

Rational Probabilistic Incoherence

Michael Caie

Syracuse University

1. Introduction

The following is a plausible principle of rationality:

PROBABILISM A rational agent's credences should always be probabilistically coherent.

To say that an agent's credences are **probabilistically coherent** is to say that such credences can be represented by a function $Cr(\cdot)$ satisfying the following constraints:

NORMALIZATION For any logical truth \top , $Cr(\top) = 1$.

NONNEGATIVITY For any proposition ϕ , $0 \leq Cr(\phi)$.

FINITE ADDITIVITY If ϕ and ψ are incompatible propositions, then $Cr(\phi \vee \psi) = Cr(\phi) + Cr(\psi)$.

It has been argued that PROBABILISM follows given the plausible assumption that our primary epistemic goal is to represent the world as accurately as possible. Joyce 1998 argues that for any probabilistically incoherent credal state C that an agent might have, there is a probabilistically coherent credal state C^* that is guaranteed to be more accurate than C no matter what the world is like, while Joyce 2009 argues that, in addition, the reverse is never true. Call this the **accuracy-dominance argument** for PROBABILISM. Since it is plausible that a rational agent should

Thanks to Fabrizio Cariani and Kenny Easwaran for helpful comments on an earlier draft of this material. Thanks also go to an anonymous referee and an editor at *Philosophical Review*, as well as the audience at the Society for Exact Philosophy.

try to have as accurate a credal state as possible, the accuracy-dominance argument, if successful, would seem to provide good reason to endorse PROBABILISM.

PROBABILISM, however, has some surprising consequences. In particular, it can be shown that there are cases in which it is impossible for an agent with moderately good access to her own credal state, and a high credence in an obvious truth, to be probabilistically coherent. If PROBABILISM is true, it follows that in certain cases rationality requires that an agent either be ignorant of her own credences or be ignorant of an obvious truth.

We might simply accept this consequence of PROBABILISM, despite its prima facie implausibility. In this essay, I'll argue that this isn't the right response. To this end, I'll show that the cases in which probabilistic coherence demands either ignorance of one's own credal state or ignorance of an obvious truth can be used to expose flaws in the accuracy-dominance argument for PROBABILISM. Once these flaws are exposed, we can see that considerations of accuracy, instead of motivating PROBABILISM, support the claim that in certain cases a rational agent's credences ought to be probabilistically *incoherent*.

The essay proceeds as follows. In section 2.1, I present a case in which an agent is guaranteed to be probabilistically incoherent given that she is moderately sensitive to her own credences and has a high credence in an obvious a priori necessary truth. The obvious a priori necessary truth in question says, in effect, that this agent's credence in a proposition ϕ is far from ϕ 's actual truth-value. Andy Egan and Adam Elga have used the fact that a probabilistically coherent agent cannot have a high credence in this type of proposition and be sensitive to her own credences to argue that, for any proposition ϕ , it is rationally impermissible for an agent to have a high credence that her credence in ϕ is far from ϕ 's actual truth-value. I argue that in the case under consideration, this conclusion is implausible, and that this gives us prima facie reason to doubt PROBABILISM.

In section 2.2, I consider a related case in which an agent is guaranteed to be probabilistically incoherent simply given a moderate sensitivity to her own credences. I show how this case can be used to undermine Egan and Elga's argument. I argue that this case gives us further reason to be skeptical of PROBABILISM.

In section 3, I consider the bearing that the case considered in section 2.1 has on the accuracy-dominance argument for PROBA-

BILISM.¹ I first outline the accuracy-dominance argument. I show, by appeal to this earlier case, that there are crucial steps in the argument that are invalid. We can grant, as the argument assumes, that a credal state is defective insofar as it is guaranteed to be less accurate than some other credal state. It doesn't follow, however, that probabilistic coherence is rationally required. For there are cases in which the most accurate credal state that an agent can have is one that is probabilistically incoherent. Assuming that an agent ought to try to be as accurate as possible, it follows that in these cases an agent ought to be probabilistically incoherent.

In section 4, I present the accuracy-dominance argument for PROBABILISM in more explicit decision-theoretic terms. In decision theory, we can sometimes argue that an act or act-type is rationally required by showing that the act or act-type is better than all the alternatives, no matter what the state of the world. In this type of case, we say that the act or act-type **dominates** its alternatives. It is well known, however, that in order to apply dominance reasoning in this way, the acts and states must be related in a particular way. We call this relation **independence**. In this section, I show that the accuracy-dominance argument for PROBABILISM, framed in explicit decision-theoretic terms, fails when it is applied to the case discussed in section 2.1 because the acts and states appealed to in this argument are not independent. I show, further, that when this defect is corrected, we can provide an accuracy-dominance argument for the conclusion that in these cases it is rationally required that the agent be probabilistically incoherent.

The case that I appeal to in sections 2–4 crucially involves a proposition such that necessarily the truth-value of the proposition depends on a particular agent's credence in that very proposition. This raises the question of whether there is some suitably restricted version of PROBABILISM that might still be made to work. Is there some large interesting class of propositions such that a rational agent's credences over those propositions must always be probabilistically coherent?

In section 5, I take up this question and argue that the answer is no. In principle, almost any proposition can be such that an agent may rationally fail to have probabilistically coherent credences in that proposition and its negation.

Finally, in section 6, I briefly look at the bearing that the case from section 2.1 has on Dutch Book arguments for PROBABILISM. I show that,

1. Although I focus on the case in sec. 2.1, the same points could all be made by appeal to the case considered in sec. 2.2.

as in the case of the accuracy-dominance argument, this case can be used to expose crucial flaws in this argument for **PROBABILISM**. Dutch Book considerations, on inspection, give us no reason to deny that an agent may rationally fail to have probabilistically coherent credences.

2. Some Prima Facie Problems for Probabilism

Call one who is reliably mistaken in her judgment about ϕ an **anti-expert** about ϕ .² Andy Egan and Adam Elga (2005) have shown:

ANTI-EXPERTISE If an agent has a high credence that she is an anti-expert about ϕ and, in addition, is sensitive to her own credence in ϕ , then she will be probabilistically incoherent.

Egan and Elga endorse **PROBABILISM**. In addition, they endorse the following principle:

RATIONAL INTROSPECTION A rational agent must be responsive to her own credal state.

ANTI-EXPERTISE shows that **PROBABILISM** and **RATIONAL INTROSPECTION** together entail that it is rationally impermissible for an agent to have a high credence in her own anti-expertise.³ Thus, according to Egan and Elga (2005, 83), “It is never rational to count oneself as an anti-expert because doing so must involve either [probabilistic] incoherence or poor access to one’s own beliefs.”⁴

I claim that certain cases of anti-expertise in fact provide us with good reason to reject **PROBABILISM**. In this section, I’ll lay the groundwork for this argument by looking at two cases of anti-expertise that give us good prima facie reason to be skeptical of **PROBABILISM**. In sections 3–4, I’ll argue that this skepticism is warranted. These cases can be used

2. Following Sorensen 1998, we can distinguish two types of anti-expertise. Focus, for the moment, on qualitative beliefs. We say that an agent is a *commissive anti-expert* about the proposition ϕ just in case either it’s the case that $\neg\phi$ and the agent believes ϕ , or it’s the case that ϕ and the agent believes $\neg\phi$, that is, just in case $(\neg\phi \wedge B(\phi)) \vee (\phi \wedge B(\neg\phi))$. We say that an agent is an *omissive anti-expert* just in case either it’s the case that $\neg\phi$ and the agent believes ϕ , or it’s the case that ϕ and the agent doesn’t believe ϕ , that is, just in case $\phi \leftrightarrow \neg B\phi$. If we switch to talking about credences, we can then distinguish varying degrees of commissive and omissive anti-expertise.

3. This follows given a plausible multipremise closure principle for rational obligations: $\phi, \psi \models \gamma \Rightarrow O\phi, O\psi \models O\gamma$.

4. This same claim is defended in the case of qualitative belief in Sorensen 1998.

to show that in certain situations a probabilistically *incoherent* credal state maximizes credal accuracy. This gives us, I think, good reason to reject **PROBABILISM**.

In section 2.1, I show that there are cases in which it is an *obvious a priori necessary truth* that a particular agent is an anti-expert about some proposition ϕ . This shows that **PROBABILISM** can require a rational agent to satisfy the following disjunctive obligation: either the agent must have low credence in an obvious a priori necessary truth or the agent must be insensitive to her own credal state. I agree with Egan and Elga that an agent should never be counted as irrational in virtue of having good epistemic access to her own credal state. According to Egan and Elga, the moral to draw from this case is that the agent should have low credence that she is an anti-expert about ϕ . However, it is prima facie rather implausible that an agent could be counted as irrational for having a high credence in this type of obvious truth. Since the disjunctive obligation entailed by **PROBABILISM** is prima facie implausible, this type of case provides us with prima facie reason to be skeptical of this putative rational constraint.

In section 2.2, I show that there are cases in which an agent who is sensitive to her own credal state will be guaranteed to be probabilistically incoherent. I first show how this type of case can be used to undermine Egan and Elga's argument that the self-ascription of anti-expertise is irrational. Egan and Elga infer this conclusion from the fact that **PROBABILISM** and **RATIONAL INTROSPECTION** are jointly incompatible with the claim that it is rational to believe that one is an anti-expert. However, since **PROBABILISM** and **RATIONAL INTROSPECTION** are, in certain cases, jointly unsatisfiable, a plausible ought-implies-can principle tells us that we must reject at least one of these claims, and so reject at least one of the premises of Egan and Elga's argument. I next note that this type of case provides further prima facie reason to reject **PROBABILISM**. For it tells us that if **PROBABILISM** is true, then in certain cases, it is rationally required that an agent have *poor* access to her own credences. However, it is prima facie implausible that an agent could be rationally required to have poor epistemic access to her own credal state, and so this case gives us additional reason to be skeptical of **PROBABILISM**.

2.1.

Consider an agent who we'll call 'Yuko'. Let us suppose that Yuko introduces the name '(*)' and stipulates that it refer to the following interpreted sentence:

Yuko's credence that (*) is true isn't greater than or equal to 0.5.⁵

We'll use ' Cr_y ' to abbreviate 'Yuko's credence that...'. The above can then be represented as:

$$(*) \quad \neg Cr_y T(*) \geq 0.5$$

As an instance of the T-schema, we have:

$$(1) \quad T(*) \leftrightarrow \neg Cr_y T(*) \geq 0.5$$

If classical logic is correct (and I'll assume here that it is), then we shouldn't accept *every* instance of the T-schema.⁶ As is well known, there are certain instances of this schema, for example, instances involving liar sentences, that are inconsistent given classical logic. We should certainly reject *these* biconditionals. In the vast majority of cases, however, there is no conflict with classical logic. Given the intuitive plausibility of these biconditionals, if there is no logical reason to reject an instance of this schema, we should, I think, endorse it.⁷ (1) is perfectly consistent with classical logic. We should, therefore, accept this claim.⁸

5. Two points. First, we should think of the name 'Yuko' as having a first-personal mode of presentation for Yuko. Second, we achieve sentential self-reference here via stipulation as in Kripe 1975. This could, however, also be achieved by a coding technique such as Gödel-numbering.

6. It is worth noting that if we give up the assumption that classical logic is correct, then there are interesting ways of treating the types of cases we'll be looking at in this section. See Caie 2012 for a discussion of how to treat similar cases that arise for qualitative belief using nonclassical resources.

7. There are perfectly general non-ad hoc treatments of the truth-predicate for which this holds. This will, for example, hold according to the theory KF. See Field 2008 for a description of this theory.

8. For those who are worried about the truth of (1), let me note that this case could easily be run without appeal to a truth-predicate. What is required for this case is that there be some proposition ϕ for which we have $\phi \leftrightarrow \neg Cr_y(\phi) \geq 0.5$. The claim that (*) is true is a particularly simple case, but there are other propositions that could work. For example, we might assume that Yuko is constituted so that whether she will make a free throw in basketball depends on how confident she is that she will make the shot. We can assume that she will make the shot, but only if she is less confident that she will make it than that she will miss. This is no doubt an unusual situation, but there doesn't seem to be anything impossible about things being this way. Instead of setting up the case using the claim that

We make the following assumptions about Yuko:

- (2) $Cr_y(T(*) \leftrightarrow \neg Cr_y T(*) \geq 0.5) \geq 1 - \varepsilon$
- (3) $[Cr_y T(*) \geq 0.5] \rightarrow [Cr_y(Cr_y T(*) \geq 0.5) > 0.5 + \varepsilon]$
- (4) $[\neg Cr_y T(*) \geq 0.5] \rightarrow [Cr_y(\neg Cr_y T(*) \geq 0.5) > 0.5 + \varepsilon]$

(1) expresses the claim that Yuko is an anti-expert about the truth of (*).⁹ (2), then, expresses the claim that Yuko has credence at least $1 - \varepsilon$ in her own anti-expertise about the truth of (*). If $\varepsilon = 0$, then Yuko is completely certain of her own anti-expertise with respect to the truth of (*) and is at least mildly sensitive to her own credence in the truth of (*). As ε increases, Yuko's credence in her own anti-expertise with respect to the truth of (*) decreases, and her assumed sensitivity to her credence in the truth of (*) increases. These assumptions are jointly satisfiable as long as $0 \leq \varepsilon \leq 0.5$.¹⁰

We can show:

From (2)–(4), it follows that Yuko is probabilistically incoherent.

