Arguments For—Or Against—Probabilism?

Alan Hájek

1 Introduction

On Mondays, Wednesdays, and Fridays, I call myself a *probabilist*.[1] In broad outline I agree with probabilism's key tenets: that

(1) an agent's beliefs come in degrees, which we may call *credences;* and that
(2) these credences are rationally required to conform to the probability calculus.

Here, 'the probability calculus' refers to at least the finite fragment of Kolmogorov's theory, according to which probabilities are non-negative, normalized (with a top value of 1), and finitely additive. Probabilism is a simple, fecund theory. Indeed, it achieves such an elegant balance of simplicity and strength that, in the spirit of Ramsey's and Lewis's accounts of 'law of nature', I am inclined to say that probabilism codifies the synchronic laws of epistemology.[2] Or so I am inclined on those days of the week.

But on the remaining days of the week I am more critical of probabilism. A number of well-known arguments are offered in its support, but each of them is inadequate. I do not have the space here to spell out all of the arguments, and all of their inadequacies. Instead, I will confine myself to four of the most important arguments—the Dutch Book, representation theorem, calibration, and gradational accuracy arguments—and I will concentrate on a particular inadequacy in each of them, in its most familiar form.

I think it is underappreciated how structurally similar these four arguments for probabilism are. Each begins with a mathematical theorem that adverts to credences

A. Hájek (✉)
Research School of Social Sciences, Australian National University
Canberra, ACT 0200, Australia
e-mail: alanh@coombs.anu.edu.au

[1] Much like Earman (1992), p. 1.

[2] The Ramsey/Lewis account has it that a law of nature is a theorem of the best theory of the universe—the true theory that best balances simplicity and strength. I say 'synchronic' laws of epistemology to allow for there being further 'diachronic' laws about how credences should update in the face of evidence.

F. Huber, C. Schmidt-Petri (eds.), *Degrees of Belief*, Synthese Library 342,
DOI 10.1007/978-1-4020-9198-8_9,

or degrees of belief, and that has the form of a *conditional* with an *existentially quantified consequent*. The antecedent speaks of some agent's credences violating the probability calculus. The consequent states the existence of *something putatively undesirable* that awaits such an agent, some way in which the agent's lot is *worse* than it could be by obeying the probability calculus, in a way that allegedly impugns her rationality. In each case, I will not question the theorem. But each argument purports to derive probabilism from the theorem. And it is underappreciated that in each case the argument, as it has been standardly or canonically presented, is invalid.[3]

The trouble in each case is that there is a *mirror-image* theorem, equally beyond dispute, that undercuts probabilism; if we focus on *it*, we apparently have an argument *against* probabilism, of exactly equal strength to the original argument *for* probabilism. The original theorem provides good news for probabilism, but the mirror-image theorem provides bad news. Once all this news is in, it provides no support for probabilism. The probabilist must then look elsewhere for more good news. I discuss some ways in which it has been found, or I attempt to provide it myself—but even then it is alloyed.

2 The Dutch Book Argument[4]

The Dutch Book argument assumes that credences can be identified with corresponding betting prices. Your degree of belief in X is p iff you are prepared to buy or sell at \$$p$ a bet that pays \$1 if p, and nothing otherwise. We may call p the price that you consider *fair* for the bet on X—at that price, you are indifferent between buying and selling the bet, and thus you see no advantage to either side. The betting interpretation, of course, involves a good deal of idealization, but I won't begrudge it here. (I begrudge it enough elsewhere—see Eriksson and Hájek 2007.) Instead, I will question the validity of the argument.

The centerpiece of the argument, as it has repeatedly been stated, is the following theorem, which I will not dispute.

Dutch Book Theorem

> If you violate probability theory, there exists a set of bets, each of which you consider fair, which collectively guarantee your loss.

---

[3] I say 'invalid' to convey that the fault with each argument is that the conclusion does not *follow from* the theorem, rather than that the theorem is false. There's a sense in which any argument for a necessary truth p is valid—even 'not p ∴ p'. After all, it is not possible for the premises of the argument to be true and the conclusion false. So if probabilism is a necessary truth, then the argument 'Snow is white ∴ probabilism' is valid in this sense. But philosophers often use 'invalid' in a different sense, according to which an argument is invalid if it is missing key steps needed to show us that its conclusion follows from its premises. This is the sense that I intend in this paper.

[4] This section streamlines an argument given in my (2005), which concentrated solely on the Dutch Book argument.

Call an agent who violates probability theory *incoherent*.[5] Call a set of bets, each of which you consider fair, and which collectively guarantee your loss, a *Dutch Book* against you. The Dutch Book theorem tells us that if you are incoherent, there exists a Dutch Book against you. Note the logical form: a conditional with an existentially quantified consequent. The antecedent speaks of a violation of probability theory; the consequent states the existence of something bad that follows from such a violation. We will see this form again and again.

So much for the theorem. What about the argument for probabilism? It is so simple that it can be presented entirely in words of one syllable:[6]

> You give some chance to *p:* it is the price that you would pay for a bet that pays a buck if *p* is true and nought if *p* is false. You give some chance to *q:* it is the price that you would pay for a bet that pays a buck if *q* is true and nought if *q* is false. And so on. Now, if you failed to live up to the laws of chance, then you could face a dire end. A guy—let's make him Dutch—could make a set of bets with you, each fair by your lights, yet at the end of the day you would lose, come what may. What a fool you would be! You should not tempt this fate. So you should bet in line with the laws of chance.

This argument is invalid. For all the Dutch Book theorem tells us, you may be just as susceptible to Dutch Books if you *obey* probability theory. Maybe the world is an unkind place, and we're all suckers! (Compare: it's certain that *if you pursue a career in philosophy, you will eventually die*; but *that's* hardly a reason to avoid a career in philosophy.) This possibility is ruled out by the surprisingly neglected, yet equally important Converse Dutch Book theorem: if you obey probability theory, then there does *not* exist a Dutch Book against you. So far, so good for probabilism.

But nothing can rule out the following mirror-image theorem, since it is clearly true. With an eye to the financial gains that are in the offing, let's call it the

Czech Book Theorem

> If you *violate* probability theory, there exists a set of bets, each of which you consider fair, which collectively guarantee your *gain*.

The proof of the theorem is easy: just rewrite the proof of the original Dutch Book theorem, replacing 'buying' by 'selling' of bets, and vice versa, throughout. You thereby turn the original 'Dutch Bookie' who milks you into a 'Czech Bookie' whom you milk. Call a set of bets, each of which you consider fair, and which collectively guarantee your gain, a *Czech Book* for you. The Czech Book theorem tells us that if you are incoherent, there exists a Czech Book for you. It is a simple piece of mathematics, and there is no disputing it.

So much for the theorem. I now offer the following argument *against* probabilism, again in words of one syllable. It starts as before, then ends with a diabolical twist:

---

[5] de Finetti used the word 'incoherent' to mean 'Dutch bookable', while some other authors use it as I do. It will be handy for me to have this word at my disposal even when I am not discussing Dutch books.

[6] The homage to George Boolos will be obvious to those who know his (1994).

> . . . Now, if you failed to live up to the laws of chance, then you could face a sweet end. A guy—let's make him Czech—could make a set of bets with you, each fair by your lights, yet at the end of the day you would *win*, come what may. What a brain you would be! You should seek this fate. So you should bet out of line with the laws of chance.

This argument is invalid. For all the Czech Book theorem tells us, you may be just as open to Czech Books if you obey probability theory. Maybe the world is a kind place, and we're all winners! (Compare: it's certain that *if you pursue a career in philosophy, you will be happy at some point in your life*; but *that's* hardly a reason to pursue a career in philosophy.) This possibility is ruled out by the surprisingly neglected, yet equally important Converse Czech Book theorem: if you obey probability theory, then there does *not* exist a Czech Book for you.[7] So far, so bad for probabilism.

