Imprecise Bayesianism and Global Belief Inertia

Aron Vallinder

Abstract

Traditional Bayesianism requires that an agent's degrees of belief be represented by a real-valued, probabilistic credence function. However, in many cases it seems that our evidence is not rich enough to warrant such precision. In light of this, some have proposed that we instead represent an agent's degrees of belief as a set of credence functions. This way, we can respect the evidence by requiring that the set, often called the agent's credal state, includes all credence functions that are in some sense compatible with the evidence. One known problem for this evidentially-motivated imprecise view is that in certain cases, our imprecise credence in a particular proposition will remain the same no matter how much evidence we receive. In this paper I argue that the problem is much more general than has been appreciated so far, and that it's difficult to avoid it without compromising the initial evidentialist motivation.

1 Introduction

In the orthodox Bayesian framework, agents must have precise degrees of belief, in the sense that these degrees of belief are represented by a real-valued credence function. This may seem implausible in several respects. In particular, one might think that our evidence is rarely rich enough to justify this kind of precision—choosing one number over another as our degree of belief will often be an arbitrary decision with no basis in the evidence. For this reason, Joyce ([2010]) suggests that we should

represent degrees of belief by a set of credence functions instead.[1] This way, we can avoid arbitrariness by requiring that the set contains all credence functions that are, in some sense, compatible with the evidence. However, this requirement creates a new difficulty. The more limited our evidence is, the greater the number of credence functions compatible with it will be. In certain cases, the number of compatible credence functions will be so vast that the range of our credence in some propositions will remain the same no matter how much evidence we subsequently go on to obtain. This is the problem of belief inertia. Joyce is willing to accept this implication, but I will argue that the phenomenon is much more widespread than he seems to realize, and that there is therefore decisive reason to abandon his view.

In the next section, I introduce the traditional Bayesian formalism and provide some reason for thinking that its precision may be problematic. In Section 3, I present Joyce's preferred alternative—imprecise Bayesianism—and attempt to spell out its underlying evidentialist motivation. In particular, I suggest an account of what it means for a credence function to be compatible with a body of evidence. After that, in Section 4, I introduce the problem of belief inertia via an example from Joyce. I also prove that one strategy for solving the problem (suggested but not endorsed by Joyce) is unsuccessful. Section 5 argues that the problem is far more general than one might think when considering Joyce's example in isolation. The argument turns on the question of what prior credal state an evidentially motivated imprecise Bayesian agent should have. I maintain that, in light of her motivation for rejecting precise Bayesianism, her prior credal state must include all credence functions that satisfy some very weak constraints. However, this means that the problem of belief inertia is with us from the very start, and that it affects almost all of our beliefs. Even those who are willing to concede certain instances of belief inertia should find this general version unacceptable. Finally, in Section 6 I consider a few different ways for an imprecise Bayesian to respond. The upshot is that we must give up the very strong form of evidentialism and allow that the choice of prior credal state is to a large extent subjective. However, this move greatly decreases the imprecise Bayesian's dialectical advantage over the precise subjective Bayesian.

2   Precision and Its Problems

Traditional Bayesianism, as I will understand it here, makes the following two normative claims:

> Probabilism
> A rational agent's degrees of belief are represented by a credence function

---

[1] Although Joyce is my main target in this essay, the view is of course not original to him. For an influential early exponent, see Levi ([1980]).

c which assigns a real number $c(P)$ to each proposition P in some Boolean algebra Ω. The credence function c respects the axioms of probability theory:

1. $c(P) \geq 0$ for all $P \in \Omega$.
2. If $\top$ is a tautology, then $c(\top) = 1$.
3. If P and Q are logically incompatible, then $c(P \vee Q) = c(P) + c(Q)$.

Conditionalization
A rational agent updates her degrees of belief over time by conditionalizing her credence function on all the evidence she has received. If E is the strongest proposition an agent with credence function c_0 at t_0 learns between t_0 and t_1, then her new credence function c_1 is given as $c_1(\cdot) = c_0(\cdot \mid E)$.