To see this, first assume: $Cr_y T(*) \geq 0.5$. By (3), we have:

$Cr_y(Cr_y T(*) \geq 0.5) > 0.5 + \varepsilon$. From (2), we know that if Yuko is probabilistically coherent, then $|Cr_y \neg T(*) - Cr_y(Cr_y T(*) \geq 0.5)| \leq \varepsilon$. Thus, assuming that Yuko is probabilistically coherent, we have: $Cr_y \neg T(*) > 0.5$. Given our original assumption, this guarantees that $Cr_y T(*) + Cr_y \neg T(*) > 1$. But probabilistic coherence requires, in general, that $Cr_y(\phi) + Cr_y(\neg\phi) = 1$. On the assumption that Yuko has credence greater than or equal to 0.5 in the truth of (*), it follows that Yuko is probabilistically incoherent.

(*) is true, then, we could use the claim that Yuko will make the relevant free throw. We'll look in more detail at this type of case in sec. 5. For now, however, I'll stick with (1).

9. More precisely, it expresses the claim that she is an ommissive anti-expert about (*). See n. 2. Note that, for any agent, we can construct a sentence like (*) such that it is a truth that the agent is an anti-expert about the proposition expressed by that sentence. Every agent, then, is an anti-expert about some claim.

10. If $\varepsilon > 0.5$, then by (3) and (4), we would either have to have $Cr_y(Cr_y T(*) \geq 0.5) > 1$ or $Cr_y(\neg Cr_y T(*) \geq 0.5) > 1$. But, I assume, it is impossible for an agent to have credence greater than 1 in any proposition. If one rejects this assumption, then one will hold that there are more values for ε for which (2)–(4) are jointly satisfiable.

Next, assume: $\neg Cr_y T(*) \geq 0.5$. By (4), we have: $Cr_y(\neg Cr_y T(*) \geq 0.5) > 0.5 + \varepsilon$. From (2), we know that if Yuko is probabilistically coherent, then $|Cr_y T(*) - Cr_y(\neg Cr_y T(*) \geq 0.5)| \leq \varepsilon$. Thus, assuming that Yuko is probabilistically coherent, we have: $Cr_y T(*) > 0.5$. But this is incompatible with our assumption that $\neg Cr_y T(*) \geq 0.5$. It follows that, on the assumption that Yuko doesn't have credence greater than or equal to 0.5 in the truth of (*), Yuko is probabilistically incoherent.

Since it follows that Yuko is probabilistically incoherent both on the assumption that $Cr_y T(*) \geq 0.5$ and on the assumption that $\neg Cr_y T(*) \geq 0.5$, we can conclude that Yuko is probabilistically incoherent.

Rational obligations, I assume, are closed under logical consequence. That is, we have: $\phi \models \psi \Rightarrow O\phi \models O\psi$. We have seen that Yuko will satisfy the requirements imposed by PROBABILISM only if either (2) fails to hold, (3) fails to hold, or (4) fails to hold. If PROBABILISM is true, then it follows that Yuko is rationally obligated to be such that one of (2)–(4) fail.

For (2) to fail, Yuko must have credence in (1) below $1 - \varepsilon$. For (3) to fail, Yuko must have credence greater than or equal to 0.5 in the truth of (*), but have at best $0.5 + \varepsilon$ credence that her credence in the truth of (*) falls in this range. For (4) to fail, Yuko must fail to have credence greater than or equal to 0.5 in the truth of (*), but have at best $0.5 + \varepsilon$ credence that her credence in the truth of (*) fails to fall in this range. PROBABILISM thus prohibits Yuko from both being confident that (1) expresses a truth and being somewhat sensitive to her own credence in the truth of (*).

This should, I think, strike you as rather surprising.

First, note that (1) is an obvious truth that is, plausibly, a priori knowable for Yuko. For given that Yuko stipulates that the name '(*)' refers to ' $\neg Cr_y T(*)$ ', it would seem to follow that she is thereby in a position to know that (1) is true without any further investigation. We may compare this case to Kripke's famous case of the standard meter stick. Let S be the standard meter stick. We're told to imagine someone introducing the term 'one meter' by stipulating: "By 'one meter' we will mean the length of S at t_0 ." Kripke asks, "What, then, is the epistemological status of the statement 'Stick S is one meter long at t_0 ' for someone who has fixed the reference of the metric system by reference to stick S?" He answers, plausibly, "It would seem that he knows it a priori. For if he

used stick S to fix the referent of ‘one meter’, then as a result of this kind of ‘definition’ . . . he knows automatically, without further investigation, that S is one meter long” (Kripke 1972, 56). This seems quite plausible to me. And just as Kripke’s agent is able to know a priori that S is one meter long in virtue of his introduction of the term ‘one meter’, so too will Yuko be able to know a priori that (1) is true in virtue of her introduction of the term ‘(*)’.¹¹

Since (1) is an obvious truth that is a priori knowable by Yuko, it would seem to be an epistemic defect for Yuko to fail to have a high credence in (1). But it would also seem to be an epistemic defect for an agent to fail to be sensitive to her own credal state. It follows, then, that PROBABILISM may require Yuko to exhibit at least one of these apparent epistemic defects. It is, however, *prima facie* implausible that rationality could require an agent to choose between such epistemic vices.

As we’ve noted, the conclusion that Egan and Elga draw from this type of case is that Yuko should avoid having a high credence in (1). Of course, *we* have a high credence in (1). And it seems that Yuko has the same evidence that we have, in virtue of which (1) is a manifest truth. A natural worry, then, about Egan and Elga’s claim is that it appears to countenance different responses to the evidence in favor of the truth of (1) depending on who it is that is considering this evidence. But if it would be an epistemic defect for us to have a low credence in (1) given knowledge of which sentence ‘(*)’ refers to, how could the same not be true of Yuko?

Egan and Elga (2005, 86–87) consider this type of worry. To illustrate the worry they consider the case of Professor X, an unfortunate geologist who is an anti-expert about geology. As they note, *we* could rationally have high credence that Professor X is a geological anti-expert given an appropriate list detailing her past inaccurate geological judgments. Moreover, if Professor X were presented with a list detailing the same past inaccuracies of some other individual, it would seem that she could rationally come to have a high credence that this individual is an anti-expert about geology. So, the question that arises is: couldn’t Professor X

11. To be clear there are important differences between this case and Kripke’s meter stick case. In particular, Kripke uses the meter stick case to argue that there are a priori contingent truths. (1), however, is both a priori and necessary. The relevant point of similarity is just that in certain circumstances if an agent introduces a term in a certain way, certain propositions are then a priori knowable for that individual.

rationally have high credence that she is an anti-expert about geology if it were revealed to her that this was a list of her own past inaccuracies?

According to Egan and Elga, the answer is no. Instead, the correct response for Professor X, were she to be given such evidence of her own anti-expertise, is to revise her opinions about geology so that she is no longer an anti-expert.

In the case of Professor X, this is, I think, the right thing to say. For in this case, we may suppose, it is possible for Professor X to change her credences in light of the evidence of her past anti-expertise so that she is no longer an anti-expert about geology.

Indeed, in many cases it will be possible for an agent to respond to evidence of her own anti-expertise in such a way as to eliminate her anti-expertise. For example, an agent's evidence may support the claim that it is raining as well as the claim that she has low credence that it is raining. Her evidence, then, supports the claim that she is an anti-expert about whether it is raining. In this case, the right response to this evidence would seem to be for the agent to become appropriately confident that it is raining, that is, for her to adjust her credences so that she is no longer an anti-expert about rain.

However, although Egan and Elga's suggested response to evidence of anti-expertise seems plausible in the case of Professor X and in the case of our anti-expert about rain, importantly the same is not true in the case of Yuko. The reason for this is that (1) is a necessary truth that holds regardless of what Yuko's credence is in (*).¹² Since Yuko is necessarily an anti-expert with respect to the truth of (*), she cannot respond to evidence of her own anti-expertise in such a way as to make it the case that she is not an anti-expert about (*). The following is a plausible general constraint on principles of rationality:

OUGHT-CAN It must always be possible for an agent to meet the requirements imposed by rationality.

Given that Yuko cannot respond to evidence for the truth of (1) by making it false, it follows from OUGHT-CAN that she is not rationally

12. Note that (1) is necessary only on the assumption that (*) refers rigidly to an *interpreted* sentence. If we took (*) to simply refer to a string of graphemes, then, despite the fact that in the actual world (*) is true just in case Yuko doesn't have credence at or above 0.5 that (*) is true, this need not hold at some other world in which those graphemes have a different meaning. That (*) refers to the interpreted sentence, however, was part of Yuko's stipulation in introducing the name.

required to so respond. Egan and Elga's suggested response to evidence of anti-expertise, therefore, doesn't apply in the case of Yuko.

How then should Yuko respond to the evidence in virtue of which (1) is a manifest truth? Since she cannot make (1) false, the natural answer is that she ought to respond to such evidence in the way that we do, namely, by having a high credence in this proposition.

On inspection, then, it would seem that Yuko could be highly confident in her own anti-expertise about the truth of (*) and sensitive to her own credal state without thereby being irrational. **PROBABILISM**, however, demands, implausibly, that she either turn a blind eye to the truth of (1) or to her own credences. Given the implausibility of this claim, we have some reason to doubt that **PROBABILISM** is true.

This, of course, doesn't provide anything like a conclusive argument against **PROBABILISM**. For while it is certainly *prima facie* plausible that Yuko could have high credence in (1) and be sensitive to her own credal state without thereby being irrational, **PROBABILISM** is also *prima facie* plausible. At present, then, we have a set of *prima facie* plausible claims that are jointly incompatible, but we don't have clear, decisive reason for rejecting **PROBABILISM** in particular. In sections 3–4, I'll show that on closer inspection the case of Yuko provides us with the materials to develop a strong argument against **PROBABILISM**. Before doing so, however, let's look at another case that provides additional motivation for the rejection of **PROBABILISM**.

2.2.

Let '(#)' name the following sentence:

Hiro's credence in the proposition expressed by (#) isn't greater than or equal to 0.5.

We'll use ' Cr_h ' to abbreviate 'Hiro's credence in ...' and we'll use ' ρ ' to abbreviate 'the proposition expressed by'. The above can then be represented as:

$$(4) \quad \neg Cr_h \rho(\#) \geq 0.5$$

Note that since both '(#)' and ' $\neg Cr_h \rho(\#) \geq 0.5$ ' refer to the same sentence, we have:

$$(5) \quad \rho(\#) = \rho(\neg Cr_h \rho(\#) \geq 0.5)$$

We'll assume the following facts about Hiro's introspective powers. We'll assume that if Hiro has credence greater than or equal to 0.5 in the proposition expressed by (#), then Hiro has credence greater than 0.5 in the proposition that he has credence greater than or equal to 0.5 in the proposition expressed by (#). We'll also assume that if Hiro does not have credence greater than or equal to 0.5 in the proposition expressed by (#), then Hiro has credence greater than 0.5 in the proposition that he does not have credence greater than or equal to 0.5 in the proposition expressed by (#). We can represent these assumptions as follows:

- (6) $[Cr_h\rho(\#) \geq 0.5] \rightarrow [Cr_h(\rho'Cr_h\rho(\#) \geq 0.5') > 0.5]$
 (7) $[\neg Cr_h\rho(\#) \geq 0.5] \rightarrow [Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') > 0.5]$

We can show:

From (5)–(7), it follows that Hiro is probabilistically incoherent.

Assume: $\neg Cr_h\rho(\#) \geq 0.5$. By (5), we can substitute ' $\rho'\neg Cr_h\rho(\#) \geq 0.5'$ ' for ' $\rho(\#)$ ' within attitude ascriptions *salva veritate*. Thus, from our assumption and (5), we have: $\neg Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') > 0.5$. But from our assumption, it follows, given (7), that we have: $Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') > 0.5$.

Since the assumption that $\neg Cr_h\rho(\#) \geq 0.5$ leads to a contradiction, it follows, given (5)–(7), that $Cr_h\rho(\#) \geq 0.5$, that is, that Hiro has credence at least as great as 0.5 in the proposition expressed by (#). But now we can show that Hiro is doomed to probabilistic incoherence.

For probabilistic coherence requires that Hiro's credence in the proposition expressed by (#) and his credence in its negation sum to one, that is, that $Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') + Cr_h(\rho'Cr_h\rho(\#) \geq 0.5') = 1$. But given that Hiro has credence at least as great as 0.5 in the proposition expressed by (#), we can show that it follows that $Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') + Cr_h(\rho'Cr_h\rho(\#) \geq 0.5') > 1$.

From $Cr_h\rho(\#) \geq 0.5$, it follows, given (5), that: $Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') \geq 0.5$. But from $Cr_h\rho(\#) \geq 0.5$ and (6), it follows that: $Cr_h(\rho'Cr_h\rho(\#) \geq 0.5') > 0.5$. Thus, we have: $Cr_h(\rho'\neg Cr_h\rho(\#) \geq 0.5') + Cr_h(\rho'Cr_h\rho(\#) \geq 0.5') > 1$.

There are a few important lessons that we can draw from this case. First, this case shows that one should not endorse both **PROBABILISM**

and RATIONAL INTROSPECTION. As we've seen, Hiro will satisfy the requirements imposed by PROBABILISM only if either (6) or (7) fail to hold. According to RATIONAL INTROSPECTION, however, (6) and (7) are requirements of rationality. It follows, then, that Hiro is unable to meet all of the requirements imposed by PROBABILISM and RATIONAL INTROSPECTION. Given this, if one were to endorse both PROBABILISM and RATIONAL INTROSPECTION, then one would be committed to the existence of a rational dilemma. In particular, one would be committed to the claim that rationality requires of a certain agent that he have good access to his own credences and to the claim that rationality requires of this same agent that he have poor access to his own credences. It follows, then, from OUGHT-CAN, that one shouldn't endorse both of these putative constraints on rationality. Of course, one might simply bite the bullet here and accept that an agent may sometimes be faced with a rational dilemma. But this seems to me to be poorly motivated. I think we will do better if we let OUGHT-CAN guide our judgments in this case and infer that PROBABILISM and RATIONAL INTROSPECTION aren't both correct.