Let's take stock, putting the theorems side by side:

Iff you violate probability theory, there exists a *specific bad thing* (a Dutch Book against you).

Iff you violate probability theory, there exists a *specific good thing* (a Czech Book for you).

The Dutch Book argument sees the incoherent agent's glass as half empty, while the Czech Book argument sees it as half full. If we focus on the former, probabilism prima facie looks compelling; but if we focus on the latter, the denial of probabilism prima facie looks compelling.

2.1 Saving the Dutch Book Argument[8]

If you survey the vast literature on Dutch Book arguments, you will find that most presentations of it focus solely on bets bought or sold at exactly your *fair* prices, bets that by your lights bestow no advantage to either side. (See, e.g., Adams (1962), Adams and Rosenkrantz (1980), Armendt (1992), Baillie (1973), Carnap (1950, 1955), Christensen (1991, 1996, 2001), de Finetti (1980), Döring (2000), Earman (1992), Gillies (2000), Howson and Urbach (1993), Jackson and Pargetter (1976), Jeffrey (1983, 1992), Kaplan (1996), Kemeny (1955), Kennedy and Chihara (1979), Lange (1999), Lehman (1955), Maher (1993), Mellor (1971), Milne (1990), Rosenkrantz (1981), Seidenfeld and Schervish (1983), Skyrms (1986), van Fraassen (1989), Waidacher (1997), Weatherson (1999), and Williamson (1999)). But bets that you consider fair are not the only ones that you accept; you also accept bets that you consider *favourable*—that is, better than fair. You are prepared to sell a given bet at higher prices, and to buy it at lower prices, than your fair price. This

[7] Proof: Suppose for reductio that you obey probability theory, and that there is a set of bets, each of which you consider fair, which collectively guarantee your gain. Then swapping sides of these bets, you would still consider each fair, yet collectively they would guarantee your loss. This contradicts the Converse Dutch Book theorem.

[8] This sub-section mostly repeats the corresponding section of my (2005)—for this move in the dialectic I have nothing more, nor less, to say than I did there.

observation is just what we need to break the symmetry that deadlocked the Dutch Book argument and the Czech Book argument.

Let us rewrite the theorems, replacing 'fair' with 'fair-or-favourable' throughout, and see what happens:

Dutch Book theorem, revised:

If you violate probability theory, there exists a set of bets, each of which you consider fair-or-favourable, which collectively guarantee your loss.

Converse Dutch Book theorem, revised:

If you obey probability theory, there does not exist a set of bets, each of which you consider fair-or-favourable, which collectively guarantee your loss.

Czech Book theorem, revised:

If you violate probability theory, there exists a set of bets, each of which you consider fair-or-favourable, which collective guarantee your gain.

Converse Czech Book theorem, revised:

If you obey probability theory, there does not exist a set of bets, each of which you consider fair-or-favourable, which collectively guarantee your gain.

The first three of these revisions are true, obvious corollaries of the original theorems. Indeed, the revised versions of the Dutch Book theorem and the Czech Book theorem follow immediately, because any bet that you consider fair you ipso facto consider fair-or-favourable. The revised version of the Converse Dutch Book theorem also follows straightforwardly from the original version.[9]

But the revised version of the Converse Czech Book theorem is not true: if you obey probability theory, there *does* exist a set of bets, each of which you consider fair-or-favourable, which collectively guarantee your gain. The proof is trivial.[10] The revision from 'fair' to 'fair-or-favourable' makes all the difference. And with the failure of the revised version of the Converse Czech Book theorem, the corresponding revised version of the Czech Book argument is invalid. There were no Czech Books for a coherent agent, because Czech Books were defined in terms of *fair* bets. But there are other profitable books besides Czech Books, and incoherence is not required in order to enjoy those. Opening the door to fair-or-favourable bets opens the door to sure profits for the coherent agent. So my parody no longer goes through when the Dutch Book argument is cast in terms of fair-or-favourable bets, as it always should have been.

---

[9] Proof. Suppose you obey probability theory. Suppose for reductio that there does exist a set of bets, each of which you consider fair-or-favourable, that collectively guarantee a loss; let this loss be $L > 0$. Then you must regard at least one of these bets as favourable (for the Converse Dutch Book theorem assures us that if you regarded them all as fair, then there could not be such guaranteed loss). That is, at least one of these bets is sold at a higher price, or bought at a cheaper price, than your fair price for it. For each such bet, replacing its price by your fair price would increase your loss. Thus, making all such replacements, so that you regard all the bets as fair, your guaranteed loss is even greater than L, and thus greater than 0. This contradicts the Converse Dutch Book theorem. Hence, we must reject our initial supposition, completing the reductio. We have proved the revised version of the Converse Dutch Book theorem.

[10] Suppose you obey the probability calculus; then if T is a tautology, you assign $P(T) = 1$. You consider fair-or-favourable paying less than \$1—e.g., 80 cents—for a bet on T at a \$1 stake, simply because you regard it as favourable; and this bet guarantees your gain.

I began this section by observing that most of the presenters of the Dutch Book argument formulate it in terms of your fair prices. You may have noticed that I left Ramsey off the list of authors.[11] His relevant remarks are confined to 'Truth and Probability', and what he says is somewhat telegraphic:

> If anyone's mental condition violated these laws [of rational preference, leading to the axioms of probability], his choice would depend on the precise form in which the options were offered him, which would be absurd. He could have a book made against him by a cunning bettor and would then stand to lose in any event ... Having degrees of belief obeying the laws of probability implies a further measure of consistency, namely such a consistency between the odds acceptable on different propositions as shall prevent a book being made against you (1980/1931, 42).

Note that Ramsey does not say that all of the bets in the book are individually considered fair by the agent. He leaves open the possibility that some or all of them are considered better than fair; indeed 'acceptable' odds is synonymous with 'fair-or-favourable' odds. After all, one would accept bets not only at one's fair odds, but also at better odds. Ramsey again:

> By proposing a bet on *p* we give the subject a possible course of action from which so much extra good will result to him if *p* is true and so much extra bad if *p* is false. Supposing the bet to be in goods and bads instead of in money, he will take a bet at any better odds than those corresponding to his state of belief; in fact his state of belief is measured by the odds he will just take; ... (1980/1931, 37).

It was the subsequent authors who restricted the Dutch Book argument solely to fair odds. In doing so, they sold it short.[12]

2.2 *'The Dutch Book Argument Merely Dramatizes an Inconsistency in the Attitudes of an Agent Whose Credences Violate Probability Theory'*

So is it a happy ending for the Dutch Book argument after all? Unfortunately, I think not. What exactly does the argument show? Taken literally, it is supposed to show that an incoherent agent is susceptible to *losing money*. Understood this naïve way, it is easily rebutted—as various authors have noted, the agent can just refuse to bet when approached by a Dutch bookie. To put the old point a novel way, in that case the susceptibility is *masked*. The underlying basis for the agent's betting

---

[11] Skyrms (1986) was on the list, but not Skyrms (1980, 1984, or 1987). For example, in his (1987) he notes that an agent will buy or sell contracts 'at what he considers the fair price or better' (p. 225), and in his (1980), he explicitly states the Dutch Book theorem in terms of 'fair or favourable' bets (p. 118). Shimony (1955), Levi (1974), Kyburg (1978), Armendt (1993), Douven (1999), and Vineberg (2001) also leave open that the bets concerned are regarded as favourable. It is hard to tell whether certain other writers on the Dutch Book argument belong on the list or not (e.g., Ryder 1981, Moore 1983).