Some philosophers within the Bayesian tradition have taken issue with the precision required by probabilism. For one thing, it may appear descriptively inadequate. It seems implausible to think that flesh-and-blood human beings have such fine-grained degrees of belief.[2] However, even if this psychological obstacle could be overcome, Joyce ([2010]) argues that precise probabilism should be rejected on normative grounds, because our evidence is rarely rich enough to justify having precise credences. His point is perhaps best appreciated by way of example. Consider the following case, adapted from (Bradley [unpublished]).

Three Urns
There are three urns in front of you, each of which contains a hundred marbles. You are told that the first urn contains fifty black and fifty white marbles, and that all marbles in the second urn are either black or white, but you don't know their ratio. You are given no further information about marble colours in the third urn. For each urn i, what credence should you have in the proposition B_i that a marble drawn at random from that urn will be black?

Here I will understand a random draw simply as one where each marble in the urn has an equal chance of being drawn. That makes the first case straightforward. We know that there are as many black marbles as there are white ones, and that each of them has an equal chance of being drawn. Hence we should apply some chance-credence principle and set $c(B_1) = 0.5$.[3] The second case is not so clear-cut.

---

[2] Whether this is implausible will depend on what kind of descriptive claim one thinks is involved in ascribing a precise degree of belief to an agent. See for instance (Meacham and Weisberg [2011]).

[3] Hardcore subjectivists may insist that, even in this case, any probabilistically coherent credence assignment is permissible.

Some will say that any credence assignment is permissible, or at least that a wide range of them are. Others will again try to identify a unique credence assignment as rationally required, typically via an application of the principle of indifference. They will claim that we have no reason to consider either black or white as more likely than the other, and that we should therefore give them equal consideration by setting $c(B_2) = 0.5$. However, as is well-known, the principle of indifference gives inconsistent results depending on how we partition the space of possibilities.[4] This becomes even more evident when we consider the third urn. In the first two cases we knew that all marbles were either black or white, but now we don't even have that piece of information. So in order to apply the principle of indifference, we must first settle on a partition of the space of possible colours. If we settle on the partition {black, not black}, the principle of indifference gives us $c(B_3) = 0.5$. If we instead think that the partition is given by the eleven basic colour terms of the English language, the principle of indifference tells us to set $c(B_3) = 1/11$.

How can we determine which partition is appropriate? In some problem cases, the principle's adherents have come up with ingenious ways of identifying a privileged partition.[5] However, Joyce ([2005], p. 170) argues that even if this could be done across the board (which seems doubtful), the real trouble runs deeper. The principle of indifference goes wrong by always assigning precise credences, and hence the real culprit is (precise) probabilism. In the first urn case, our evidence is rich enough to justify a precise credence of 0.5. But in the second and third cases, our evidence is so limited that any precise credence would constitute a leap far beyond the information available to us. Adopting a precise credence in these cases would amount to acting as if we have evidence we simply do not possess, regardless of whether that precise credence is based merely on personal opinion, or whether it has been derived from some supposedly objective principle.

The lesson Joyce draws from this example is therefore that we should only require agents to have imprecise credences. This way we can respect our evidence even when that evidence is ambiguous, partial, or otherwise limited. My target in this paper will be this sort of evidentially motivated imprecise Bayesianism. In the next section I present the view and clarify the evidentialist argument for adopting it.

3 Imprecise Bayesianism and Respecting Ambiguous Evidence

Joyce's ([2010], p. 287) imprecise Bayesianism makes the following two normative claims:

---

[4] Widely discussed examples include Bertrand's ([1889]) paradox, and van Fraassen's ([1989]) cube factory.
[5] See for example (Jaynes [1973]).

Imprecise Probabilism
A rational agent's degrees of belief are represented by a credal state C, which is a set of credence functions. Each $c \in C$ assigns a real number $c(P)$ to each proposition P in some Boolean algebra Ω. Furthermore, each $c \in C$ respects the axioms of probability theory.