The case of Hiro, then, serves to undermine Egan and Elga's argument for the claim that it is always irrational for an agent to have high credence in her own anti-expertise. For, according to Egan and Elga, the reason that we should endorse this latter claim is that it follows from PROBABILISM and RATIONAL INTROSPECTION. However, since we shouldn't accept both of these principles, the fact that they entail that it is irrational for an agent to self-ascribe anti-expertise gives us no reason to accept this claim.

The next point to note is that the case of Hiro provides us with further reason to be skeptical of PROBABILISM. For, if PROBABILISM is true, it follows that it is a requirement of rationality that Hiro be such that either (i) he has credence greater than or equal to 0.5 in the proposition expressed by (#) but has at best 0.5 credence, that is, is at best agnostic, that his credence in this proposition is in this range, or (ii) he fails to have credence greater than or equal to 0.5 in the proposition expressed by (#), but has at best 0.5 credence, that is, is at best agnostic, that his credence in this proposition fails to be in this range. PROBABILISM thus demands that Hiro be insensitive to his own credal state. But it is quite implausible that an agent could be rationally required to have poor epistemic access to his own credal state. The fact that PROBABILISM requires us to accept such an implausible conclusion, then, provides another reason to doubt

the claim that a rational agent must always have probabilistically coherent credences.

Let me close this section by addressing one potential worry about this case. Why, one may ask, should we assume that (#) does in fact express a proposition that could serve as the object of Hiro's doxastic attitudes? I take it that if there is a worry here about the existence of an appropriate proposition, it stems from the self-referential nature of (#). In response, let me say the following.

First, we should fix on some diagnostic tests for whether a sentence ϕ expresses a proposition. I take it that a sufficient condition for ϕ to express a proposition is if ϕ can be embedded under metaphysical modal operators in a way that results in a true sentence. For the resultant sentence could be true only if it expressed a proposition; and such a sentence could express a proposition only if its component sentences expressed propositions. A sentence's failure to express a proposition is something that infects any sentence of which it is a part.

Given this, we can show that a sentence is not, in general, precluded from expressing a proposition in virtue of the fact that it contains a term that purports to refer to the proposition expressed by that sentence.

One way to achieve sentential self-reference is via stipulation, as in the case of (#). Another is via a definite description that picks out the sentence in which the definite description occurs. Imagine, for example, that in room 301 there is a single blackboard, and on that blackboard is written the following sentence: 'The proposition expressed by the sentence on the blackboard in room 301 is not true'. In this case, the definite description 'the sentence on the blackboard in room 301' refers to the very sentence of which that definite description is a constituent. And so the definite description 'the proposition expressed by the sentence on the blackboard in room 301' purports to refer to the proposition expressed by that sentence.

To argue that this sentence does indeed express a proposition, it suffices to argue that this sentence can embed under a metaphysical modal operator in a way that results in a true sentence. But it seems fairly obvious that this is the case. Consider a possible world in which the sentence written on the blackboard in 301 is ' $2 + 2 = 5$ '. We assume that in this world the proposition expressed by the sentence written on the blackboard in 301 is just the proposition that $2 + 2 = 5$. In this world, then, the proposition expressed by the sentence written on the blackboard in room 301 is not true. But then it follows that it is possible that the proposition expressed by the sentence written on the blackboard in room 301 is

not true.¹³ It would seem, then, that we can embed ‘The proposition expressed by the sentence on the blackboard in room 301 is not true’ under the operator ‘It is possible that . . .’ and get a true sentence. And if that’s the case, then ‘The proposition expressed by the sentence on the blackboard in room 301 is not true’ must express a proposition.

The above reflections show that a sentence is not barred from expressing a proposition simply in virtue of containing a term that purports to refer to the proposition expressed by that sentence. The fact that (#) contains such a term does not, then, provide us with good reason to deny that this sentence expresses a proposition. In the absence of a more convincing reason to the contrary, I think we should therefore assume that (#) does indeed express a proposition that can serve as the object of Hiro’s doxastic attitudes. And given that this is case, we have good reason to reject Egan and Elga’s argument that it is irrational to self-ascribe anti-expertise, and, furthermore, we have good reason to be skeptical of PROBABILISM.

Having noted some *prima facie* problems for PROBABILISM, I’ll now turn to showing how the case of Yuko can be used to mount a strong direct argument for the claim that in certain situations it is rational for an agent to have probabilistically incoherent credences.

3. Probabilism and Accuracy

We can assess a qualitative doxastic state in terms of how accurate it is. Consider an agent’s attitude toward a single proposition ϕ . If ϕ is true, we can say:

Believing ϕ is more accurate than being agnostic about ϕ , and being agnostic about ϕ is more accurate than disbelieving ϕ .

While, if ϕ is false, we can say:

Disbelieving ϕ is more accurate than being agnostic about ϕ , and being agnostic about ϕ is more accurate than believing ϕ .

It’s plausible to think that our primary epistemic goal in forming beliefs is to represent matters as accurately as we can. In forming beliefs, we aim to have true beliefs and avoid having false beliefs.

We may appeal to the fact that accuracy is our primary epistemic goal to justify certain claims about doxastic rationality. For example, we

13. Of course, this is only true on the *de dicto* reading of the above sentence.

may argue that it is never rational to believe $\phi \wedge \neg\phi$. Since $\phi \wedge \neg\phi$ is guaranteed to be false, we are guaranteed to be more accurate if we don't believe $\phi \wedge \neg\phi$ than if we do. Since we ought to try to be as accurate as possible in our judgments, and since this goal is best achieved by never believing $\phi \wedge \neg\phi$, we ought not believe $\phi \wedge \neg\phi$.

Just as we can assess a qualitative doxastic state for accuracy, so too can we assess a quantitative doxastic state, that is, a credal state. Consider an agent's credence in a single proposition ϕ . If ϕ is true, we can say:

A higher credence in ϕ is more accurate than a lower credence.

While, if ϕ is false, we can say:

A lower credence in ϕ is more accurate than a higher credence.

Just as it is plausible to think that our primary goal in forming qualitative doxastic attitudes is to be as accurate as we can in our judgments of truth-value, so too it is plausible that our primary goal in forming quantitative doxastic attitudes is to be as accurate as we can in our estimation of truth-values.

It is a tricky question exactly how credal accuracy should be measured. There are numerous ways of measuring the accuracy of credences in particular propositions that meet the above constraints. And there are numerous ways of measuring the accuracy of a total credal state given the accuracy of particular credences.

In Joyce 1998, it is argued, however, that for any reasonable way of measuring accuracy, the following holds:

PCA 1 For any probabilistically incoherent credal state C , there is a probabilistically coherent credal state C^* , such that C^* would be more accurate than C , no matter what the actual world is like.

And in Joyce 2009, it is further argued that for any reasonable way of measuring accuracy, the following also holds:

PCA 2 For any probabilistically coherent credal state C^* , there is no probabilistically incoherent credal state C , such that (i) C would be at least as accurate as C^* no matter what the actual world is like, and (ii) C would be more accurate than C^* given at least one possible state of the world.

Given PCA 1 and PCA 2, a powerful argument can be given for PROBABILISM. By PCA 1, if an agent has a probabilistically incoherent credal state C , there is some probabilistically coherent credal state C^* that would

have been more accurate than C no matter what the actual world is like. Assuming that accuracy is our primary epistemic goal, it follows that from an epistemic perspective, the agent should see C^* as being preferable to C . By PCA 2, there is no countervailing reason to find any probabilistically incoherent credal state preferable to C^* . Thus, from an epistemic perspective, an agent should always prefer being probabilistically coherent to being probabilistically incoherent.¹⁴

What PCA 1 and PCA 2 show, if they're correct, is that the goal of credal accuracy is best achieved by being probabilistically coherent. Assuming that one ought to try to have credences that are as accurate as possible, it follows that one ought to be probabilistically coherent.

I'm happy to say that accuracy is our primary epistemic goal. Indeed, I'll assume that this is so throughout this essay. But this idea doesn't support PROBABILISM. For, both PCA 1 and PCA 2 are false. To show this, I'll show that there are cases in which:

There is a probabilistically incoherent credal state C such that, for any probabilistically coherent credal state C^* , an agent would be less accurate were her credal state to be C^* instead of C , no matter what the actual world is like.

In certain cases, the goal of credal accuracy is best achieved by being probabilistically *incoherent*. Since one ought to try to have credences that are as accurate as possible, in such cases one ought to have probabilistically incoherent credences.

In what follows, we'll consider an agent who has credences defined over a finite algebra of propositions \mathcal{P} .¹⁵ To say that \mathcal{P} is an algebra is to say that membership in the set is closed under negation and finite disjunction. We'll represent the agent's credal state by the function $Cr(\cdot)$.

14. I find this argument quite convincing. But see Easwaran and Fitelson 2012 for an interesting argument that other epistemic goods may in certain cases rule out accuracy-dominating credal states.

15. In order to ensure finiteness, I'll assume that a proposition is identical to the set of worlds in which it is true. Of course, for certain purposes we may want to think of propositions in a more fine-grained way, but for our purposes here, nothing will be lost by taking this coarse-grained approach. I should also note that nothing essential turns on our assumption that the algebra over which the credences are defined is *finite*, but this will help simplify the presentation at certain points.

In arguing for PCA 1 and PCA 2, Joyce goes to great lengths to try to show that these claims will hold for a large number of possible ways of measuring the accuracy of credences. For the sake of simplicity, I will focus on one of these measures, but none of the points that follow turn essentially on any idiosyncratic features of this measure.

We assume, then, the following:

BRIER ACCURACY Given an agent with credences $Cr(\cdot)$ defined over an algebra of n propositions, located in a world w , the *accuracy* of the agent's credences is given by:

$$1 - \left[(1/n) \sum_{\phi \in \mathcal{P}} (Cr(\phi) - w(\phi))^2 \right]$$

Here $w(\phi)$ is the truth-value of the proposition ϕ at the world w . This will be 1 if ϕ is true at w , and 0 if it is false at w .

$$(1/n) \sum_{\phi \in \mathcal{P}} (Cr(\phi) - w(\phi))^2$$

is the so-called Brier score. Among those who think that PROBABILISM is supported by the doxastic goal of accuracy, this is often taken to be the best measure of a credal state's *inaccuracy*.¹⁶ Proponents of this measure of inaccuracy will take

$$1 - \left[(1/n) \sum_{\phi \in \mathcal{P}} (Cr(\phi) - w(\phi))^2 \right]$$

to provide our measure of accuracy. Given the popularity and plausibility of this view, it is a nice case to focus on.

Assume that there are n propositions in \mathcal{P} . We can represent possible credal states as points in the space \mathbb{R}^n . A point in this space is specified by an n -tuple $\langle x_1, x_2, \dots, x_n \rangle$, such that every $x_i \in \mathbb{R}$. Pick some arbitrary bijection F from $\{x : 1 \leq x \leq n\}$ onto \mathcal{P} . We can then view the point $\langle x_1, x_2, \dots, x_n \rangle$ as representing a credal state $Cr(\cdot)$, such that $Cr(F(i)) = x_i$. That is, $\langle x_1, x_2, \dots, x_n \rangle$ represents a credal state in which the agent has credence x_i in the proposition represented by the i -th variable under the mapping F .

16. See Selten 1998 and Joyce 2009 for catalogs of the virtues of this measure. See also Leitgeb and Pettigrew 2010a, 2010b.

We can also represent possible worlds in such a space. A point $\langle x_1, x_2, \dots, x_n \rangle$ represents a possible world w just in case for every i such that $1 \leq i \leq n$, $w(F(i)) = x_i$.¹⁷ A point representing a possible world will be such that each $x_i \in \{0, 1\}$; although not every distribution of zeros and ones will necessarily represent a genuine possibility. Let's label the set of points in \mathbb{R}^n representing possible worlds \mathcal{W} .

The set of probabilistically coherent credal states can be identified as the convex hull of \mathcal{W} .¹⁸ This is the set of points in \mathbb{R}^n that can be written as weighted sums of members of \mathcal{W} , with the weightings summing to one.¹⁹ Let's label this set \mathcal{C} .

Finally, we can define the following measure on \mathbb{R}^n . Let $x = \langle x_1, x_2, \dots, x_n \rangle$, and $y = \langle y_1, y_2, \dots, y_n \rangle$. We say:

$$B(x, y) =_{df} 1 - \left[(1/n) \sum_{i=1}^n (x_i - y_i)^2 \right]$$

We're now in a position to state the arguments for PCA 1 and PCA 2. The arguments for these claims rely on the following mathematical results:²⁰

THEOREM 1 Given any point in $x \in \mathbb{R}^n - \mathcal{C}$, there is a point $y \in \mathcal{C}$, such that for every $w \in \mathcal{W}$, $B(y, w) > B(x, w)$.²¹

THEOREM 2 Given any point $x \in \mathcal{C}$, there is no point $y \in \mathbb{R}^n - \mathcal{C}$ such that (i) for every $w \in \mathcal{W}$, $B(y, w) \geq B(x, w)$, and (ii) for some $w \in \mathcal{W}$, $B(y, w) > B(x, w)$.²²

17. What I'm calling "possible worlds" are, of course, not maximally specific metaphysical possibilities. Instead they are sets of such possibilities that agree on the members of \mathcal{P} .

18. The proof of this identification is originally due to Bruno de Finetti. See de Finetti 1974. For a simple proof, see Pred et al. 2009, Proposition 1.

19. A little more pedantically: Let W be a function listing the members of \mathcal{W} , that is, a bijective function from some interval $[1, n]$ of \mathbb{N}^+ onto \mathcal{W} . The convex hull of \mathcal{W} is the set of points x such that there is some set Λ of nonnegative numbers such that $\sum_{\lambda_i \in \Lambda} \lambda_i = 1$ for which $x = \sum_{i=1}^n \lambda_i W(i)$.