[12] This ends the sub-section that was lifted from my (2005); the remainder of this paper is again new.

dispositions—her relevant mental state—is unchanged, but she protects them from ever being triggered. Note well: *she* protects them; this is not even a case of masking that is hostage to external influences (as are some of the well-known examples in the dispositions literature).[13] The protection is due to another disposition *of her own*. Nor need we necessarily look far to find the masking disposition. It may simply be her disposition to do the maths, to notice the net loss that taking all the bets would accrue, and thus to shun them. Her disposition to take the bets because she finds them individually favourable is masked by her disposition to refuse them because she can see that they collectively lose. Now, you may say that she has *inconsistent* dispositions, to accept the bets under one mode of presentation and to shun them under another, and that she is *ipso facto* irrational. That's surely a better interpretation of the lesson of the Dutch Book argument, and we are about to consider it properly. But here I am merely rebutting the naïve interpretation that takes literally the lesson of the monetary losses.

So let's consider the more promising interpretation of the argument, also originating with Ramsey, which regards the susceptibility as symptomatic of a deeper defect. Recall his famous line: 'If anyone's mental condition violated these laws [of rational preference, leading to the axioms of probability], his choice would depend on the precise form in which the options were offered him, which would be absurd.' Authors such as Skyrms (1984) and Armendt (1993) regard this is as the real insight of the Dutch Book argument: an agent who violates probability theory would be guilty of a kind of double-think, 'divided-mind inconsistency' in Armendt's phrase. Such authors downplay the stories of mercenary Dutch guys and sure monetary losses; these are said merely to dramatize that underlying state of inconsistency. Skyrms describes the Dutch Book theorem as 'a striking corollary' of an underlying inconsistency inherent in violating the probability axioms (1984, 22). The inconsistency is apparently one of regarding a particular set of bets both as fair (since they are regarded individually as fair) and as unfair (since they collectively yield a sure loss).

Notice that put this way, there is no need to replace talk of 'fair' bets with 'fair-or-favourable' bets, the way there was before. But we could do so: the inconsistency equally lies in regarding the same set of bets both as fair-or-favourable and as not fair-or-favourable. Moreover, there is nothing essentially *Dutch* about the argument, interpreted this way. The Czech Book theorem is an equally striking corollary of the same underlying inconsistency: regarding another set of bets both as fair (since they are regarded individually as fair) and as *better-than-fair* (since they collectively yield a sure gain). To be sure, guaranteed losses may be more *dramatic* than guaranteed gains, but the associated double-think is equally bad.

So now the real argument for probabilism seems not to stem from the Dutch Book theorem (which is merely a 'corollary'), but from another *putative theorem*, apparently more fundamental. I take it to be this: *If you violate probability theory,*

[13] I thank Jonathan Schaffer for pointing out the distinction between internally-based and externally-based masking.

there exists a set of propositions (involving bets) to which you have inconsistent attitudes. Either the Dutch Book bets or the Czech Book bets could be used to establish the existence claim. This again is a conditional with an existentially quantified consequent. Now I don't have a mirror-image theorem to place alongside it, in order to undercut it.

However, nor have I seen the *converse* of this more fundamental putative theorem; still less am I aware of anyone claiming to have proved it. It seems to be a live possibility that if you obey probability theory, then there also exists a set of propositions to which you have inconsistent attitudes—not inconsistent in the sense of being Dutch-bookable (the converse Dutch Book theorem assures us of this), but inconsistent nonetheless.[14] That is, I have not seen any argument that in virtue of avoiding the inconsistency of Dutch-bookability, at least some coherent agents are guaranteed to avoid all inconsistency. Without a proof of this further claim, it seems an open question whether probabilistically coherent agents might *also* have inconsistent attitudes (somewhere or other). Maybe non-extremal credences, probabilistic or not, necessarily manifest a kind of inconsistency. I don't believe that, but I don't see how the Dutch Book argument rules it out. The argument *needs* to rule it out in order to preempt the possibility of a partners-in-crime defence of non-probabilism: the possibility that we are all epistemically damned whatever we do. Indeed, if all intermediate credences were 'inconsistent' in this sense, then this sense of inconsistency would not seem so bad after all. I said earlier that the original Dutch Book argument, understood in terms of monetary losses, is invalid; the converse Dutch Book theorem came to its rescue (even though this theorem is surprisingly neglected). Now I am saying that the Ramsey-style Dutch Book argument, understood as dramatizing an inconsistency in attitudes, is similarly invalid; it remains to be seen if the converse of the putative theorem (italicized in the previous paragraph) will come to its rescue.

I say 'putative theorem' because its status as a theorem is less clear than before—this status is disputed by various authors. Schick (1986) and Maher (1993) question the inconsistency of the attitudes at issue regarding the additivity axiom. They reject the 'package principle', which requires one to value a set of bets at the sum of the values of the bets taken individually, or less specifically, to regard a set of bets as fair if one regards each bet individually as fair. The package principle seems especially problematic when there are interference effects between the bets in a package—e.g. the placement of one bet is correlated with the outcome of another. For example, you may be very confident that your partner is happy: you will pay 90 cents for a bet that pays a dollar if so. You may be fairly confident that the Democrats will win the next election: you will pay 60 cents for a bet that pays a dollar if they win. So by the package principle, you should be prepared to pay $1.50 for both bets. But you also know that your partner hates you betting on political matters and inevitably finds

[14] I suppose you are safe from such inconsistency if you obey probability theory trivially with a consistent assignment of 1's and 0's, corresponding to a consistent truth-value assignment. Let us confine our attention, then, to *non*-trivial probability assignments, which after all are the lifeblood of probabilism.

out as soon as you do so. So you are bound to lose the partner-is-happy bet if you package it with the Democrats bet: you are certain that you will win a maximum of a dollar, so $1.50 is a bad price for the package. This is just a variation on a problem with the betting interpretation in its own right: placing a bet on *X* can change one's probability for *X*. Still, this variation only arises for packages, not single bets.

Or consider an agent who attaches extra value to a package in which risky gambles cancel each other, compared to the gambles assessed individually. Buying insurance can be a rational instance of this. Suppose I am forced to bet on a coin toss. I may attach extra value to a package that includes both a bet on Heads and an equal bet on Tails compared to the individual bets, if I especially want to avoid the prospect of loss. We see a similar pattern of preferences in the so-called Allais 'paradox'.[15] Granted, such preferences cannot be rationalized by the lights of expected utility theory. Yet arguably they can be rational. Moreover, the package principle is even more problematic for infinite packages—see Arntzenius, Elga and Hawthorne (2004)—so the Dutch Book argument for countable additivity is correspondingly even more problematic.

This leaves us with a dilemma for the Dutch Book argument for probabilism. Either we interpret its cautionary tale of monetary losses literally, or not. In the former case, the moral of the tale seems to be false: an incoherent agent can avoid those losses simply by masking her disposition to accept the relevant bets with another disposition to reject them. In the latter case, one may question the putative *theorem* when it is stated in terms of 'inconsistent' attitudes, and there seems to be no converse theorem to guarantee that at least some probabilists avoid such inconsistency. Either way, the argument for probabilism is invalid.

3 Representation Theorem-Based Arguments

The centerpiece of the argument for probabilism from representation theorems is some version of the following theorem, which I will not dispute.

Representation Theorem

> If all your preferences satisfy certain 'rationality' conditions, then there exists a representation of you as an expected utility maximizer, relative to some probability and utility function.

(The 'rationality' constraints on preferences are transitivity, connectedness, independence, and so on.) The contrapositive gets us closer to the template that I detect in all the arguments for probabilism:

> If there does not exist a representation of you as an expected utility maximizer, relative to some probability and utility function, then there exist preferences of yours that fail to satisfy certain 'rationality' conditions.

[15] Thanks here to Kenny Easwaran.

Focusing on the probabilistic aspect of the antecedent, we have a corollary that fits the conditional-with-an-existentially-quantified-consequent form:

> If your credences cannot be represented with a probability function, then there exist preferences of yours that fail to satisfy certain 'rationality' conditions.

The antecedent involves a violation of the probability calculus; the consequent states the existence of a putatively undesirable thing that follows: some violation of the 'rationality' conditions on preferences. In short, *if your credences cannot be represented with a probability function, then you are irrational*.