Imprecise Conditionalization
A rational agent updates her credal state over time by conditionalizing each of its elements on all the evidence she has received. If E is the strongest proposition an agent with credal state C_0 at t_0 learns between t_0 and t_1, then her new credal state C_1 is given as $C_1 = \{c_0(\cdot \mid E) : c_0 \in C_0\}$.[6]

Each individual credence function thus behaves just like the credence functions of precise Bayesianism: they are probabilistic, and they are updated by conditionalization. The difference is only that the agent's degrees of belief are now represented by a set of credence functions, rather than a single one. As a useful terminological shorthand, I will write $C(P)$ for the set of numbers assigned to the proposition P by the elements of C, so that $C(P) = \{x : \exists c \in C \text{ s.t. } c(P) = x\}$. I will refer to $C(P)$ simply as the agent's credence in P.

Agents with precise credences are more confident in a proposition P than in another proposition Q if and only if their credence function assigns a greater value to P than to Q. In order to be able to make similar comparisons for agents with imprecise credences, we will adopt what I take to be the standard, supervaluationist, view and say that an imprecise believer is determinately more confident in P than in Q if and only if $c(P) > c(Q)$ for each $c \in C$. If there are $c, c' \in C$ such that $c(P) > c(Q)$ and $c'(P) < c'(Q)$, it is indeterminate which of the two propositions she regards as more likely. In general, any claim about her overall doxastic state requires unanimity among all the credence functions in order to be determinately true or false.[7]

Now, Joyce defends imprecise Bayesianism on the grounds that many evidential situations do not warrant precise credences. With his framework in place, we can respect the datum that a precise credence of 0.5 is the correct response in the first

---

[6] As stated, the update rule doesn't tell us what to do if an element of the credal state assigns zero probability to a proposition that the agent later learns. This problem is of course familiar from the precise setting. Three options suggest themselves: (i) discard all such credence functions from the posterior credal state, (ii) require that each element of the credal state the regularity principle, so that they only assign zero to doxastically impossible propositions, thereby ensuring that the situation can never arise, or (iii) introduce a primitive notion of conditional probability. For my purposes, we don't need to settle on a solution. I'll just assume that the imprecise Bayesian has some satisfactory way of dealing with these cases.

[7] This supervaluationist view of credal states is endorsed by Joyce ([2010]), van Fraassen ([1990]), and Hájek ([2003]), among others.

urn case, without thereby being forced to assign precise credences in the second and third cases as well. In these last two cases, our evidence is ambiguous or partial, and assigning precise credences would require making a leap far beyond the information available to us. This raises the question of how far in the direction of imprecision we should move in order to remain on the ground. How many credence functions must we include in our credal state before we can be said to be faithful to our evidence? Joyce answers that we should include just those credence functions that are ompatible with our evidence.[8] We can state this as:

> Evidence Grounding Thesis
> At any point in time, a rational agent's credal state includes all and only those credence functions that are compatible with the total evidence she possesses at that time.

To unpack this principle, we need a substantive account of what it takes for a credence function to be compatible with a body of evidence. One such proposal is due to (White [2010], p. 174):

> Chance Grounding Thesis
> Only on the basis of known chances can one legitimately have sharp credences. Otherwise one's spread of credence should cover the range of possible chance hypotheses left open by your evidence.

The chance grounding thesis posits a very tight connection between credence and chance. As Joyce ([2010], p. 289) points out, the connection is indeed too tight, in at least one respect. There are cases where all possible chance hypotheses are left open by our evidence, but where we should nevertheless have sharp (precise) credences. He provides the following example.

> Symmetrical Biases
> Suppose that an urn contains coins of unknown bias, and that for each coin of bias α there is another coin of bias $(1 - \alpha)$. One coin has been chosen from the urn at random. What credence should we have in the proposition H, that it will come up heads on the first flip?

Because the chance of heads corresponds to the bias of the chosen coin (whatever it is), and since (for all we know) the chosen coin could have any bias, every possible chance hypothesis is left open by the evidence. In this setup, for each $c \in C$,

---

8   Joyce writes ([2010], p. 288) that each element of the credal state is a probability function that the agent takes to be compatible with her evidence. This formulation leaves it open whether compatibility is meant to be an objective or a subjective notion; we will return to this issue later.