20. See Pred et al. 2009 for proofs of generalized versions of these theorems.

21. For a simple proof, see, for example, Williams 2012.

22. Proof sketch: Let $x \in \mathcal{C}$ and $y \in \mathbb{R}^n - \mathcal{C}$. We can show that there is some $w \in \mathcal{W}$ such that $B(x, w) > B(y, w)$. Let $H = \{z : z \cdot (x - y) = 1/2((x \cdot x) - (y \cdot y))\}$. This is the hyperplane that runs perpendicular to the vector $x - y$ containing the point $1/2(x + y)$. Let S be the half space such that $S = \{z : z \cdot (x - y) \geq 1/2((x \cdot x) - (y \cdot y))\}$. Let \dot{S} denote the interior of S . Note that $x \in \dot{S}$, and for every $z \in \dot{S}$, $B(x, z) > B(y, z)$. Since $x \in \mathcal{C}$, there is a function W listing the members of \mathcal{W} , that is, a bijective function from some interval $[1, n]$ of \mathbb{N}^+ onto \mathcal{W} and a set Λ of nonnegative numbers such that $\sum_{\lambda_i \in \Lambda} \lambda_i = 1$,

Given THEOREM 1 and BRIER ACCURACY, it is tempting to argue for PCA 1 as follows:

- (1a) By THEOREM 1, for any point x —representing a probabilistically incoherent credal state C —there is a point y —representing a probabilistically coherent credal state C^* —such that for every possible world w , $B(y, w) > B(x, w)$.
- (1b) By BRIER ACCURACY, $B(x, w)$ is a measure of the accuracy of having the credal state represented by x in world w .
- (1c) Thus, by (1a) and (1b), it follows that for any probabilistically incoherent credal state C , there is some probabilistically coherent credal state C^* , such that, for any world w , one is more accurate if one has credence C^* in w than if one has credence C in w .

PCA 1 Thus, from (1c), it follows that for any probabilistically incoherent credal state C , there is a probabilistically coherent credal state C^* , such that C^* would be more accurate than C , no matter what the actual world is like.

Similarly, given THEOREM 2 and BRIER ACCURACY, it is tempting to argue for PCA 2 as follows:

- (2a) By THEOREM 2, for any point x —representing a probabilistically coherent credal state C^* —there is no point y —representing a probabilistically incoherent credal state C —such that (i) for every possible world w , $B(y, w) \geq B(x, w)$, and (ii) for some possible world w , $B(y, w) > B(x, w)$.
- (2b) By BRIER ACCURACY, $B(y, w)$ is a measure of the accuracy of having the credal state represented by x in world w .
- (2c) Thus, by (2a) and (2b), it follows that for any probabilistically coherent credal state C^* , there is no probabilistically incoherent credal state C , such that, (i) for any world w , one is at least as accurate if one has credal state C in w as one is if one has credal state C^* in w , and (ii) for some world w , one is more accurate if one has credal state C in w than if one has credal state C^* in w .

PCA 2 Thus, from (2c), it follows that for any probabilistically coherent credal state C^* , there is no probabilistically incoherent credal state C , such that (i) C would be at least as accurate as C^* , no matter what the actual world is like, and (ii) C would be more accurate than C^* , given at least one possible state of the world.

such that $x = \sum_{i=1}^n \lambda_i W(i)$. Since $x \cdot (x - y) > 1/2((x \cdot x) - (y \cdot y))$, it follows that $\sum_{i=1}^n \lambda_i (W(i) \cdot (x - y)) > 1/2((x \cdot x) - (y \cdot y))$. This guarantees that there is some $w_i \in \mathcal{W}$ such that $w_i \in \dot{S}$. And so there is some $w_i \in \mathcal{W}$, such that $B(x, w) > B(y, w)$.

Both of these arguments, though *prima facie* plausible, are ultimately flawed. There are two problems with each argument.

The first problem is with the inference, cited in (1b) and (2b), from BRIER ACCURACY to the claim that, in our geometric model, $B(x, w)$ measures the accuracy of having a credal state x in a world w . Although this may seem completely obvious, there are good reasons to reject this inference.²³ I'm going to bracket this worry until the next section, however, since there is a more fundamental problem. For now, then, I'll assume that (1b) and (2b) are correct.

The more fundamental problem with this argument is that PCA 1 doesn't follow from (1a)–(1b), and PCA 2 doesn't follow from (2a)–(2b). Indeed, while we may accept both (1a)–(1b) and (2a)–(2b), both PCA 1 and PCA 2 are false. On the way to showing this, I'll first show that the inference from (1c) to PCA 1 and the inference from (2c) to PCA 2 are invalid.

Recall the case of Yuko.²⁴ This case featured a proposition, namely, the proposition that (*) is true, that was true just in case a certain agent did not have credence at or above 0.5 in that proposition. Consider the smallest algebra containing the proposition expressed by (*), that is, the algebra consisting of this proposition, the negation of this proposition, the logical truth \top , and the contradiction \perp . Since \top is true no matter what, and \perp false no matter what, credal accuracy will be maximized by having credence 1 in \top and credence 0 in \perp . I'll assume that Yuko has these credences in these propositions. Yuko's possible credal states, then, will differ in what credences are assigned to the proposition expressed by (*) and to its negation.

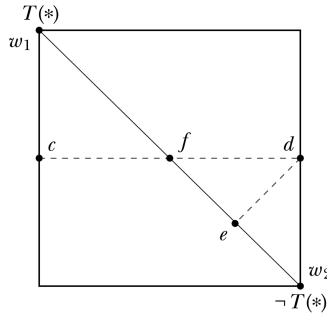
We can represent these credal states as points in \mathbb{R}^2 . We'll let x_1 represent the negation of the proposition expressed by (*), and x_2 represent the proposition expressed by (*). The point $w_1 = \langle 0, 1 \rangle$, then, represents the possible world in which the proposition expressed by (*) is true and its negation is false, while the point $w_2 = \langle 1, 0 \rangle$ represents the possible world in which this proposition is false and its negation is true.

Let's focus on the credal states in $[0, 1]^2$. We can represent these states graphically. In referring to the following graph, be sure to keep in

23. See the discussion of (3a) in the following section.

24. All of the points that follow could equally be made by appealing to the case of Hiro. I'll focus throughout, though, on the case of Yuko, leaving it to the reader to note how the same points could be made using the alternate case.

mind that $B(x, y)$ will be greater the smaller the Euclidean distance between x and y .



The probabilistically coherent states are represented by the points on the line segment between w_1 and w_2 . Consider the points $d = \langle 1, 0.5 \rangle$ and $e = \langle 0.75, 0.25 \rangle$. Point d represents a probabilistically incoherent credal state, while point e represents a probabilistically coherent credal state. e is, in fact, one of the points that accuracy-dominates d in the manner characterized by THEOREM 1. Thus we have: $\forall w B(e, w) > B(d, w)$.²⁵

In this case, however, we can see that it doesn't follow from the fact that $\forall w B(e, w) > B(d, w)$ that were Yuko to have credal state e she would be more accurate than if she were to have credal state d . The reason for this is that in this case which of w_1 or w_2 is actual depends on what Yuko's credal state is. In particular:

If Yuko were to have credal state d , then w_2 would be actual.

If Yuko were to have credal state e , then w_1 would be actual.

When asking whether Yuko would be more accurate were she to have credal state e or credal state d , the only values that we need to compare, then, are $B(d, w_2)$ and $B(e, w_1)$. And here we see: $B(d, w_2) = 0.875 > 0.4375 = B(e, w_1)$. In this case, then, despite the fact that we have $\forall w B(e, w) > B(d, w)$, it's nonetheless true that:

Yuko would be more accurate were she to have the probabilistically incoherent credal state d than were she to have the probabilistically coherent credal state e .

25. A quick calculation will verify that $B(d, w_1) = 0.375 < 0.4375 = B(e, w_1)$ and $B(d, w_2) = 0.875 < 0.9375 = B(e, w_2)$.

This case shows us that the inference from (1c) to PCA 1 isn't valid. While from $\forall w B(e, w) > B(d, w)$ we may infer (at least if we grant (1b)) that, for every world w , e is more accurate as evaluated at w than d , we can't infer from this fact that Yuko would be more accurate were she to have credal state e instead of credal state d , since which world is actual is different depending on whether Yuko has credal state d or credal state e .

This case similarly shows us that the inference from (2c) to PCA 2 isn't valid. In accordance with THEOREM 2, we have that there is no point $x \in \mathbb{R}^n - C$ such that (i) $\forall w B(x, w) \geq B(e, w)$, and (ii) $\exists w B(x, w) > B(e, w)$. We may infer from this (if we grant (2b)) that there is no probabilistically incoherent credal state x that is at least as accurate as e as evaluated at every possible world, and more accurate than e as evaluated at some possible world. But we can't infer from this fact that there is no probabilistically incoherent credal state x such that (i) no matter what the state of the world, were x to be Yuko's credal state, Yuko would be at least as accurate as she would be were her credal state to be e , and (ii) for some state of the world, were x to be Yuko's credal state, Yuko would be more accurate than she would be were her credal state to be e . For, as we've seen, Yuko would be more accurate were she to have the probabilistically incoherent credal state d than the probabilistically coherent credal state e . And this subjunctive claim holds no matter which world is actual.

Having demonstrated that the inferences from (1c) and (2c) to PCA 1 and PCA 2 are invalid, we now turn to showing that the latter claims are, in fact, false. To do this, we'll show:

The most accurate credal state that Yuko could have is represented by d .

By BRIER ACCURACY, we can measure the accuracy of a credal state $Cr(\cdot)$, located in a world w , by:

$$1 - \left[(1/n) \sum_{\phi \in \mathcal{P}} (Cr(\phi) - w(\phi))^2 \right]$$

If we think of $1 - (Cr(\phi) - w(\phi))^2$ as a measure of the accuracy of having a particular credence in the proposition ϕ at a world w , then we can think of the accuracy of a credal state as simply being the average accuracy of the credences determined by that state in particular propositions.

The first point to note is that, with respect to the proposition expressed by (*), the most accurate credence that Yuko can have is 0.5. To see this, refer back to our earlier graph. Let l be the line segment connecting points c and d . Let X be the set of points on the graph at or above l . Let Y be the set of points below l . We know the following facts:

- (8) For every point $x \in X$, were Yuko to have credal state x , then w_2 would be actual.
- (9) For every point $y \in Y$, were Yuko to have credal state y , then w_1 would be actual.

From (9), it follows that, for the members of Y , accuracy with respect to (*) increases as the value of the x_2 coordinate (that is, the vertical coordinate) increases, with this value always being $<.75$.

From (8), it follows that, for the members of X , accuracy with respect to (*) increases as the value of the x_2 coordinate decreases, with maximal accuracy being 0.75. This is reached when the x_2 coordinate is 0.5. This shows that the most accurate credence that Yuko can have in the proposition expressed by (*) is 0.5. The credal states with this property are those located on the line l .

If Yuko has a credal state located on l , then we know that w_2 is actual. In w_2 , the negation of (*) is true. Given that w_2 is actual, the most accurate that Yuko can be with respect to the negation of (*) is to have credence 1 in that proposition. Indeed, if this is the case, Yuko will be maximally accurate with respect to the negation of (*), that is, there is no other possible credal state that Yuko could have that would make her more accurate with respect to the negation of (*). Among the credal states on l , d is the only credal state in which Yuko has credence 1 in the negation of (*). This establishes the following:

d is the unique credal state that has the highest possible accuracy with respect to both the proposition expressed by (*) and its negation.

It follows that were Yuko to have some credal state other than d , she would be less accurate with respect to at least one of these propositions without there being any corresponding gain in her accuracy with respect to the other. If, for example, Yuko were to have some other credal state on l , she would be less accurate with respect to the negation of (*) without any corresponding gain in accuracy with respect to (*). And if Yuko were to have some other credal state not on l , she would be less

accurate with respect to (*) without any corresponding gain in accuracy with respect to the negation of (*).

Since the accuracy of a credal state is simply the average of the accuracy of the particular credences sanctioned by that state in particular propositions, and since credal state *d* is the unique credal state that maximizes accuracy with respect to both (*) and its negation, it follows that *d* is the most accurate credal state that Yuko could have. Were Yuko to have any other credal state, Yuko would be less accurate.

Since the probabilistically incoherent credal state represented by *d* is the most accurate credal state that Yuko could have, it follows that both PCA 1 and PCA 2 are false. Thus, the argument for PROBABILISM outlined earlier fails.

Indeed, we're now in a position to see that an appeal to the epistemic goal of credal accuracy actually motivates the rejection of PROBABILISM. For, given that one ought to try to have as accurate a credal state as one can, and given that the most accurate credal state that Yuko can have is one that is probabilistically incoherent, it follows that Yuko ought to be probabilistically incoherent.

Let me close this section by considering a possible response by a defender of PROBABILISM.

I've argued that considerations of accuracy give us good reason to hold that in certain circumstances an agent ought to have probabilistically incoherent credences. One may argue, however, that it doesn't follow from the simple fact that a credal state *C** is guaranteed to be less accurate than some other credal state *C* that *C** is therefore irrational. The reason for this is that one may hold that there are necessary truths that are a posteriori. If this is so, then, plausibly, it would not be irrational for an agent to have credence lower than 1 in such a proposition. However, if an agent has credence less than 1 in a necessary proposition, it follows that there is some other credal state that is guaranteed to be more accurate than her current credal state. Thus, we have a case in which an agent has an accuracy-dominated credal state without thereby being irrational.

If one thinks that there are necessary a posteriori truths, then what would seem to be relevant to assessing the rationality of a credal state is not whether there is some other credal state that is guaranteed to be more accurate, but whether it is a priori that there is some other credal state that is guaranteed to be more accurate.