I will dispute that probabilism follows from the original theorem, and *a fortiori* that it follows from the corollary. For note that probabilism is, in part, the stronger thesis that *if your credences violate probability theory, then you are irrational* (a restatement of what I called tenet (2) at the outset). It is clearly a stronger thesis than the corollary, because its antecedent is weaker: while 'your credences cannot be represented with a probability function' entails 'your credences violate probability theory', the converse entailment does not hold. For it is possible that your credences violate probability theory, and that nonetheless they can be *represented* with a probability function. Merely being *representable* some way or other is cheap, as we will see; it's more demanding actually to *be* that way. Said another way: it's one thing to act *as if* you have credences that obey probability theory, another thing to actually *have* credences that obey probability theory. Indeed, probabilism does not even follow from the theorem coupled with the premises that Maher adds in his meticulous presentation of his argument for probabilism, as we will also see.

The concern is that for all we know, the mere *possibility* of representing you one way or another might have less force than we want; your acting *as if* the representation is true of you does not make it true of you. To make this concern vivid, suppose that I represent your preferences with *Voodooism*. My voodoo theory says that there are warring voodoo spirits inside you. When you prefer A to B, then there are more A-favouring spirits inside you than B-favouring spirits. I interpret all of the usual rationality axioms in voodoo terms. Transitivity: if you have more A-favouring spirits than B-favouring spirits, and more B-favouring spirits that C-favouring spirits, then you have more A-favouring spirits than C-favouring spirits. Connectedness: any two options can be compared in the number of their favouring spirits. And so on. I then 'prove' Voodooism: if your preferences obey the usual rationality axioms, then there exists a Voodoo representation of you. That is, you act *as if* there are warring voodoo spirits inside you in conformity with Voodooism. Conclusion: rationality requires you to have warring Voodoo spirits in you. Not a happy result.

Hence there is a need to bridge the gap between the possibility of representing a rational agent a particular way, and this representation somehow being *correct*. Maher, among others, attempts to bridge this gap. I will focus on his presentation, because he gives one of the most careful formulations of the argument. But I suspect my objections will carry over to any version of the argument that infers the rational *obligation* of having credences that are probabilities from the mere represent*ability* of an agent with preferences obeying certain axioms.

Maher claims that the expected utility representation is *privileged*, superior to rival representations. First, he assumes what I will call *interpretivism:*

> an attribution of probabilities and utilities is correct just in case it is part of an overall interpretation of the person's preferences that makes sufficiently good sense of them and better sense than any competing interpretation does (1993, 9).

Then he maintains that, when available, an expected utility interpretation is a *perfect* interpretation:

> if a person's preferences all maximize expected utility relative to some p and u, then it provides a perfect interpretation of the person's preferences to say that p and u are the person's probability and utility functions.

He goes on to give the argument from the representation theorems:

> . . . we can show that rational persons have probability and utility functions if we can show that rational persons have preferences that maximize expected utility relative to some such functions. An argument to this effect is provided by representation theorems for Bayesian decision theory.

He then states the core of these theorems:

> These theorems show that if a person's preferences satisfy certain putatively reasonable qualitative conditions, then those preferences are indeed representable as maximizing expected utility relative to some probability and utility functions (1993, 9).

We may summarize this argument as follows:

Representation Theorem Argument

1. (Interpretivism) You have a particular probability and utility function iff attributing them to you provides an interpretation that makes:

   (i) sufficiently good sense of your preferences and
   (ii) better sense than any competing interpretation.

2. (Perfect interpretation) Any maximizing-expected-utility interpretation is a perfect interpretation (when it fits your preferences).
3. (Representation theorem) If you satisfy certain constraints on preferences (transitivity, connectedness, etc.), then you can be interpreted as maximizing expected utility.
4. The constraints on preferences assumed in the representation theorem of 3 are *rationality constraints*.

Therefore (generalizing what has been established about 'you' to 'all rational persons'),

Conclusion: [All] rational persons have probability and utility functions (1993, 9)

The conclusion is probabilism, and a bit more, what we might call *utilitism*.

According to Premise 1, a necessary condition for you to have a particular probability and utility function is their providing an interpretation of you that is *better* than any competing interpretation. Suppose we grant that the expected utility representation is a perfect interpretation when it is available. To validly infer probabilism, we need also to show that *no other interpretation is as good*. Perhaps this can be done, but nothing in Maher's argument does it. For all that he has said, there are other perfect interpretations out there (whatever that means).

Probabilism would arguably follow from the representation theorem if *all* representations of the preference-axiom-abiding agent were probabilistic representations.[16] Alas, this is not the case, for the following 'mirror-image' theorem is equally true:

> If all your preferences satisfy the same 'rationality' conditions, then you can be interpreted as maximizing non-expected utility, some rival to expected utility, and in particular as having credences that violate probability theory.

How can this be? The idea is that the rival representation compensates for your credences' violation of probability theory with some non-standard rule for combining your credences with your utilities. Zynda (2000) proves this mirror-image theorem. As he shows, if you obey the usual preference axioms, you can be represented with a sub-additive belief function, and a corresponding combination rule. For all that Maher's argument shows, this rival interpretation may also be 'perfect'.

According to probabilism, rationality requires an agent's *credences* to obey the probability calculus. We have rival ways of representing an agent whose preferences obey the preference axioms; which of these representations correspond to her *credences?* In particular, why should we privilege the probabilistic representation? Well, there may be reasons. Perhaps it is favoured by considerations of simplicity, fertility, consilience, or some other theoretical virtue or combination thereof—although good luck trying to clinch the case for probabilism by invoking these rather vague and ill-understood notions. And it is not clear that these considerations settle the issue of what rational credences *are*, as opposed to how they can be fruitfully modelled. (See Eriksson and Hájek 2007 for further discussion.) It seems to be a further step, and a dubious one at that, to reify the theoretical entities in our favourite model of credences.

It might be objected that the 'rival' representations are not *really* rival. Rather, the objection goes, they form a family of isomorphic representations, and choosing among them is merely a matter of convention; whenever there is a probabilistic representation, all of these other representations impose exactly the same laws on rational opinion, just differently expressed.[17] First, a perhaps flat-footed reply: I understand 'probabilism' to be *defined* via Kolmogorov's axiomatization of probability. So, for example, a non-additive measure is not a probability function, so understood. That said, one might want to have a broader understanding of 'probabilism', encompassing any transformation of a probability function and a correspondingly transformed combination rule for utility that yields the same ordinal representation

---

[16] Only arguably. In fact, I think that it does not follow, because the preference axioms are not all rationality constraints.

[17] I thank Hartry Field and Jim Joyce for independently offering versions of this objection to me.

of preferences. If that is what is intended, then probabilism should be stated in those terms, and not in the flat-footed way that is entirely standard. We would then want to reexamine the arguments for probabilism in that light—presumably with a revised account of 'credence' in terms of betting, a revised statement of what 'calibration' consists in, and revised axioms on 'gradational accuracy'. But I am getting ahead of myself—calibration, and gradational accuracy are just around the corner!

In any case, I believe that my main point stands, even with a more liberal understanding of 'probabilism': the representation theorem argument is invalid. We have the theorem:

> if you obey the preference axioms, then you are representable by a credence function that is a suitable transformation of a probability function.

But to be able validly to infer probabilism in a broad sense, we need the further theorem:

> if you obey the preference axioms, then you are *not* also representable by a credence function that is *not* a suitable transformation of a probability function.

It seems that the status of this is at best open at the moment. The representation theorem argument for probabilism remains invalid until the case is closed in favour of the further theorem.

4 The Calibration Argument

The centerpiece of the argument is the following theorem—another conditional with an existentially quantified consequent—which I will not dispute.

Calibration Theorem

> If c violates the laws of probability then there is a probability function c^+ that is better calibrated than c under every logically consistent assignment of truth-values to propositions.

