators for  $c_{\text{Phil}}$  and the set of  $\mathfrak{S}\mathfrak{A}$ -dominators. This suggests the following conjecture: if we assume Weak Convexity, the Bronfman objection disappears. The conjecture is given further support from the fact that another natural way of illustrating the Bronfman objection is also resolved by restricting to inaccuracy measures that are weakly convex. For  $1 < \alpha$ , let  $\mathfrak{D}^\alpha$  be the Bregman divergence generated by the strictly convex function  $\varphi(x) = x^\alpha$ , and let  $\mathfrak{J}^\alpha$  be the inaccuracy measure generated by  $\mathfrak{D}^\alpha$ .<sup>4</sup> Then for many  $1 < \alpha, \alpha'$ , there is no probabilistic credence function  $c^*$  that both  $\mathfrak{J}^\alpha$ -dominates and  $\mathfrak{J}^{\alpha'}$ -dominates  $c_{\text{Phil}}$ . For instance, there is no probabilistic credence function that  $\mathfrak{J}^2$ -dominates and  $\mathfrak{J}^3$ -dominates  $c_{\text{Phil}}$ , and no probabilistic credence function that  $\mathfrak{J}^2$ -dominates and  $\mathfrak{J}^{\frac{3}{2}}$  dominates  $c_{\text{Phil}}$ .<sup>5</sup>

<sup>3</sup>Suppose  $\mathfrak{J}$  is an inaccuracy measure. Then  $\mathfrak{J}$  is *weakly convex* if, for  $0 < \alpha < 1$ ,

$$\mathfrak{J}(\alpha c + (1 - \alpha)c', w) \leq \alpha \mathfrak{J}(c, w) + (1 - \alpha)\mathfrak{J}(c', w)$$

$\mathfrak{J}$  is *strictly convex* if the inequality is always strict.

<sup>4</sup>Thus,  $\mathfrak{J}^\alpha(c, w) = \sum_{X \in \mathcal{F}} s^\alpha(v_w(X), c(X))$ , where  $s^\alpha(0, x) = (\alpha - 1)x^\alpha$  and  $s^\alpha(1, x) = 1 - \alpha x^{\alpha-1} + (\alpha - 1)x^\alpha$ .

<sup>5</sup>Suppose  $c$  is probabilistic. Then

- $c$   $\mathfrak{J}^2$ -dominates  $c_{\text{Phil}}$  iff  $0.8419 < c(X) < 0.8515$
- $c$   $\mathfrak{J}^3$ -dominates  $c_{\text{Phil}}$  iff  $0.8789 < c(X) < 0.8858$

However, for  $1 < \alpha$ ,  $\mathcal{J}^\alpha$  is weakly convex iff  $\alpha = 2$ . Thus, if we restrict to only the weakly convex inaccuracy measures, the instance of the Bronfman objection disappears again. Is this generally the case? Or are there weakly convex inaccuracy measures  $\mathcal{J}$ ,  $\mathcal{J}'$  and non-probabilistic  $c$  such that there is no probabilistic  $c^*$  that  $\mathcal{J}$ -dominates and  $\mathcal{J}'$ -dominates  $c$ ? Unfortunately, the latter — that is, the conjecture from above is false. Let  $\mathcal{D}^\epsilon$  be the Bregman divergence generated by the strictly convex function  $\varphi(x) = \frac{1}{e^x}$ , and  $\mathcal{J}^\epsilon$  the inaccuracy measure generated by  $\mathcal{D}^\epsilon$ .<sup>6</sup> Then there is no probabilistic credence function  $c^*$  that  $\mathcal{J}^\epsilon$ -dominates and  $\mathcal{J}^2$  dominates  $c$ .<sup>7</sup> But  $\mathcal{J}^\epsilon$  and  $\mathcal{J}^2$  are both strictly convex.

Having said that, the following intriguing fact suggests that Weak Convexity might play some role in an alternative resolution of the Bronfman objection; one that does not appeal to anything as restrictive as Symmetry. Consider the following condition (Joyce, 2009, 274):

**0/1-Symmetry** If  $\mathcal{J}$  is generated by the strictly proper scoring rule  $s$ , then  $s(0, x) = s(1, 1 - x)$ .

This says that the inaccuracy of a credence is solely a function of the difference between that credence and the omniscient credence. We can easily show that, for any  $\mathcal{J}$  that is additive and continuous and satisfies Weak Convexity and 0/1 Symmetry, any non-probabilistic credence function  $c$  defined on  $X$  and  $\bar{X}$  is  $\mathcal{J}$ -dominated by the particular credence function  $c^\dagger$  defined as follows:  $c^\dagger(X) = c(X) + \frac{1-(c(X)+c(\bar{X}))}{2}$  and  $c^\dagger(\bar{X}) = c(\bar{X}) + \frac{1-(c(X)+c(\bar{X}))}{2}$ .<sup>8</sup> How to generalise this result is an open question. But it suggests a route to answering the Bronfman objection that is more permissive than Symmetry.