It is a deep and difficult question whether there are necessary a posteriori propositions. Assuming that there is at least one a priori neces-

sary truth, for example, $\phi \vee \neg\phi$, this requires that there be more than one necessary proposition. This view, then, requires that propositions not be identified with sets of metaphysically possible worlds. Whether such an identification is correct or not is a controversial issue. However, despite the unclear status of necessary a posteriori propositions, it is worth considering how our arguments stand if we allow for their existence. The short answer is that nothing of importance changes.

First, let us note that even if we think (i) that it is a priori accuracy-dominance that is relevant to the rational assessment of a credal state, and (ii) that this may differ from simple accuracy-dominance, we can see, given what we've already established in this section, that this still won't help the defender of PROBABILISM. A defense of PROBABILISM, framed in terms of a priori accuracy dominance, would appeal not to PCA 1 and PCA 2, but instead to the following claims:

PCA 1* For any probabilistically incoherent credal state C , there is a probabilistically coherent credal state C^* , such that *it is a priori* that C^* would be more accurate than C , no matter what the actual world is like.

PCA 2* For any probabilistically coherent credal state C^* , there is no probabilistically incoherent credal state C , such that (i) *it is a priori* that C would be at least as accurate as C^* no matter what the actual world is like, and (ii) *it is a priori* that C would be more accurate than C^* given at least one possible state of the world.

This revised argument for PROBABILISM, however, won't fare any better than the original argument framed simply in terms of accuracy-dominance. For it follows from the fact that d is the most accurate credal state that Yuko could have that PCA 1* is false. For if PCA 1* were true, it would have to be a priori that there is some credal state that is more accurate than d , no matter what the actual world is like. But this isn't even true, let alone a priori!

Second, as I argued in section 2.1, (1) is a priori for Yuko. Moreover, it would seem to be not just a priori that (1) is true but also a priori that (1) is necessary. Given that Yuko stipulates that '(*)' refers to a certain interpreted sentence, she is in a position to know that (1) holds at all possible worlds without having to undertake any further investigation. The fact that d is the most accurate credal state that Yuko could have follows from the necessity of (1). Thus, it is also a priori knowable for Yuko

that d is the most accurate credal state that she could have.²⁶ It follows, then, that even if one thinks that considerations of accuracy have rational force only insofar as they are a priori accessible, we still have good reason to hold that Yuko ought to be probabilistically incoherent.²⁷

4. Accuracy and Decision Theory

In this section, I'll present the accuracy-dominance argument for PROBABILISM in more explicit decision-theoretic terms. Doing so helps highlight where the dominance argument for PROBABILISM goes wrong and how dominance reasoning may be used to argue against PROBABILISM.²⁸

Call a quadruple: $\mathbf{D} = \langle \mathbf{A}, \mathbf{S}, \mathbf{U}, \mathbf{C} \rangle$, a **decision problem**. Both \mathbf{A} and \mathbf{S} are sets of propositions. We call \mathbf{A} the set of **acts** and \mathbf{S} the set of **states**. Think of the members of \mathbf{A} as propositions describing various acts that an agent may undertake.²⁹ Think of the members of \mathbf{S} as propositions describing various ways the world might be that are relevant to the outcomes that would obtain were the acts described by the members of \mathbf{A} to be performed. We assume that both \mathbf{S} and \mathbf{A} form partitions of the space of possible worlds. \mathbf{U} is a **utility function** that assigns to propositions of the form $A_i \wedge S_j$ a number that measures the utility that would result for the agent were the act described by A_i to be performed in state S_j . Finally, \mathbf{C} is a credence function that is defined on an algebra containing all propositions of the form $A_i \wedge S_j$.³⁰

26. I'm assuming that the logical consequences of a priori truths are a priori. Note that this shows that PCA 2* is also false.

27. It is worth noting that either version of the accuracy-dominance argument tells against Egan and Elga's claim that it is irrational for an agent to self-ascribe anti-expertise. For since (1) is necessary, Yuko maximizes accuracy by having credence (1) in this proposition, and since (1) is a priori, it follows that it is a priori that having credence 1 in (1) maximizes accuracy.

28. See Pettigrew 2011a, 2011b, for a helpful survey of some of the uses of decision-theoretic machinery in epistemology.

29. We'll be somewhat promiscuous with what we consider an act. In particular, we'll count an agent's coming to have a particular credal state as an act. This shouldn't, though, be seen as an endorsement of a questionable doxastic voluntarism.

30. As we'll see, there are further constraints that we will want to put on decision problems. For now, however, it is useful to simply think about decision problems as having this minimal structure.

Given a decision problem \mathbf{D} , we say:

An act A_1 **strongly dominates** an act A_2 (in \mathbf{D}) just in case for every $S_i \in \mathbf{S}$, $\mathbf{U}(A_1 \wedge S_i) > \mathbf{U}(A_2 \wedge S_i)$.

An act A_1 **weakly dominates** an act A_2 (in \mathbf{D}) just in case (i) for every $S_i \in \mathbf{S}$, $\mathbf{U}(A_1 \wedge S_i) \geq \mathbf{U}(A_2 \wedge S_i)$, and (ii) for some $S_i \in \mathbf{S}$, $\mathbf{U}(A_1 \wedge S_i) > \mathbf{U}(A_2 \wedge S_i)$.

Let \mathcal{A} and \mathcal{B} be sets of acts. We say:

\mathcal{A} **act-dominates** \mathcal{B} just in case (i) for every $B \in \mathcal{B}$, there is some $A \in \mathcal{A}$ such that A strongly dominates B , and (ii) there is no $B \in \mathcal{B}$ such that, for some $A \in \mathcal{A}$, B weakly dominates A .

Here are two putative norms that we might appeal to, given a decision problem, to single out a certain option or set of options as rationally obligatory.³¹

DOMINANCE 1 If A_i strongly dominates all other members of \mathbf{A} , then A_i is rationally required.

DOMINANCE 2 If \mathcal{A} and \mathcal{B} partition \mathbf{A} , and \mathcal{A} act-dominates \mathcal{B} , then it is rationally required that one choose some option in \mathcal{A} .

We can now present Joyce’s argument for **PROBABILISM** using this decision theoretic framework. We can represent an agent’s epistemic situation as a decision problem \mathbf{D} . We let \mathbf{A} , the set of “acts” available to an agent, be the set of propositions describing possible credal states, given an algebra \mathcal{P} , that a particular agent could have. We let \mathbf{S} , the set of states, be the set of propositions describing possible distributions of truth-values for the members of \mathcal{P} .³² We let \mathbf{U} be a measure of the agent’s epistemic utility, which we take to be measured by the accuracy of an agent’s credal state given a particular distribution of truth-values. We, then, argue as follows:

31. Note that it’s crucial that for **DOMINANCE 1**, we assume that A_i strongly dominates *every other member* of \mathbf{A} , and that for **DOMINANCE 2**, we assume that \mathcal{A} and \mathcal{B} partition \mathbf{A} .

32. Note that if, for the reasons canvassed in the previous section, one thinks that it is a priori accuracy-dominance that is relevant to the rational assessment of a credal state, then one can think of the states not as metaphysically possible truth-value distributions, but instead as epistemically possible truth-value distributions, that is, truth-value distributions that could obtain so far as one knows a priori. Given that (1) is a priori for Yuko, (1) will hold in all states so construed. And, given that this holds, the arguments that follow will go through unaffected.

- (3a) If A is the credal state represented by x , and S the state represented by w , then, by BRIER ACCURACY, $\mathbf{U}(A \wedge S) = B(x, w)$.
- (3b) By (3a), THEOREM 1, and THEOREM 2, it follows that, relative to \mathbf{D} , the set of probabilistically incoherent credal states are act-dominated by the set of probabilistically coherent credal states.
- (3c) By (3b) and DOMINANCE 2, it follows that an agent is rationally required to have a probabilistically coherent credal state.

As with the argument in the previous section, we can locate two problems with this argument for PROBABILISM. Again, it will help in getting clear on where the argument breaks down to focus on the case of Yuko.

In accordance with the above argument, we can think of Yuko as facing the following epistemic decision problem, \mathbf{D}_Y^1 . We let \mathbf{A}_Y^1 be the set of propositions describing possible credences that Yuko could have in the proposition expressed by (*) and in its negation.³³ We let \mathbf{S}_Y^1 be the set of propositions describing possible distributions of truth-values for these propositions. \mathbf{S}_Y^1 will, of course, have two members: S_1 , in which the proposition expressed by (*) is true and its negation is false, and S_2 , in which these truth-values are reversed. Finally, we assume that $\mathbf{U}_Y^1(A_i \wedge S_i) = B(x, w)$, where x is the point in \mathbb{R}^2 representing A_i and w is the point representing S_i .

The first problem in the above argument is with (3a).³⁴ Grant BRIER ACCURACY. That is, grant that, given an agent with credences $Cr(\cdot)$, located in a world w , the accuracy of the agent's credences is given by:

$$1 - \left[(1/n) \sum_{\phi \in \mathcal{P}} (Cr(\phi) - w(\phi))^2 \right]$$

Still, it doesn't follow that, in the decision problem at hand, if x represents credal state A , and w represents a world state S , the epistemic utility of $A \wedge S$ is given by $B(x, w)$.

The reason for this is that in this decision problem not all conjunctions of the form $A \wedge S$ describe genuine possibilities, that is, possible situations in which Yuko has credal state A in state S . For example, let

33. We'll continue to assume that Yuko has credence 1 in \top and credence 0 in \perp .

34. The problem that arises here is the same problem that arises with premises (1b) and (2b) in the argument in the previous section. I earlier noted that there is a problem with these premises but deferred in-depth discussion. The points that follow should make it clear why appeal to (1b) and (2b) in our earlier argument is problematic.

A_e be the proposition that Yuko has the credal state represented by point e in our earlier graph, and let S_2 be the state represented by point w_2 . We know that the conjunction $A_e \wedge S_2$ is impossible. Of course, we can assign a number to this conjunction by using the measure B defined on \mathbb{R}^2 . But this number does not represent the epistemic utility of the possible situation in which Yuko has the credal state represented by A_e in state S_2 ; for there simply is no possible situation in which Yuko has this credal state and S_2 obtains.

One way of bringing out the problem here is to note that if we were to say that $B(x, w)$ always measures the epistemic utility of $A \wedge S$ (where A is the credal state represented by x and S the state represented by w), then we would be committed to inconsistent assignments of epistemic utility to sets of possible worlds. To see this, let A_d be the proposition that Yuko has the credal state represented by point d in our earlier graph and let S_1 be the state represented by w_1 . Since both $A_e \wedge S_2$ and $A_d \wedge S_1$ are impossible, they both describe the same set of possible worlds, namely, the null set. But it's easy to verify that $B(e, w_2) \neq B(d, w_1)$. Even if we could make sense of an assignment of utilities to the null set of worlds (and I doubt we can), we should surely want to hold that this utility is unique. Taking $B(x, w)$ to measure epistemic utility wouldn't allow for this.

We can draw a lesson from this first problem:

If we want to model an agent's epistemic position as a decision problem, we should make sure that we choose our states so that they are compatible with each of the agent's possible credal states.

In a moment we'll see how to do this, but first let's look at the second problem with the above argument.

The second problem can be located in the appeal to DOMINANCE 2. Dominance reasoning is certainly plausible. After all, if some option (or set of options) is better than the alternatives *no matter what the world is like*, how could it not be better *tout court*? It is well known, however, that one needs to be careful in how one sets up a decision problem if dominance reasoning is not to lead us astray.³⁵ Consider the following situation:

35. See Jeffrey 1983 and Joyce 1999 for discussion of some ways in which dominance reasoning may fail.

Bounty A large sum of money has been stolen from a local crime boss, and you've been framed. There's a bounty on your head paying an exorbitant sum of money in return for your death. You can either flee to the mountains or stay home. If you stay at home, you're very likely to be shot, and you know this. If you flee, though, there's a decent chance you'll escape alive, and you know this. You would, however, prefer to live at your house than in the mountains. You'd also prefer, somewhat, to be killed at home than to be killed in the mountains. Of course, you strongly prefer living to dying (whether in the mountains or at home). What should you do?

We might represent this situation using the following decision problem that we'll label \mathbf{D}_B . In \mathbf{D}_B there are two acts available to you: staying home, and fleeing to the mountains. And there are two possible states: in one state you are killed, in another state you live. The utilities can be represented by the following matrix:

	Die	Live
Stay	1	5
Flee	0	3

Applying either DOMINANCE 1 or DOMINANCE 2 to this decision problem yields the verdict that the rational thing to do is to stay at home.

But this is clearly wrong. If you stay at home, then you're almost certain to be killed, while if you flee, there's a good chance you may escape with your life; and you know these facts. Since you'd prefer to live rather than die, you should flee.

Pretty much everyone agrees that in this type of case, dominance reasoning leads us astray. It turns out, however, to be a matter of some controversy exactly why this reasoning fails. Everyone agrees that in order to apply dominance reasoning to a decision problem, the acts and states must in some sense be **independent**. However, there is disagreement about exactly what this condition of act-state independence amounts to.

According to **evidential decision theorists**, in order to apply dominance reasoning to a decision problem, the acts and states must be *probabilistically independent*.³⁶ That is, for each act A and state S , we must have: $\mathbf{C}(S | A) = \mathbf{C}(S)$.

36. See Jeffrey 1983 for the canonical development of evidential decision theory.

According to **causal decision theorists**, in order to apply dominance reasoning to a decision problem, the acts and states must be *causally independent*.³⁷ That is, for each act A and state S , A must neither causally promote nor hinder S .