Calibration is a measure of how well credences match corresponding relative frequencies. Suppose that you assign probabilities to some sequence of propositions—for example, each night you assign a probability to it raining the following day, over a period of a year. Your assignments are *(perfectly) calibrated* if proportion p of the propositions to which you assigned probability p are true, for all p. In the example, you are perfectly calibrated if it rained on 10% of the days to which you assigned probability 0.1, on 75% of the days to which you assigned probability 0.75, and so on. More generally, we can measure how well calibrated your assignments are, even if they fall short of perfection.

The clincher for probabilism is supposed to be the calibration theorem. If you are incoherent, then you can figure out a priori that you could be better calibrated by being coherent instead. Perfect calibration, moreover, is supposed to be A Good Thing, and a credence function that is better calibrated than another one is thereby supposed to be superior in at least one important respect. Thus, the argument concludes, you should be coherent. See Joyce (2004) for a good exposition of this style of argument for probabilism (although he does not endorse it himself).

I argue elsewhere (MS) that perfect calibration may be A Rather Bad Thing, as does Seidenfeld (1985) and Joyce (1998). More tellingly, the argument, so presented, is invalid.

I will not quarrel with the calibration theorem. Nor should the probabilist quarrel with the following 'mirror-image' theorem:

> If c violates the laws of probability then there is a *non*-probability function c^+ that is better calibrated than c under every logically consistent assignment of truth-values to propositions.

Think of c^+ as being more coherent than c, but not entirely coherent. If c assigns, say, 0.2 to rain and 0.7 to not-rain, then an example of such a c^+ is a function that assigns 0.2 to rain and 0.75 to not-rain. If you are incoherent, then you know a priori that you could be better calibrated by *staying in* coherent, but in some other way.[18]

So as it stands, the calibration argument is invalid. Given that you can improve your calibration situation *either* by moving to some probability function *or* by moving to some other non-probability function, why do you have an incentive to move to a probability function? The answer, I suppose, is this. If you moved to a non-probability function, you would only recreate your original predicament: you would know a priori that you could do better by moving to a probability function. Now again, you could *also* do better by moving to yet another non-probability function. But the idea is that moving to a non-probability function will give you no rest; it can never be a stable stopping point. Still, the argument for probabilism is invalid as it stands. To shore it up, we had better be convinced that at least some probability functions *are* stable stopping points.

The following converse theorem would do the job:

> If c *obeys* the laws of probability then there is *not* another function c^+ that is better calibrated than c under every logically consistent assignment of truth-values to propositions.

I offer the following near-trivial proof: Let P be a probability function. P can be perfectly calibrated—just consider a world where the relative frequencies are exactly

---

[18] To be sure, the mirror-image theorem gives you no advice as to which non-probability function you should move to. But nor did the calibration theorem give you advice as to which probability function you should move to. Moreover, for all the theorem tells us, you can *worsen* your calibration index, come what may, by moving from a non-probability function to a 'wrong' probability function. Here's an analogy (adapted from Aaron Bronfman and Jim Joyce, personal communication). Suppose that you want to live in the best city that you can, and you currently live in an American city. I tell you that for each American city, there is a better Australian city. (I happen to believe this.) It does not follow that you should move to Australia. If you do not know *which* Australian city or cities are better than yours, moving to Australia might be a backward step. You might choose Coober Pedy.

That said, the calibration argument may still be probative, still diagnostic of a defect in an incoherent agent's credences. To be sure, she is left only with the general admonition to become coherent, without any advice on how specifically to do so. Nevertheless, the admonition is non-trivial. Compare: when an agent has inconsistent *beliefs*, logic may still be probative, still diagnostic of a defect in them. To be sure, she is left only with the general admonition to become coherent, without any advice on how specifically to do so. Nevertheless, the admonition is non-trivial.

as P predicts, as required by calibration. (If P assigns some irrational probabilities, then the world will have to provide infinite sequences of the relevant trials, and calibration will involve agreement with limiting relative frequencies.) At that world, no other function can be better calibrated than P. Thus, P cannot be beaten by some other function, come what may, in its calibration index—for short, P is not *calibration-dominated*. Putting this result together with the calibration theorem, we have the result that *probability functions are exactly the functions that are not calibration-dominated*.

The original calibration argument for probabilism, as stated above, was invalid, but I think it can be made valid by the addition of this theorem. However, this is not yet a happy ending for calibrationists. If you are a fan of calibration, surely what matters is being well calibrated in the *actual* world, and being coherent does not guarantee that.[19] A coherent weather forecaster who is wildly out of step with the actual relative frequencies can hardly plead that at least he is perfectly *in* step with the relative frequencies *in some other possible world!* (Compare: someone who has consistent but wildly false beliefs can hardly plead that at least his beliefs are true *in some other possible world!*)

5 The Gradational Accuracy Argument

Joyce (1998) rightly laments the fact that ‘probabilists have tended to pay little heed to the one aspect of partial beliefs that would be of most interest to epistemologists: namely, their role in representing the world’s state’ (576). And he goes on to say: ‘I mean to alter this situation by first giving an account of what it means for a system of partial beliefs to accurately represent the world, and then explaining why having beliefs that obey the laws of probability contributes to the basic epistemic goal of accuracy.’

The centerpiece of his ingenious (1998) argument is the following theorem—yet another conditional with an existentially quantified consequent—which I will not dispute.

Gradational Accuracy Theorem

> if c violates the laws of probability then there is a probability function c^+ that is strictly more accurate than *c under every logically consistent assignment of truth-values to propositions* (Joyce 2004, 143).

[19] Seidenfeld (1985) has a valuable discussion of a theorem due to Pratt and rediscovered by Dawid that may seem to yield the desired result. Its upshot is that if an agent is probabilistically coherent, and updates by conditionalization after each trial on feedback information about the result of that trial, then in the limit calibration is achieved almost surely (according to her own credences). This is an important result, but it does not speak to the case of an agent who is coherent but who has not updated on such an infinite sequence of feedback information, and indeed who may never do so (e.g., because she never gets such information).

Joyce gives the following account of the argument. It

> relates probabilistic consistency to the *accuracy* of graded beliefs. The strategy here involves laying down a set of axiomatic constraints that any reasonable gauge of accuracy for confidence measures should satisfy, and then showing that probabilistically inconsistent measures are always less accurate than they need to be (2004, 142).

Saying that incoherent measures are 'always less accurate than they need to be' suggests that they are always unnecessarily inaccurate—that they always *could* be more accurate. But this would not distinguish incoherent measures from coherent measures that assume non-extremal values—that is, coherent measures that are not entirely opinionated. After all, such a measure *could* be more accurate: an opinionated measure that assigns 1 to the truth and 0 to all false alternatives to it, respectively, *is* more accurate. Indeed, if a coherent measure *P* assumes non-extremal values, then *necessarily* there exists another measure that is more accurate than *P*: in each possible world there exists such a measure, namely an opinionated measure that gets all the truth values right in that world. More disturbingly, if a coherent measure *P* assumes non-extremal values, then necessarily there exists an *incoherent* measure that is more accurate than *P*: for example, one that raises *P*'s non-extremal probability for the truth to 1, while leaving its probabilities for falsehoods where they are. (The coherent assignment <1/2, 1/2> for the outcomes of a coin toss, <Heads, Tails>, is less accurate than the incoherent assignment <1, 1/2> in a world where the coin lands Heads, and it is less accurate than the incoherent assignment <1/2, 1> in a world where the coin lands Tails.) But this had better not be an argument against the rationality of having a coherent intermediate-valued credence function—that would hardly be good news for probabilism!

The reversal of quantifiers in Joyce's theorem appears to save the day for probabilism. It isn't merely that:

> if your credences violate probability theory, in each possible world there exists a probability function that is more accurate than your credences.

More than that, by his theorem we have that:

> if your credences violate probability theory, there exists a probability function such that in each possible world, *it* is more accurate than your credences.