## References

- Banerjee, A., Guo, X., & Wang, H. (2005). On the Optimality of Conditional Expectation as a Bregman Predictor. *IEEE Transactions of Information Theory*, 51, 2664–69.
- Briggs, R. A., & Pettigrew, R. (ms). Conditionalization. Unpublished manuscript.
- Bronfman, A. (ms.). A Gap in Joyce’s Argument for Probabilism.
- D’Agostino, M., & Dardanoni, V. (2009). What’s so special about Euclidean distance? A characterization with applications to mobility and spatial voting. *Social Choice and Welfare*, 33(2), 211–233.
- de Finetti, B. (1974). *Theory of Probability*, vol. I. New York: John Wiley & Sons.
- Greaves, H. (2013). Epistemic Decision Theory. *Mind*, 122(488), 915–952.

- 
- $c$   $\mathcal{J}^{\frac{3}{2}}$ -dominates  $c_{\text{Phil}}$  iff  $0.8008 < c(X) < 0.8255$ .

No two of these overlap.

<sup>6</sup>Thus,  $\mathcal{J}^\epsilon(c, w) = \sum_{X \in \mathcal{F}} \epsilon(v_w(X), c(X))$ , where  $\epsilon(0, x) = 1 - \frac{1+x}{e^x}$  and  $\epsilon(1, x) = \frac{1}{e} - \frac{x}{e^x}$ .

<sup>7</sup>Suppose  $c$  is probabilistic. Then

- $c$   $\mathcal{J}^2$ -dominates  $c_{\text{Phil}}$  iff  $0.8419 < c(X) < 0.8515$
- $c$   $\mathcal{J}^\epsilon$ -dominates  $c_{\text{Phil}}$  iff  $0.8272 < c(X) < 0.8348$ .

These do not overlap.

<sup>8</sup>Given a credence function  $c$  defined on  $\mathcal{F} = \{X, \bar{X}\}$ , let  $\bar{c}$  be defined as follows:  $\bar{c}(X) = 1 - c(\bar{X})$  and  $\bar{c}(\bar{X}) = 1 - c(X)$ . Then, by 0/1 Symmetry,  $\mathcal{J}(c, w) = \mathcal{J}(\bar{c}, w)$  for all  $w$ . Since  $\mathcal{J}$  is weakly convex, the set of dominators of any credence function is weakly convex. Thus, all points on the straight line from  $c$  to  $\bar{c}$  that lie strictly between the end points are in the dominator of  $c$  (and of  $\bar{c}$ ). This straight line intersects with the probabilistic credence functions at  $c^\dagger$ .  $\square$

- Greaves, H., & Wallace, D. (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, 115(459), 607–632.
- Grünwald, P. D., & Dawid, A. P. (2004). Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory. *The Annals of Statistics*, 32(4), 1367–1433.
- Hájek, A. (2008). Arguments For—Or Against—Probabilism? *The British Journal for the Philosophy of Science*, 59(4), 793–819.
- Joyce, J. M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, 65(4), 575–603.
- Joyce, J. M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In F. Huber, & C. Schmidt-Petri (Eds.) *Degrees of Belief*. Springer.
- Kivinen, J., & Warmuth, M. K. (1999). Boosting as entropy projection. In *COLT '99: Proceedings of the twelfth annual conference on computational learning theory*, (pp. 134–44).
- Kullback, S. (1959). *Information theory and Statistics*. John Wiley & Sons.
- Lafferty, J. (1999). Additive models, boosting, and inference for generalized divergences. In *COLT '99: Proceedings of the twelfth annual conference on computational learning theory*, (pp. 125–133).
- Magdalou, B., & Nock, R. (2011). Income Distributions and Decomposable Divergence Measures. *Journal of Economic Theory*, 146(6), 2440–2454.
- Maher, P. (2002). Joyce's Argument for Probabilism. *Philosophy of Science*, 69(1), 73–81.
- Pettigrew, R. (2016a). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.
- Pettigrew, R. (2016b). The population ethics of belief: in search of an epistemic Theory X. *Noûs*.
- Pettigrew, R. (ms). What is justified credence? Unpublished manuscript.
- Pettigrew, R. (ta). Making Things Right: the true consequences of decision theory in epistemology. In K. Ahlstrom-Vij, & J. Dunn (Eds.) *Epistemic Consequentialism*. Oxford: Oxford University Press.
- von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 2<sup>nd</sup> ed.
- Yablo, S. (1993). Paradox without Self-Reference. *Analysis*, 53(4), 251–2.