In the decision problem we've used to model **Bounty**, our acts and states are neither probabilistically nor causally independent. Given this fact, causal and evidential decision theorists will agree that dominance reasoning shouldn't be sanctioned in this case.

In the case of \mathbf{D}_Y^1 , the decision problem we've used to model Yuko's epistemic situation, one should certainly reject the appeal to dominance reasoning if one is a causal decision theorist. For it's clear that the acts and states in this decision problem are not causally independent. Recall that the possible states are truth-value distributions for the proposition expressed by (*) and its negation, while the acts are possible credence distributions in these propositions. Since which state is actual depends on what Yuko's credence in (*) is, Yuko's acts will causally influence which state obtains. It follows that if one is a causal decision theorist, then one should reject the appeal to DOMINANCE 2 in (3c).

In the case of evidential decision theory, matters are a bit more subtle since, in order to know whether we can apply dominance reasoning, we need to make some assumptions about the agent's credal state. We can show, however, that in a large class of reasonable cases, the appeal to dominance reasoning will be illicit by the lights of evidential decision theory. And the reason for this is that there are a large number of reasonable credal states that Yuko could have that would make the acts and states in \mathbf{D}_Y^1 probabilistically dependent.

For example, assume that Yuko is aware of the way in which the state of the world is dependent on her credal state. In particular, assume that Yuko's credences are such that: $Cr_y(S_1 | A_e) = 1$ and $Cr_y(S_1 | A_d) = 0$. Since Yuko can't have both credence 1 and credence 0 in S_1 , it follows that the acts and states in this decision problem will not be probabilistically independent. In such cases, then, if one is an evidential decision theorist, one should reject the appeal to DOMINANCE 2 in (3c).

37. For developments of causal decision theories, see, for example, Joyce 1999 and Lewis 1981.

The lesson to be drawn here is the following:

If we want to model an agent's epistemic position as a decision problem and apply dominance reasoning, we should choose our states so that they are independent of the agent's possible credal states.

There is a simple way of reformulating the decision problem representing Yuko's epistemic situation that lets us address both of the defects in the preceding argument. Instead of representing our states as possible distributions of truth-values for the proposition expressed by (*) and its negation, as we did in \mathbf{D}_Y^1 , we should instead take our states to be **dependence hypotheses**.

A dependence hypothesis is a proposition that states for each possible act, what utility the agent would gain were that act to be performed. Let $[U = u]$ be a proposition specifying that an agent's utility is u . We may think of a dependence hypothesis as a (possibly infinite) conjunction of nonbacktracking counterfactuals of the form $A_i \square \rightarrow [U = u]$, containing exactly one conjunct for each act A_i .

In \mathbf{D}_Y^1 there were two states representing the two possible distributions of truth-values for (*) and its negation. If, instead, we carve up the space of possible worlds by grouping together worlds that make true the same counterfactuals connecting credal states and epistemic utilities, then there will only be one state in our decision problem. For both of our possible worlds w_1 and w_2 agree about which world would be actual were Yuko to have a particular credal state. For example, both w_1 and w_2 agree that were Yuko to have credal state e , w_1 would be actual. Thus both w_1 and w_2 will agree about what epistemic value Yuko would have were she to have a particular credal state.

Instead of representing Yuko's epistemic situation by the decision problem \mathbf{D}_Y^1 , we should represent it by the following alternative decision problem \mathbf{D}_Y^2 . Let S_y^2 be the dependence hypothesis specifying how Yuko's epistemic utility counterfactually depends on her credal state. We let \mathbf{S}_Y^2 be the singleton set consisting of S_y^2 . We let \mathbf{A}_Y^2 be the set of possible credence distributions that Yuko can have in the proposition expressed by (*) and its negation. Finally, we let $\mathbf{U}_Y^2(A \wedge S_y^2) = u \leftrightarrow (S_y^2 \vDash A \square \rightarrow [U = u])$, that is, just in case $A \square \rightarrow [U = u]$ is one of the conjuncts of S_y^2 .

Note that the members of \mathbf{A}_Y^2 are all compatible with S_y^2 and are all causally independent of S_y^2 . This is guaranteed since S_y^2 is true in every possible world.

Moreover, S_y^2 will be probabilistically independent of each member of \mathbf{A} , given the assumption that Yuko gives credence 1 to S_y^2 both unconditionally and conditional on each member of \mathbf{A} . Given that S_y^2 is true in every possible world, it's reasonable to assume that Yuko is not rationally precluded from having a credal state that satisfies this constraint. We'll assume that Yuko's credal state does satisfy this constraint.

Since the acts and states in \mathbf{D}_y^2 are independent (in either of the relevant senses), we can apply dominance reasoning to this decision problem to draw conclusions about what sort of credal state Yuko ought to have. And what dominance reasoning tells us here is that Yuko ought to be probabilistically incoherent.

To see this, recall that in the previous section we showed that the credal state represented by d maximizes accuracy in the following sense: were Yuko to have any other credal state, she would be less accurate than she would be if she were to have the credal state represented by d . There is, then, a nonbacktracking counterfactual $A_d \square \rightarrow [U = u]$, such that $S_y^2 \models A_d \square \rightarrow [U = u]$, and such that for any other $A \in \mathbf{A}_y^2$, if $S_y^2 \models A \square \rightarrow [U = u^*]$, then $u^* < u$. It follows that for every $A \in \mathbf{A}_y^2$ such that $A \neq A_d$, $\mathbf{U}_y^2(A_d \wedge S_y^2) > \mathbf{U}_y^2(A \wedge S_y^2)$. By DOMINANCE 1, then, it follows that Yuko ought to have the credal state represented by A_d . And since A_d is probabilistically incoherent, it follows that Yuko ought to be probabilistically incoherent.

5. Rational Probabilistic Incoherence Extended

I've argued that there are cases in which an agent ought to have probabilistically incoherent credences—at least assuming that credal accuracy is our primary epistemic goal. Instead of supporting PROBABILISM, considerations of credal accuracy give us good reason to reject this principle.

The case that we focused on, however, is unusual. This case involves a proposition ϕ such that necessarily ϕ is true just in case a certain agent's credence in ϕ is less than a particular value. Most propositions aren't like this. Most propositions are such that their truth-values aren't tied in this way to what our credences are in those propositions. This raises the question whether there is an interesting restricted version of PROBABILISM that we may still endorse. Is there some large algebra \mathcal{P} such that, if we restrict our attention to \mathcal{P} , it is true that a rational agent must always have probabilistically coherent credences in those propositions?

I'll argue that the answer to this question is no. For almost any proposition ϕ , there is some possible situation in which an agent's credences in ϕ and $\neg\phi$ may be rationally probabilistically incoherent.

We've seen that an agent's epistemic situation with respect to an algebra of propositions \mathcal{P} can be represented as a decision problem \mathbf{D} in which the set of states \mathbf{S} consists of maximal specifications of how the agent's epistemic utility counterfactually depends on her credences in the members of \mathcal{P} . Given such a decision problem, we can appeal to principles of rational decision making, such as DOMINANCE 1 and DOMINANCE 2, to argue that certain credal states, or sets of credal states, are rationally obligatory. Dominance reasoning, however, can be only applied to a limited range of decision problems. Where there is no act that dominates all its competitors, DOMINANCE 1 falls silent. Where there is no set of acts that act-dominates all its competitors, DOMINANCE 2 falls silent. To determine the rational act(s) in these cases, we need a more general principle of rational decision making.

Call a decision problem **proper** if \mathbf{S} is a set of dependence hypotheses, and \mathbf{C} is a credence function that is probabilistically coherent over the smallest algebra containing every proposition of the form $A_i \wedge S_j$. Given a proper decision problem \mathbf{D} , we can define the **causal expected utility** of an act A , $U_C(A)$, as follows:

$$U_C(A) =_{df} \sum_{S \in \mathbf{S}} \mathbf{C}(S)U(A \wedge S)$$

We can further define the **evidential expected utility** of an act A , $U_E(A)$, as follows:

$$U_E(A) =_{df} \sum_{S \in \mathbf{S}} \mathbf{C}(S|A)U(A \wedge S)$$

We can now formulate two additional putative principles of rational decision making:

- CUP** If \mathbf{C} is a rational credal state and $U_C(A_1) > U_C(A_2)$, then A_1 is rationally preferable to A_2 .
- EUP** If \mathbf{C} is a rational credal state and $U_E(A_1) > U_E(A_2)$, then A_1 is rationally preferable to A_2 .

Causal decision theorists will endorse CUP, while evidential decision theorists will endorse EUP. Since evidential and causal expected utilities can come apart, CUP and EUP will sometimes give contradictory

verdicts. In what follows, however, we can remain neutral on the question of which of these two principles we should endorse.

I'll now argue, by appeal to CUP and EUP, that even in cases in which there is no necessary connection between the truth of a proposition and an agent's credence in that proposition, an agent may be rationally probabilistically incoherent. Even for such propositions, there are cases in which, by a rational agent's own lights, accuracy is better achieved by being probabilistically incoherent than probabilistically coherent.

Consider the proposition that Yuko will make a particular free throw. Name this proposition **FT**. Clearly, there is no necessary connection between the truth-value of FT and Yuko's credence in this proposition.

Consider the smallest algebra containing FT. This consists of \top , \perp , FT, and \neg FT. We'll assume that Yuko has credence 1 in \top and credence 0 in \perp . Yuko's possible credal states, then, differ with respect to this algebra just over the credences assigned to FT and \neg FT.

We may think of Yuko's epistemic situation regarding this class of propositions as a decision problem \mathbf{D}_Y^3 . The set of acts \mathbf{A}_Y^3 is the set of possible credences that Yuko could have in FT and \neg FT. Unlike with \mathbf{D}_Y^2 , however, there will be more than one dependence hypothesis stating how Yuko's epistemic utility depends counterfactually on members of \mathbf{A}_Y^3 . The reason for this is that, unlike with (*), there is no necessary connection between the truth-value of FT and Yuko's credence in FT.

Assume that as a matter of contingent fact Yuko is an extremely accurate free-throw shooter but only when her credence is less than 0.5 that she will make the free throw. In particular, assume that the following counterfactual claims hold:

- (10) If Yuko were to have credence less than 0.5 in FT, then FT would be true. $[(Cr_Y(\text{FT}) < 0.5) \Box \rightarrow \text{FT}]$
- (11) If Yuko were to have credence greater than or equal to 0.5 in FT, then FT would be false. $[(Cr_Y(\text{FT}) \geq 0.5) \Box \rightarrow \neg \text{FT}]$

It's easy to see that (10) and (11) together entail, for any $A \in \mathbf{A}_Y^3$, a counterfactual: $A \Box \rightarrow [U = u]$, where $[U = u]$ specifies the agent's epistemic utility given the credences specified in A .³⁸ (10) and (11), thus, jointly

38. Justification: For every $A \in \mathbf{A}_Y^3$, either $A \models Cr_Y(\text{FT}) < 0.5$ or $A \models Cr_Y(\text{FT}) \geq 0.5$. This ensures that $(Cr_Y(\text{FT}) < 0.5) \Box \rightarrow \text{FT}$ and $(Cr_Y(\text{FT}) \geq 0.5) \Box \rightarrow \neg \text{FT}$ jointly entail, for every $A \in \mathbf{A}_Y^3$, either $A \Box \rightarrow \text{FT}$ or $A \Box \rightarrow \neg \text{FT}$. The accuracy of members of \mathbf{A}_Y^3 is deter-

entail a particular dependence hypothesis and are jointly incompatible with all other dependence hypotheses. Since they both obtain, they settle which dependence hypothesis obtains. Call this dependence hypothesis S_y^3 . According to S_y^3 , Yuko's epistemic utility depends on her credences in FT and its negation in exactly the same manner that her epistemic utility depends on her credences in (*) and its negation according to S_y^2 .

To show that Yuko may rationally fail to have probabilistically coherent credences in FT and its negation, we make the following assumptions:

- (12) Yuko's credence in S_y^3 is 0.9.
- (13) Yuko's credal state is probabilistically coherent over the smallest algebra containing every proposition of the form $A_i \wedge S_j$, for $A_i \in \mathbf{A}_Y^3$ and $S_j \in \mathbf{S}_Y^3$.
- (14) Yuko's credences are such that the members of \mathbf{A}_Y^3 and the members of \mathbf{S}_Y^3 are probabilistically independent.
- (15) The credal profile ascribed in (12)–(14) is rational.

There are numerous credal states that satisfy (12)–(14). Shortly, I will discuss the further assumption that this credal profile may be rational. First, however, I'll show that from (12)–(15), it follows by either CUP or EUP that Yuko may rationally fail to have probabilistically coherent credences in FT and \neg FT.

Let A_d be the member of \mathbf{A}_Y^3 according to which Yuko has credence 0.5 in FT and credence 1 in its negation. Let \mathbf{A}_P be the set of probabilistically coherent credences in FT and its negation. We first show:

From (12)–(13), it follows that, for every $A \in \mathbf{A}_P$, $U_C(A_d) > U_C(A)$.

We assume that \mathbf{C} is a probabilistically coherent credence distribution over the smallest algebra containing every proposition of the form $A_i \wedge S_j$, for $A_i \in \mathbf{A}_Y^3$ and $S_j \in \mathbf{S}_Y^3$.

The epistemic utility of having the credal state represented by A_d in state S_y^3 is 0.875. This gives us a lower bound on the causal expected utility of A_d . If $C(S_y^3) = x$, then $U_C(A_d) \geq x(0.875)$.

mined solely by whether or not FT is true or false. We have, then, for every $A \in \mathbf{A}_Y^3$, $A \wedge \text{FT} \models [U = u]$ and $A \wedge \neg \text{FT} \models [U = u^*]$, for some propositions $[U = u]$ and $[U = u^*]$. It follows, then, from the fact that $(Cr_y(\text{FT}) < 0.5) \Box \rightarrow \text{FT}$ and $(Cr_y(\text{FT}) \geq 0.5) \Box \rightarrow \neg \text{FT}$ entail, for every $A \in \mathbf{A}_Y^3$, either $A \Box \rightarrow \text{FT}$ or $A \Box \rightarrow \neg \text{FT}$, that they entail, for every $A \in \mathbf{A}_Y^3$, some proposition of the form $A \Box \rightarrow [U = u]$.