The key is that the *same* probability function outperforms your credences in each possible world, if they are incoherent. Thus, by the lights of gradational accuracy you would have nothing to lose and everything to gain by shifting to that probability function. So far, so good. But we had better be convinced, then, that at least some *coherent* intermediate-valued credences do not face the same predicament—that they *cannot* be outperformed in each possible world by a single function (probability, or otherwise). Well, let's see.

With the constraints on reasonable gauges of accuracy in place, Joyce (1998) proves the gradational accuracy theorem. He concludes: 'To the extent that one accepts the axioms, this shows that the demand for probabilistic consistency follows from the purely epistemic requirement to hold beliefs that accurately represent the world' (2004, 143).

Let us agree for now that the axioms are acceptable. (Maher 2003 doesn't.) I have already agreed that the theorem is correct. But I do not agree that the demand for probabilistic consistency follows.

Again, we have a 'mirror-image' theorem:

> if c violates the laws of probability then there is a *non*-probability function c^+ that is strictly more accurate than c *under every logically consistent assignment of truth-values to propositions.*

(As with the corresponding calibration theorem, the trick here is to make c^+ *less* incoherent than c, but incoherent nonetheless.) If you are incoherent, then your beliefs could be made more accurate by moving to another incoherent function. Why, then, are you under any rational obligation to move instead to a coherent function? The reasoning, I gather, will be much as it was in our discussion of the calibration theorem. Stopping at a non-probability function will give you no rest, because by another application of the gradational accuracy theorem, you will again be able to do better by moving to a probability function. To be sure, you will *also* be able to do better by moving to yet another non-probability function. But a non-probability function can never be a stable stopping point: it will always be strictly accuracy-dominated by some other function.[20]

I contend that Joyce's (1998) argument is invalid as it stands. As before, to shore it up we had better be convinced that at least some probability functions *are* stable stopping points. The following converse theorem would do the job:

> If c *obeys* the laws of probability then there is *not* another function c^+ that is strictly more accurate than c under every logically consistent assignment of truth-values to propositions.

Things have moved quickly in this area recently. As it turns out, Joyce (this volume) reports some results (by Lindley and Lieb et al.), and he proves a very elegant result of his own, that entail this theorem. In these results, constraints of varying strength are imposed on the 'reasonable' gauges of accuracy that will 'score' credence functions. The weakest such constraints on such *scoring rules* that Joyce considers appear in his paper's final theorem.[21]

To understand it, we need some terminology and background, following Joyce. Consider a finite partition $\boldsymbol{X} = \langle X_1, X_2, \ldots, X_N\rangle$ of propositions. Our agent's degrees of belief are represented by a *credence function* $\boldsymbol{b}$ (not necessarily a probability function) that assigns a real number $\boldsymbol{b}(X)$ in [0, 1] to each $X \in \boldsymbol{X}$. The N-dimensional cube $\mathcal{B}_X = [0, 1]^N$ then contains all credence functions defined on $\boldsymbol{X}$. $\mathcal{P}_X$ is the set of all probability functions defined on $\boldsymbol{X}$, which properly includes the set $\mathcal{V}_X$ of all consistent truth-value assignments to elements of $\boldsymbol{X}$—I will call these *worlds*. We may define a *scoring rule* $\boldsymbol{S}$ on $\boldsymbol{X}$ which, for each $\boldsymbol{b}$ in $\mathcal{B}_X$ and $\boldsymbol{v}$ in $\mathcal{V}_X$, assigns a real number $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v}) \geq 0$ measuring the *epistemic disutility* of

[20] I thank Selim Berker and Justin Fisher for independently suggesting this point. It inspired the similar point in the previous section, too.

[21] I had completed a draft of this paper before I had a chance to see Joyce's new article, with its new theorem. I thank him for sending it to me.

holding the credences $\boldsymbol{b}$ when the truth-values across $\boldsymbol{X}$ are given by $\boldsymbol{v}$. As in golf, higher scores are worse! In the special case in which accuracy is the only dimension along which epistemic disutility is measured, $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v})$ measures the inaccuracy of $\boldsymbol{b}$ when the truth-values are given by $\boldsymbol{v}$. Say that a credence function $\boldsymbol{b}$ in $\boldsymbol{B}_X$ is *strongly dominated by* $\boldsymbol{b}^*$ *according to* $\boldsymbol{S}$ if $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v}) > \boldsymbol{S}(\boldsymbol{b}^*, \boldsymbol{v})$ for every world $\boldsymbol{v}$, and say that $\boldsymbol{b}$ is *weakly dominated by* $\boldsymbol{b}^*$ *according to* $\boldsymbol{S}$ if $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v}) \geq \boldsymbol{S}(\boldsymbol{b}^*, \boldsymbol{v})$ for every $\boldsymbol{v}$ and $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v}) > \boldsymbol{S}(\boldsymbol{b}^*, \boldsymbol{v})$ for some $\boldsymbol{v}$. In English: $\boldsymbol{b}$ is weakly dominated by $\boldsymbol{b}^*$ according to $\boldsymbol{S}$ iff $\boldsymbol{b}$ scores at least as badly as $\boldsymbol{b}^*$ in every world, and strictly worse in at least one world, according to $\boldsymbol{S}$. Say that $\boldsymbol{b}$ *is weakly dominated according to* $\boldsymbol{S}$ if there is some function $\boldsymbol{b}^*$ such that $\boldsymbol{b}$ is weakly dominated by $\boldsymbol{b}^*$ according to $\boldsymbol{S}$.

Being weakly dominated according to a reasonable scoring rule is an undesirable property of a function: holding that function is apparently precluded, since there is another function that is guaranteed to do no worse, and that could do better, by $\boldsymbol{S}$'s lights.[22] The constraint of *coherent admissibility* on a reasonable scoring rule $\boldsymbol{S}$ is that $\boldsymbol{S}$ will never attribute the undesirable property of being weakly dominated to a *coherent* credence function:

> Coherent Admissibility. No coherent credence function is weakly dominated according to $\boldsymbol{S}$.

The constraint of *truth-directedness* on $\boldsymbol{S}$ is that $\boldsymbol{S}$ should favour a credence function over another at a world if the former's assignments are uniformly closer than the latter's to the truth values in that world:

> Truth Directedness. If $\boldsymbol{b}$*'s* assignments are uniformly closer than $\boldsymbol{c}$*'s* to the truth values according to $\boldsymbol{v}$, then $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v}) < \boldsymbol{S}(\boldsymbol{c}, \boldsymbol{v})$.

Now we can state the final theorem of Joyce's paper (this volume):

> Theorem. Let $\boldsymbol{S}$ be a scoring rule defined on a partition $\boldsymbol{X} = \langle X_n \rangle$. If $\boldsymbol{S}$ satisfies TRUTH DIRECTEDNESS and COHERENT ADMISSIBILITY, and if $\boldsymbol{S}(\boldsymbol{b}, \boldsymbol{v})$ is finite and continuous for all $\boldsymbol{b}$ in $\mathcal{B}_X$ and $\boldsymbol{v} \in \mathcal{V}_X$, then

(i). every incoherent credence function is strongly dominated according to $\boldsymbol{S}$ and, moreover, is strongly dominated by some coherent credence function, and
(ii). no coherent credence function is weakly dominated according to $\boldsymbol{S}$.

(i). is the counterpart to the original gradational accuracy theorem, and it is striking that the domination of incoherent functions follows from the non-domination of coherent functions, and seemingly weak further assumptions. (ii). entails the converse theorem that I have contended was missing from Joyce's (1998) argument.