Among the probabilistically coherent credal states, the most accurate state in S_y^3 will be the state represented by point f in our geometric model. If Yuko has this credal state, she will have credence 0.5 in FT and credence 0.5 in its negation. Let A_f be the proposition according to which Yuko has this credal profile. The epistemic utility of having the credal state represented by A_f in state S_y^3 is 0.75. This gives us an upper bound on the causal expected utility for $A \in \mathbf{A}_P$. If $C(S_y^3) = x$, then for any $A \in \mathbf{A}_P$, $U_C(A) \leq x(0.75) + (1 - x)1$.

Given this lower bound on the expected utility of A_d and this upper bound on the expected utility of members of \mathbf{A}_P , we can show that there are values of x , such that if $C(S_y^3) = x$, then for every $A \in \mathbf{A}_P$, $U_C(A_d) > U_C(A)$.

To show this, we first calculate the value for x at which the lower bound for A_d equals the upper bound for members of \mathbf{A}_P . To do this, we set $x(0.875) = x(0.75) + (1 - x)1$, and solve for x . A quick calculation shows that this equality holds when $x = (1/1.125) \approx 0.89$. It follows that whenever $x > 0.89$, the lower bound for A_d will be greater than the upper bound for the members of \mathbf{A}_P . Thus, if $C(S_y^3) > 0.89$, then for every $A \in \mathbf{A}_P$, $U_C(A_d) > U_C(A)$.

From (12)–(13), it therefore follows that for every $A \in \mathbf{A}_P$, $U_C(A_d) > U_C(A)$. Given this fact and our assumption that the credal profile ascribed by (12)–(13) is rational, it follows from CUP that A_d is rationally preferable to every $A \in \mathbf{A}_P$. If one endorses CUP, then one should hold that Yuko is not rationally required to have probabilistically coherent credences.

From (14), it follows that for every $A \in \mathbf{A}$, $U_E(A) = U_C(A)$. Thus, we have that $U_E(A_d) > U_E(A)$ for every probabilistically coherent credal state $A \in \mathbf{A}_P$. Given this fact and (15), it follows from EUP that A_d is rationally preferable to every $A \in \mathbf{A}_P$. If one endorses EUP, then one should hold that Yuko is not rationally required to have probabilistically coherent credences.

Nothing in this argument turns on any features of FT that couldn't, at least in principle, be shared by almost any other contingent proposition. Assuming that this argument works, we have, then, a fairly general recipe for generating cases in which an agent may rationally fail to have probabilistically coherent credences in a contingent proposition and its negation.

The one substantive assumption that a proponent of CUP or EUP may question in order to block this argument is (15), that is, the assumption that the credal profile imposed by (12) – (14) is rational. If one rejects this claim, then one must hold that it is impossible for Yuko to have a credal state satisfying (12) – (14) without being guilty of a rational failure. Why would this be?

There are three options:

- One could hold that Yuko would be irrational in virtue of having a credal state that is probabilistically coherent over the smallest algebra containing every proposition of the form $A_i \wedge S_j$ for $A_i \in \mathbf{A}_Y^3$ and $S_j \in \mathbf{S}_Y^3$.
- One could hold that Yuko would be irrational in virtue of such a credence distribution making the members of \mathbf{A}_Y^3 and \mathbf{S}_Y^3 probabilistically independent.
- One could hold that Yuko would be irrational in virtue of having a high credence in S_y^3 .

I can't see any even remotely plausible reasons for endorsing either of the first two options. I'll put these claims, then, to the side and move on to consideration of the third option.

If Yuko were to have a high credence in S_y^3 , then she would have a high credence that no matter what her credence was in FT, it was bound to be far from the actual truth-value, that is, she would have high credence in her own anti-expertise about FT. One may, therefore, try to justify the claim that it is irrational for Yuko to have a high credence in S_y^3 by appealing to the premise that it is *in general* irrational for an agent to have a high credence in her own anti-expertise.

As we saw in section 2, this general claim is endorsed by Andy Egan and Adam Elga. However, I argued there that there is little reason to accept this general claim. In section 2.1, I showed that there are cases where it is an obvious necessary truth that an agent is an anti-expert about a certain proposition. In such cases, it is extremely plausible that an agent may rationally have a high credence in her own anti-expertise. And in section 2.2, I argued that the premises from which Egan and Elga infer the general claim that it is irrational to self-ascribe anti-expertise are jointly unacceptable. Given these considerations, we shouldn't endorse this general claim.

Of course, one may still hold that in this particular case it is irrational for Yuko to have a high credence in her own anti-expertise. This seems to me, however, to be poorly motivated. S_y^3 is a truth, and there

is no reason that Yuko could not have sufficient evidence to make this truth highly likely. If she were to be in possession of such evidence, what would be the rational response? In section 2, I noted that one possible response—that recommended by Egan and Elga—to evidence of anti-expertise is for an agent to adjust her credences so that she is no longer an anti-expert about the relevant proposition. However, I noted that this is only plausibly the rational response in cases in which an agent can in fact make it false that she is an anti-expert by adopting some particular credence. In the case of (1), for example, this was not possible, and so it seemed that this could not be the correct response to Yuko's evidence of her own anti-expertise concerning (*). Importantly, the same is true in this case as well. Given the truth of S_y^3 , Yuko is guaranteed to be an anti-expert about FT no matter what her credence is in this proposition. So, since Yuko cannot adopt credences that will make it false that she is an anti-expert about FT, it seems that were she to be in possession of strong evidence in favor of S_y^3 , the rational option for her—as for any other individual in possession of such evidence—would be to have an appropriately high credence in this proposition.

Although one could try to block the argument for the claim that Yuko may rationally fail to have probabilistically coherent credences in FT and its negation by holding that it is impossible for Yuko to have a credal state satisfying (12)–(14) without thereby being guilty of a rational failure, the prospects for this option, on inspection, look quite dim.

Even for a contingent proposition such as FT, then, it would seem that there can be situations in which, by a rational agent's own lights, accuracy isn't maximized by having probabilistically coherent credences. Assuming that an agent ought to try to make her credences as accurate as possible, in such cases an agent may rationally fail to have probabilistically coherent credences. **PROBABILISM**, then, doesn't just fail when we look at propositions such as that expressed by (*) that concern an agent's own credences. In principle, almost any proposition could be such that an agent could rationally fail to have credences in that proposition and its negation that sum to 1.

6. Dutch Books

I've argued that considerations of accuracy, instead of supporting **PROBABILISM**, in fact support the claim that an agent may in certain circumstances rationally fail to have probabilistically coherent credences. Since

I'm inclined to agree with Joyce that credal accuracy is our primary epistemic goal, this gives us, I think, good reason to reject **PROBABILISM**.

There are, however, other arguments that have been offered for **PROBABILISM** that make no appeal to considerations of accuracy. The most famous of these is the so-called Dutch Book argument.³⁹ At first glance, it may seem that Dutch Book considerations tell against my claim that it is rational for Yuko to have credence 0.5 in $T(*)$ and credence 1 in $\neg T(*)$. I'll argue in this section, however, that on closer inspection, this proves not to be the case. As with the accuracy-dominance argument, the case of Yuko can be used to expose crucial flaws in the Dutch Book argument for **PROBABILISM**. While I'm inclined to think that the sorts of considerations adduced in Dutch Book arguments have limited probative force for claims about epistemic rationality, still, those who do find such considerations persuasive do not have reason to reject my claim that it may be rational for an agent to be probabilistically incoherent. Indeed, I'll argue that those who are persuaded of **PROBABILISM** by Dutch Book considerations in fact have good reason to endorse the claim that probabilistic incoherence can be rational.

The Dutch Book argument for **PROBABILISM** can be thought of as resting on the following three claims:

SURE-LOSS AVOIDANCE An agent ought to have credences that do not lead the agent to value as fair a set of bets that guarantee a loss of money.

DUTCH BOOK If an agent is probabilistically incoherent, then the agent will value as fair a set of bets that will guarantee her a loss of money.

CONVERSE DUTCH BOOK An agent will value as fair a set of bets that will guarantee her a loss of money only if she is probabilistically incoherent.

From **DUTCH BOOK** and **CONVERSE DUTCH BOOK**, it follows that an agent will value as fair a sure-loss set of bets just in case the agent fails to have probabilistically coherent credences.⁴⁰ Thus, given these two claims,

39. The Dutch Book argument for **PROBABILISM** was first presented in Ramsey 1931.

40. A worry about **CONVERSE DUTCH BOOK** arises if one thinks that there are a posteriori necessities. In this case, unless one reads 'logical truth' in the principle **NORMALIZATION** as meaning 'necessary truth', it would seem that one could be probabilistically coherent but assign a credence lower than 1 to such a proposition and so be

it follows from SURE-LOSS AVOIDANCE that an agent ought to have probabilistically coherent credences.

There are a number of worries that one may have with this argument.⁴¹ Perhaps the most natural worry is the following. PROBABILISM is a principle of *epistemic* rationality. If this argument is to succeed, then, it would need to be the case that the notion of obligation involved in SURE-LOSS AVOIDANCE is epistemic in nature. However, SURE-LOSS AVOIDANCE is most naturally read as a principle of *practical* rationality. There is, then, a nontrivial challenge to justify the principle SURE-LOSS AVOIDANCE when this is understood in epistemic terms. We can, however, bracket this worry. For even if we ignore this worry, we can show that the Dutch Book argument provides us with little reason to endorse PROBABILISM.

Let an agent have credence p in ϕ . The Dutch Book argument assumes that an agent will value as fair a bet that costs $p \times \$n$ and pays $\$n$ in the event ϕ and pays $\$0$ otherwise. To say that the agent values such a bet as fair is to say that the agent sees no advantage in either buying or selling this bet. Let's grant this assumption.⁴² It is easy to show, then, that if Yuko has credence 0.5 in $T(*)$ and credence 1 in $\neg T(*)$, she will value as fair the following sure-loss set of bets:

BET 1 Yuko pays $\$0.5$. She receives $\$1$ if $T(*)$ and 0 otherwise.

BET 2 Yuko pays $\$1$. She receives $\$1$ if $\neg T(*)$, 0 otherwise.

BET 1 and BET 2 together guarantee a loss of $\$0.5$. For if Yuko buys both of these bets, she will have paid out a combined $\$1.50$, but she will only receive back $\$1$. However, Yuko will value BET 1 as fair given that she has credence 0.5 in $T(*)$, and she will value BET 2 as fair given that she has credence 1 in $\neg T(*)$. Thus, if Yuko has credence 0.5 in $T(*)$ and credence

subject to a sure-loss bet. If one takes seriously the possibility of a posteriori necessities, then one should read 'guaranteed loss of money' as 'a priori guaranteed loss of money' in the above claims. I note that given that (1) is a priori for Yuko, the arguments that follow can all be applied mutatis mutandis to this alternate interpretation of the Dutch Book argument. I leave the verification of this, though, to the interested reader.

41. See Hájek 2008 for references to the vast literature on Dutch Book arguments. For an expression of the following worry, see Joyce 1998.

42. This is only plausible assuming that the agent's assignment of utility to gains of money is linear. While this may not be psychologically plausible, it is harmless for our purposes. In principle, we could trade in talk of money for talk of utility. Having noted this, I'll continue to work with bets framed in monetary terms.

1 in $\neg T(*)$, then she will value as fair a set of bets that guarantee her a loss of money.

A proponent of the Dutch Book argument for PROBABILISM may point to this fact as evidence that Yuko's credences in $T(*)$ and $\neg T(*)$ are irrational, despite my earlier argument to the contrary. To conclude this, however, would be hasty. For we can show:

No matter what credence Yuko has in $T(*)$, she will value as fair a bet that guarantees her a loss of money.

First, assume that $Cr_y T(*) < 0.5$. Yuko, then, will value the following bet as fair:

BET 3 Yuko sells a ticket for $Cr_y T(*) \times \$1$. She pays \$1 if $T(*)$ and \$0 otherwise.

This bet, however, guarantees Yuko a loss. For given that $Cr_y T(*) < 0.5$, it follows that $T(*)$ holds. And so given that $Cr_y T(*) < 0.5$, Yuko will be forced to pay out \$1. And since $(Cr_y T(*) \times \$1) < \1 , Yuko will be guaranteed to pay out more than she receives in selling the bet. If $Cr_y T(*) < 0.5$, it follows that Yuko will value a bet as fair, despite the fact that the bet guarantees her a loss.

Next assume that $Cr_y T(*) \geq 0.5$. Yuko, then, will value the following bet as fair:

BET 4 Yuko pays $Cr_y T(*) \times \$1$. She receives \$1 if $T(*)$ and \$0 otherwise.

This bet also guarantees Yuko a loss. For given that $Cr_y T(*) \geq 0.5$, it follows that $\neg T(*)$ holds. Thus Yuko will receive nothing from this bet, but given that $Cr_y T(*) \geq 0.5$, it follows that $(Cr_y T(*) \times \$1) > 0$. If $Cr_y T(*) \geq 0.5$, it follows that Yuko will value a bet as fair, despite the fact that the bet guarantees her a loss.

From these two cases, it follows that whatever Yuko's credence is in $T(*)$ she is bound to value as fair some sure-loss bet.