So does this lay the matter to rest, and give us a compelling argument for probabilism? Perhaps not. For (ii). *just is* the constraint of Coherent Admissibility on scoring rules—the rules have been *pre-selected* to ensure that they favour coherent

[22] I say 'apparently', because sub-optimal options are sometimes acceptable. Consider cases in which there are infinitely many options, with no maximal option—e.g. the longer one postpones the opening of Pollock's (1983) 'EverBetter wine', the better it gets. And a satisficing conception of what rationality demands might permit sub-optimality even when optimality can be achieved.

credence functions. In short, Coherent Admissibility is question-begging with respect to the converse theorem. (By contrast, the other constraints of Truth Directedness, continuity, and finiteness are not.) Another way to see this is to introduce into the debate Mr. Incoherent, who insists that credences are rationally required to *violate* the probability calculus. Imagine him imposing the mirror-image constraint on scoring rules:

> Incoherent Admissibility. No incoherent credence function is weakly dominated according to *S*.

Then (i). would be rendered false, and of course the mirror-image of (ii). would be trivially true, since it *just is* the constraint of Incoherent Admissibility. From a neutral standpoint, which prejudges the issue in favour of neither coherence nor incoherence, offhand it would appear that Incoherent Admissibility is on all fours with Coherent Admissibility.

So how would Joyce convince Mr. Incoherent that Coherent Admissibility is the correct constraint to impose, and not Incoherent Admissibility, using premises that they ought to share? Joyce offers the following argument that *any* coherent credence function can be rationally held (under suitable conditions), and that this in turn limits which scoring rules are acceptable. He maintains that it is 'plausible' that

> there are conditions under which *any* coherent credence function can be rationally held... After all, for any assignment of probabilities $\langle p_n \rangle$ to $\langle X_n \rangle$ it seems that a believer could, in principle, have evidence that justifies her in thinking that each X_n has p_n as its objective chance. Moreover, this could exhaust her information about *X*'s truth-value. According to the 'Principal Principle' of Lewis (1980), someone who knows that the objective chance of X_n is p_n, and who does not possess any additional information that is relevant to questions about X_n's truth-value, should have p_n as her credence for X_n. Thus, $\langle p_n \rangle$ is the rational credence function for the person to hold under these conditions. In light of this, one might argue, the following restriction on scoring rules should hold:
>
> Minimal Coherence: An epistemic scoring rule should never preclude, *a priori*, the holding of any coherent set of credences.

(263–297, this volume).

So we have here a putative reason to impose Coherent Admissibility on a reasonable scoring rule *S*. It obviates the putatively unacceptable situation in which a coherent credence function is precluded, insofar as it is weakly dominated according to *S*—unacceptable, since for any coherent assignment of credences, one could have evidence that it corresponds to the objective chances. To complete the argument, we apparently have no parallel reason for imposing Incoherent Admissibility on *S*, for one could *not* have evidence that an incoherent assignment of credences corresponds to the objective chances.

However, is not clear to me that any assignment of probabilities could correspond to the objective chances, still less that one could have evidence for any particular assignment that this is the case. There may necessarily be chance gaps to which some other probability functions could nevertheless assign values. For example, propositions *about* the chance function (at a time) might be 'blind spots' to the function itself but could be assigned probabilities by some other function. Perhaps there are no higher-order chances, such as the chance of: ⌜the chance of Heads is 1/2⌝, even though ⌜the chance of Heads is 1/2⌝ is a proposition, and thus fit to be assigned a value by *some* probability functions.[23] Or perhaps such higher-order chances are defined, but they are necessarily 0 or 1; and yet some other probability function could easily assign an intermediate value to 'the chance of Heads is 1/2'.

Moreover, it *is* clear to me that *not* any coherent credence function can be rationally held. For starters, *any coherent credence function that violates the Principal Principle* cannot be—and presumably Joyce agrees, given his appeal to it in his very argument. Indeed, if there are any constraints on rational credences that go beyond the basic probability calculus, then coherent violations thereof are counterexamples to Joyce's opening claim in the quoted passage. Choose your favourite such constraint—the Reflection Principle, or regularity, or the principle of indifference, or what have you. My own favourite is a prohibition on Moore paradoxical credences, such as my assigning high credence to '*p* & my credence in *p* is low' or to '*p* & I don't assign high credence to *p*'. If *epistemically rational* credence is more demanding than *coherent* credence, then there will be coherent credences that are rationally precluded. More power to an epistemic scoring rule, I say, if it precludes the holding of them!

So I am not persuaded by this defence of the Coherent Admissibility constraint, as stated. And to the extent that one is moved by this defence, it would seem to provide a more direct argument for coherence—from the coherence of chances and the Principal Principle—without any appeal to scoring rules.

Now, perhaps a slight weakening of the constraint *can* be justified along the lines of Joyce's argument. After all, some of the problematic cases that I raised involved *higher-order probability* assignments of one kind or another (higher order chances, the Principal Principle, the Reflection Principle, and the prohibition on Moore paradoxical credences), and the others (regularity and the principle of indifference) are quite controversial. So perhaps Joyce's argument goes through if we restrict our attention to partitions $\langle X_n \rangle$ of *probability-free* propositions, and to purely first-order probability assignments to them.[24] Then it seems more plausible that any coherent assignment of credences across such a partition should be admissible, and that a scoring rule that ever judges such an assignment to be weakly dominated is unreasonable.

The trouble is that then Joyce would seem to lose his argument *for probabilism* tout court, as opposed to a watered-down version of it. Probabilism says that *all* credences are rationally required to conform to the probability calculus—not merely that credences *in probability-free propositions* are so required. Consider, then, a credence function that is coherent over probability-free propositions, but that is wildly incoherent over higher-order propositions. It is obviously defective by probabilist lights, but the concern is that its defect will go undetected by a scoring rule that is confined to probability-free propositions. And how probative are scoring rules that

[23] Thanks here to Kenny Easwaran and (independently) Michael Titelbaum.

[24] Thanks here to Jim Joyce and (independently) Kenny Easwaran.

are so confined, when an agent's total epistemic state is *not* so confined, and should be judged in its entirety?

6 Conclusion

I began by confessing my schizophrenic attitude to probabilism. I have argued that the canonical statements of the major arguments for it have needed some repairing. Why, then, am I sympathetic to it at the end of the day, or at least at the end of some days? Partly because I think that to some extent the arguments *can* be repaired, and I have canvassed some ways in which this can be done, although to be sure, I think that some other problems remain. To the extent that they can be repaired, they provide a kind of triangulation to probabilism. And once we get to probabilism, it provides us with many fruits. Above all, it forms the basis of a unified theory of decision and confirmation—it combines seamlessly with utility theory to provide a fully general theory of rational action, and it illuminates or even resolves various hoary paradoxes in confirmation theory.[25] I consider that to be the best argument for probabilism. Sometimes, though, I wonder whether it is good enough.

So on Mondays, Wednesdays and Fridays, I call myself a probabilist. But as I write these words, today is Saturday.[26]

References

Adams, Ernest W. (1962), "On Rational Betting Systems", *Archive für Mathematische Logik und Grundlagenforschung* 6, 7–29, 112–128.

Adams, Ernest W. and Roger D. Rosenkrantz (1980), "Applying the Jeffrey Decision Model to Rational Betting and Information Acquisition", *Theory and Decision* 12, 1–20.

Armendt, Brad (1992), "Dutch Strategies for Diachronic Rules: When Believers See the Sure Loss Coming", *PSA 1992*, vol. 1, eds. D. Hull, M. Forbes, K. Okruhlik, East Lansing: Philosophy of Science Association, 217–229.

Armendt, Brad (1993), "Dutch Books, Additivity and Utility Theory", *Philosophical Topics* 21, No. 1, 1–20.

Arntzenius, Frank, Adam Elga, and John Hawthorne (2004), "Bayesianism, Infinite Decisions, and Binding", *Mind* 113, 251–83.

Baillie, Patricia (1973), "Confirmation and the Dutch Book Argument", *The British Journal for the Philosophy of Science* 24, 393–397.

Boolos, George (1994), "Gödel's Second Incompleteness Theorem Explained in Words of One Syllable", *Mind* 103, 1–3.