This fact undermines the Dutch Book argument for PROBABILISM in two ways. First, it provides a counterexample to CONVERSE DUTCH BOOK.⁴³ According to CONVERSE DUTCH BOOK, if an agent

43. Thanks to an anonymous referee for noting this fact.

were to have probabilistically coherent credences, then she would not value as fair a sure-loss set of bets. We have seen, however, that this isn't so. Our proof that Yuko is guaranteed to value as fair some sure-loss bet made no assumption that she is probabilistically incoherent. CONVERSE DUTCH BOOK, however, is essential to the Dutch Book argument. For without CONVERSE DUTCH BOOK, we have no guarantee that considerations of sure-loss bets do not tell equally against probabilistically coherent credences as against probabilistically incoherent credences.

Second, the fact that Yuko is subject to valuing as fair a sure-loss bet, no matter what her credence, gives us good reason to reject SURE-LOSS AVOIDANCE. According to this principle, an agent ought to avoid having credences that lead to valuing as fair a sure-loss set of bets. However, we have seen that it is impossible for Yuko to meet this requirement. From OUGHT-CAN, it follows that we should reject SURE-LOSS AVOIDANCE.

Having noted these defects, I think we should conclude that the Dutch Book argument provides us with little reason to endorse PROBABILISM, and so provides us with little reason to hold that it would be irrational for Yuko to have credence 0.5 in $T(*)$ and credence 1 in $\neg T(*)$.

It is worth looking in a little more detail to see why exactly the Dutch Book argument fails. As in the case of the accuracy-dominance argument, the Dutch Book argument relies on certain mathematical results, and there are certain assumptions that are required in order to move from these mathematical results to the premises of the Dutch Book argument. The case of Yuko can be used to show—as in the case of the accuracy-dominance argument—that these assumptions fail in certain situations.

Let us again consider an agent who has credences defined over some finite algebra \mathcal{P} consisting of n propositions. We let \mathcal{W} denote the set of possible truth-value assignments to members of \mathcal{P} . For each $w \in \mathcal{W}$ and $\phi \in \mathcal{P}$, we let $w(\phi) = 1$ just in case ϕ is true at w and $w(\phi) = 0$ just in case ϕ is false at w . We let \mathcal{C} be the set of probabilistically coherent credal functions. For each $c \in \mathcal{C}$ and $\phi \in \mathcal{P}$, we let $c(\phi) = x$ just in case x is the credence assigned to ϕ by c . Let b be some function from \mathcal{P} to \mathbb{R} . And let there be some arbitrary ordering on the members of \mathcal{P} , so that ϕ_i denotes the i -th member of \mathcal{P} under this ordering. It can be shown that the following hold:⁴⁴

44. For a general proof of these theorems, see Paris 2001. While DB THEOREM can

DB THEOREM If $b \notin \mathcal{C}$, then there exists $s_1, \dots, s_n \in \mathbb{R}$ such that for every $w \in \mathcal{W}$, $\sum_{i=1}^n s_i(w(\phi_i) - b(\phi_i)) < 0$.

CDB THEOREM If there exists $s_1, \dots, s_n \in \mathbb{R}$ such that for every $w \in \mathcal{W}$, $\sum_{i=1}^n s_i(w(\phi_i) - b(\phi_i)) < 0$, then $b \notin \mathcal{C}$.

To see the connection between these mathematical theorems, on the one hand, and DUTCH BOOK and CONVERSE DUTCH BOOK, on the other, first note that if b is a function representing a credal state and w is a world, then $s_i(w(\phi_i) - b(\phi_i))$ is the outcome of a bet on the proposition ϕ_i at world w bought at stake s_i . Thus $\sum_{i=1}^n s_i(w(\phi_i) - b(\phi_i))$ is the outcome of a set of bets over the propositions in \mathcal{P} at stakes s_1, \dots, s_n .

Assume, for the moment, that for any agent and any possible credence distribution b over an algebra \mathcal{P} that this agent may have and any assignment of truth-values w to the members of \mathcal{P} , there is some possible world in which the agent has those credences and the members of \mathcal{P} have those truth-values. Given this assumption, it follows from DB THEOREM that if one fails to be probabilistically coherent, then there is some set of bets that one will view as fair that collectively guarantee a loss. That is, given this assumption, DB THEOREM entails DUTCH BOOK. Similarly, given this assumption, it follows from CDB THEOREM that if there is some set of bets that one views as fair but that collectively guarantee a loss, then one's credences must be probabilistically incoherent. That is, given this assumption, CDB THEOREM entails CONVERSE DUTCH BOOK.

By now, however, we know that this crucial assumption can fail. As we've already had occasion to note, the case of Yuko shows that there are situations in which certain credence distributions over an algebra are incompatible with certain distributions of truth-values over that very same algebra. It was this fact that served to undermine the accuracy-dominance argument for PROBABILISM, and it is for this very same reason that CONVERSE DUTCH BOOK fails in the case of Yuko despite the correctness of CDB THEOREM.

To see this, assume that Yuko has probabilistically coherent credences in $T(*)$ and $\neg T(*)$. For definiteness, assume that $Cr_y T(*) = 0.2$ and $Cr_y \neg T(*) = 0.8$. In accordance with CDB THEOREM, for any set of bets that Yuko may view as fair given these credences, there is some distribution of truth-values for $T(*)$ and $\neg T(*)$ such that, given this distri-

be traced back to Ramsey 1931, versions of CDB THEOREM were first proved by Lehman (1955) and Shimony (1955).

bution of truth-values, these bets would not yield a negative value. So, in particular, consider the following bet:

BET 5 Yuko pays \$0.2. She receives \$1 if $T(*)$ and \$0 otherwise.

There is clearly some distribution of truth-values for $T(*)$ and $\neg T(*)$ such that this bet yields a nonnegative outcome, namely, the distribution according to which $T(*)$ is true and $\neg T(*)$ is false. However, it doesn't follow from this fact that in valuing this bet as fair, Yuko is not valuing as fair a bet that guarantees a loss. For, crucially, the truth-value distribution that yields the nonnegative outcome on this bet is one that isn't compatible with Yuko's having credence 0.2 in $T(*)$.

Interestingly, the incompatibility of certain credence distributions with certain distributions of truth-values does not undermine the inference from DB THEOREM to DUTCH BOOK. The reason for this is that DB THEOREM tells us:

If b is not probabilistically coherent, then there are bets that are fair by the lights of b that lead to a loss *for every* possible distribution of truth-values.

Given the quantification over all possible truth-value distributions, it follows that even if some of these distributions are incompatible with the credal distribution b , it is still true that b will value as fair a set of bets that guarantees a loss. We can see the difference between this case and the case of CDB THEOREM by contraposing the latter. Looked at this way, the theorem tells us:

If b is probabilistically coherent, then for any bets deemed fair by b , *there is some* possible distribution of truth-values that do not lead to a loss.

In this case, given the existential quantification, it does matter if certain truth-value distributions are incompatible with b , for it may be that the truth-value distributions that do not lead to a loss are in fact incompatible with b . Indeed, this is exactly what happens in the case of Yuko.

Let me close this section by tentatively offering an additional argument for rejecting PROBABILISM. I've argued that we should reject SURE-LOSS AVOIDANCE, since it is incompatible with OUGHT-CAN. In cases like Yuko's, there is no way for an agent to avoid valuing as fair a sure-loss set of bets. A natural question that arises is what is the rational response for an agent such as Yuko who is guaranteed to be subject to a losing set of bets. I suggest that if one was initially attracted to SURE-LOSS AVOIDANCE, then one should endorse the following claim:

LOSS MINIMIZATION If an agent is guaranteed to value as fair some set of losing bets, then the agent should, if possible, have a credal state that minimizes her possible loss.

Given LOSS MINIMIZATION, it follows that Yuko ought to have credence 0.5 in $T(*)$ and credence 1 in $\neg T(*)$. For, it is these credences that will minimize Yuko's potential losses. To see this, recall that whatever Yuko's credence is in $T(*)$, she will be guaranteed to value as fair a losing bet.

If $Cr_y T(*) < 0.5$, then she is guaranteed a loss of $\$(1 - Cr_y T(*)$.

And if $Cr_y T(*) \geq 0.5$, then she is guaranteed a loss of $\$Cr_y T(*)$.

It is easy to see, given these two facts, that her guaranteed loss on bets on $T(*)$ is minimized by having credence 0.5 in this proposition. And given that she has this credence in $T(*)$, it follows that $\neg T(*)$ holds. Thus by having credence 1 in $\neg T(*)$, she ensures that she will not be subject to a loss by a bet on this proposition. This is, in fact, the best possible situation that Yuko could be in with respect to bets on $\neg T(*)$. It follows from these two facts that Yuko minimizes her potential loss by having credence 0.5 in $T(*)$ and credence 1 in $\neg T(*)$. By LOSS MINIMIZATION, it follows that these are the credences that Yuko ought to have.

I offer this argument tentatively since I think that LOSS MINIMIZATION is subject to the same worries that I voiced at the outset of this section about SURE-LOSS AVOIDANCE. For LOSS MINIMIZATION is most naturally read as a principle of practical rationality. It is, however, less clear that it is plausible as a principle of epistemic rationality. But it is the epistemic version of this principle that we require. Nonetheless, despite this potential shortcoming, the argument is worth highlighting. For the argument shows that if one was initially convinced of PROBABILISM by Dutch Book considerations, then one should in fact agree with me that it is rational for Yuko to be probabilistically incoherent.

Dutch Book considerations do not, then, provide us with reason to reject the claim that it can be rational for an agent to have probabilistically incoherent credences. Indeed, insofar as these considerations tell us anything about epistemic rationality, they would seem to motivate the claim that probabilistic incoherence can be epistemically rational.

7. Conclusion

I began this essay by showing that PROBABILISM has a surprising consequence. In certain cases, PROBABILISM demands that an agent either be insensitive to her own credal state or be ignorant of an obvious a priori necessary truth. Looking more closely at this case, we have seen that it provides us with the material to mount a strong case against PROBABILISM.

A *prima facie* compelling argument for PROBABILISM claims that probabilistic coherence is rationally required because it serves the goal of representing the world as accurately as possible. I've argued that the central premise of this argument is false. In certain cases, credal accuracy is best served by being probabilistically incoherent. Considerations of accuracy, instead of providing us with a reason to accept PROBABILISM, provide us with a reason to reject this principle.

References

- Caie, Michael. 2012. "Belief and Indeterminacy." *Philosophical Review* 121, no. 1: 1–54.
- de Finetti, Bruno. 1974. *Theory of Probability*. Vol. 1. New York: John Wiley and Sons.
- Easwaran, Kenny, and Branden Fitelson. 2012. "An 'Evidentialist' Worry about Joyce's Argument for Probabilism." *Dialectica* 66, no. 3: 425–33.
- Elga, Adam, and Andy Egan. 2005. "I Can't Believe I'm Stupid." *Philosophical Perspectives* 19: 79–93.
- Field, Hartry. 2008. *Saving Truth from Paradox*. New York: Oxford University Press.
- Hájek, Alan. 2008. "Dutch Book Arguments." In *The Oxford Handbook of Rational and Social Choice*, ed. Raul Anand, Prasanta Pattanaik, and Clemens Puppe, 173–95. New York: Oxford University Press.
- Jeffrey, Richard. 1983. *The Logic of Decision*. 2nd ed. Chicago: Chicago University Press.
- Joyce, James. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65: 575–603.
- . 1999. *The Foundations of Causal Decision Theory*. New York: Cambridge University Press.
- . 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, ed. F. Huber and C. Schmidt-Petri, 263–97. Dordrecht: Springer.
- Kripke, Saul. 1972. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

- . 1975. "Outline of a Theory of Truth." *Journal of Philosophy* 72, no. 19: 690–716.
- Lehman, R. Sherman. 1955. "On Confirmation and Rational Betting." *Journal of Symbolic Logic* 20, no. 3: 251–62.
- Leitgeb, Hannes, and Richard Pettigrew. 2010a. "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science* 77: 201–35.
- . 2010b. "An Objective Justification of Bayesianism 2: The Consequences of Minimizing Inaccuracy." *Philosophy of Science* 77: 236–72.
- Lewis, David. 1981. "Causal Decision Theory." *Australasian Journal of Philosophy* 59: 5–30.
- Paris, J. B. 2001. "A Note on the Dutch Book Method." In *Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications*, 301–6. Ithaca, NY: Shaker.
- Pettigrew, Richard. 2011a. "An Improper Introduction to Epistemic Utility Theory." In *EPSA Philosophy of Science: Amsterdam 2009*, ed. Henk de Regt, Stephan Hartmann, and Samir Okasha, 287–301. Netherlands: Springer.
- . 2011b. "Epistemic Utility Arguments for Probabilism." In *Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, ed. Edward N. Zalta, plato.stanford.edu/archives/win2011/entries/epistemic-utility.
- Pred, Joel, Sanjeev Kulkarni, Robert Seiringer, Elliott H. Lieb, Daniel Osherson, and H. Vincent Poor. 2009. "Probabilistic Coherence and Proper Scoring Rules." *IEEE Transactions on Information Theory* 55, no. 10: 4786–92.
- Ramsey, F. P. 1931. "Truth and Probability." In *Foundations of Mathematics and Other Essays*, ed. R. B. Braithwaite, 156–78. London: Routledge.
- Selten, Reinhard. 1998. "Axiomatic Characterization of the Quadratic Scoring Rule." *Experimental Economics* 1, no. 1: 43–61.
- Shimony, A. 1955. "Coherence and the Axioms of Confirmation." *Journal of Symbolic Logic* 20: 1–28.
- Sorensen, Roy. 1988. *Blindspots*. Oxford: Oxford University Press.
- Williams, J. R. G. 2012. "Generalized Probabilism: Dutch Books and Accuracy Domination." *Journal of Philosophical Logic* 41, no. 5: 811–40.