Carnap, Rudolf (1950), *Logical Foundations of Probability*, Chicago: University of Chicago Press.

Carnap, Rudolf (1955), *Notes on Probability and Induction* (UCLA, Philosophy 249, Fall Semester 1955), Unpublished Manuscript, Carnap Archive, University of Pittsburgh, RC-107-23-01.

---

[25] See Earman (1992), Jeffrey (1992), and Howson and Urbach (1993), among others.

[26] For very helpful comments I am grateful to: Selim Berker, David Chalmers, James Chase, John Cusbert, Lina Eriksson, James Ladyman, Aidan Lyon, Ralph Miles, Katie Steele, Michael Titelbaum, and especially Kenny Easwaran (who also suggested the name 'Czech Book'), Justin Fisher, Franz Huber, Carrie Jenkins, Jim Joyce, and Andrew McGonigal.

Christensen, David (1991), "Clever Bookies and Coherent Beliefs." *The Philosophical Review* C, No. 2, 229–247.

Christensen, David (1996), "Dutch-Book Arguments Depragmatized: Epistemic Consistency for Partial Believers", *The Journal of Philosophy*, 450–479.

Christensen, David (2001), "Preference-Based Arguments for Probabilism." *Philosophy of Science* 68 (3): 356–376.

de Finetti, Bruno (1980), "Probability: Beware of Falsifications", in *Studies in Subjective Probability*, eds. H. E. Kyburg, Jr. and H. E. Smokler, New York: Robert E. Krieger Publishing Company.

Döring, Frank (2000), "Conditional Probability and Dutch Books", *Philosophy of Science* 67 (September), 39, 1–409.

Douven, Igor (1999), "Inference to the Best Explanation Made Coherent", *Philosophy of Science* 66 (Proceedings), S424–S435.

Earman, John (1992), *Bayes or Bust?*, Cambridge, MA: MIT Press.

Eriksson, Lina and Alan Hájek (2007), "What Are Degrees of Belief?", *Studia Logica* 86, July, (Formal Epistemology I), 185–215, ed. Branden Fitelson.

Gillies, Donald (2000), *Philosophical Theories of Probability*, New York: Routledge.

Hájek, Alan (2005), "Scotching Dutch Books?", *Philosophical Perspectives* 19 (issue on Epistemology), John Hawthorne (ed.), 139–51.

Hájek, Alan (MS), "A Puzzle About Partial Belief".

Howson, Colin and Peter Urbach (1993),*Scientific Reasoning: The Bayesian Approach*. La Salle, IL: Open Court.

Jackson, Frank and Robert Pargetter (1976), "A Modified Dutch Book Argument", *Philosophical Studies* 29, 403–407.

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

Jeffrey, Richard (1992), *Probability and the Art of Judgment*, Cambridge: Cambridge University Press.

Joyce, James M. (1998), "A Non-Pragmatic Vindication of Probabilism", *Philosophy of Science* 65, 575–603.

Joyce, James M. (2004), "Bayesianism", in Alfred R. Mele and Piers Rawlings, *The Oxford Handbook of Rationality*, Oxford: Oxford University Press.

Joyce, James M. (this volume), "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief", in *Degrees of Belief*, eds. Franz Huber and Christoph Schmidt-Petri, Springer.

Kaplan, Mark (1996), *Decision Theory as Philosophy*, Cambridge: Cambridge University Press.

Kemeny, John (1955), "Fair Bets and Inductive Probabilities", *Journal of Symbolic Logic*, 20: 263–273.

Kennedy, Ralph and Charles Chihara (1979), "The Dutch Book Argument: Its Logical Flaws, Its Subjective Sources", *Philosophical Studies* 36, 19–33.

Kyburg, Henry (1978), "Subjective Probability: Criticisms, Reflections and Problems", *Journal of Philosophical Logic* 7, 157–180.

Lange, Marc (1999), "Calibration and the Epistemological Role of Bayesian Conditionalization", *The Journal of Philosophy*, 294–324.

Lehman, R. (1955), "On Confirmation and Rational Betting", *Journal of Symbolic Logic* 20, 251–262.

Levi, Isaac (1974), "On Indeterminate Probabilities", *Journal of Philosophy*71, 391–418.

Lewis, David (1980), "A Subjectivist's Guide to Objective Chance", in *Studies in Inductive Logic and Probability*, Vol II., ed. Richard C. Jeffrey, University of California Press.

Maher, Patrick (1993), *Betting on Theories*, Cambridge: Cambridge University Press.

Maher, Patrick (2003), "Joyce's Argument for Probabilism", *Philosophy of Science* 69, 73–81

Mellor, D. H. (1971), *The Matter of Chance*, Cambridge: Cambridge University Press.

Milne, Peter (1990), "Scotching the Dutch Book Argument", *Erkenntnis* 32, 105–126.

Moore, P. G. (1983), "A Dutch Book and Subjective Probabilities", *The British Journal for the Philosophy of Science* 34, 263–266.

Pollock, John (1983), "How Do You Maximize Expectation Value?", *Nous* 17/3 (September), 409–421.
Ramsey, Frank P., (1980/1931), "Truth and Probability", in Foundations of *Mathematics and other Essays*, R. B. Braithwaite (ed.), Routledge & P. Kegan 1931, 156–198; reprinted in *Studies in Subjective Probability*, H. E. Kyburg, Jr. and H. E. Smokler (eds.), 2nd ed., R. E. Krieger Publishing Company, 1980, 23–52; reprinted in *Philosophical Papers*, D. H. Mellor (ed.) Cambridge: Cambridge University Press, 1990.
Rosenkrantz, Roger D. (1981), *Foundations and Applications of Inductive Probability*, Atascadero, CA: Ridgeview Publishing Company.
Ryder, J. M. (1981), "Consequences of a Simple Extension of the Dutch Book Argument", *The British Journal for the Philosophy of Science* 32, 164–167.
Schick, Frederic (1986), "Dutch Bookies and Money Pumps", *Journal of Philosophy* 83, 112–119.
Seidenfeld, Teddy (1985), "Calibration, Coherence, and Scoring Rules", *Philosophy of Science* 52, No. 2 (June), 274–294.
Seidenfeld, Teddy and Mark J. Schervish (1983), "A Conflict Between Finite Additivity and Avoiding Dutch Book", *Philosophy of Science* 50, 398–412.
Shimony, Abner (1955), "Coherence and the Axioms of Confirmation", *Journal of Symbolic Logic* 20, 1–28.
Skyrms, Brian (1980), "Higher Order Degrees of Belief", in *Prospects for Pragmatism*, ed. D. H. Mellor, Cambridge: Cambridge University Press, 109–37.
Skyrms, Brian (1984), *Pragmatics and Empiricism*, New Haven, CT: Yale University.
Skyrms, Brian (1986), *Choice and Chance*, 3rd ed., Belmont, CA: Wadsworth.
Skyrms, Brian (1987), "Coherence", in N. Rescher (ed.), *Scientific Inquiry in Philosophical Perspective*, Pittsburgh, PA: University of Pittsburgh Press, 225–242.
van Fraassen, Bas (1989), *Laws and Symmetry*, Oxford: Clarendon Press.
Vineberg, Susan (2001), "The Notion of Consistency for Partial Belief", *Philosophical Studies* 102 (February), 281–96.
Waidacher, C. (1997), "Hidden Assumptions in the Dutch Book Argument", *Theory and Decision* 43 (November), 293–312.
Weatherson, Brian (1999), "Begging the Question and Bayesians", *Studies in History and Philosophy of Science* 30A, 687–697.
Williamson, Jon (1999), "Countable Additivity and Subjective Probability", *The British Journal for the Philosophy of Science* 50, No. 3, 401–416.
Zynda, Lyle (2000), "Representation Theorems and Realism About Degrees of Belief", *Philosophy of Science* 67, 45–69.