the credence assignment $c(H)$ is given as the expected value of a corresponding probability density function (pdf), f_c, defined over the possible chance hypotheses: $c(H) = \int_0^1 x \cdot f_c(x)\,dx$. The information that, for any α, there are as many coins of bias α as there are coins of bias $(1 - \alpha)$ translates into the requirement that for each $a, b \in [0, 1]$ and for every f_c,

$$\int_a^b f_c(x)\,dx = \int_{1-b}^{1-a} f_c(x)\,dx. \tag{1}$$

Any f_c which satisfies this constraint will be symmetrical around the midpoint, and will therefore have an expected value of 0.5. This means that $c(H) = 0.5$ for each $c \in C$. Thus we have a case where all possible chance hypotheses are left open by the evidence, but where we should still have a precise credence.[9]

Nevertheless, something in the spirit of the chance grounding thesis looks like a natural way of unpacking the evidence grounding thesis. In Joyce's example, each possible chance hypothesis is indeed left open by the evidence, but we do know that every pdf f_c must satisfy constraint (**??**) for each $a, b \in [0, 1]$. So any f_c which doesn't satisfy this constraint will be incompatible with our evidence. And similarly for any other constraints our evidence might impose on f_c. In the case of a known chance hypothesis, the only pdf compatible with the evidence will be the one that assigns all weight to that known chance value. Similarly, if the chance value is known to lie within some particular range, then the only pdfs compatible with the evidence will be those that are equal to zero everywhere outside of that range. However, as Joyce's example shows, these are not the only ways in which our evidence can rule out pdfs. More generally, evidence can constrain the shape of the compatible pdfs. In light of this, we can propose the following revision.

> Revised Chance Grounding Thesis
> A rational agent's credal state contains all and only those credence functions that are given as the expected value of some probability density function over chance hypotheses that satisfies the constraints imposed by her evidence.

Just like White's original chance grounding thesis, my revised formulation posits

---

[9] An anonymous referee suggested that it might make a difference whether the coin that is to be flipped has been chosen yet or not. If it has not yet been chosen, a precise credence of 0.5 seems sensible in light of one's knowledge of the setup. If instead it has already been chosen, then it has a particular bias, and since the relevant symmetry considerations are no longer in play, one's credence should be maximally imprecise: $[0, 1]$. However, one might argue that rationally assigning a precise credence of 0.5 when the coin has not yet been chosen does not constitute a counterexample to the original chance grounding thesis, by arguing that the proposition 'The next coin to be flipped will come up heads' has an objective chance of 0.5. My argument won't turn on this, so I'm happy to go along with Joyce and accept that we have a counterexample to the chance grounding thesis.

an extremely tight connection between credence and chance. For any given body of evidence, it leaves no freedom in the choice of which credence functions to include in one's credal state. Because of the way compatibility is understood, there will always be a fact of the matter about which credence functions are compatible with one's evidence, and hence about which credence functions ought to be included in one's credal state.

The question, then, is whether we should settle on this formulation, or whether we can change the requirements without thereby compromising the initial motivation for the imprecise model. In his discussion of the chance grounding thesis, Joyce ([2010], p. 288) claims that even when the error in White's formulation has been taken care of, as I proposed to do with my revision, the resulting principle is not essential to the imprecise proposal. Instead, he thinks it is merely the most extreme view an imprecise Bayesian might adopt. Now, this is certainly correct as a claim about imprecise Bayesianism in general. One can accept both imprecise probabilism and imprecise conditionalization without accepting any claim about how knowledge of chance hypotheses, or any other kind of evidence, should constrain which credence functions are to be included in the credal state. However, on the evidentially motivated proposal that Joyce advocates himself, it's not clear whether any other way of specifying what it means for a credence function to be compatible with one's evidence could be defended.

One worry you might have about the revised chance grounding thesis is that far from all constraints on rational credence assignments appear to be mediated by information about chance hypotheses. In many cases, our evidence seems to rule out certain credence assignments as irrational, even though it's difficult to see which chance hypotheses we might appeal to in explaining why this is so. Take for instance the proposition that my friend Jakob will have the extraordinarily spicy *phaal* curry for dinner tonight. I know that he loves spicy food, and I've had *phaal* with him a few times in the past year. In light of my evidence, some credence assignments seem clearly irrational. A value of 0.001 certainly seems too low, and a value of 0.9 certainly seems too high. However, we don't normally think of our credence in propositions of this kind as being constrained by information about chances. If this is correct, then the revised chance grounding thesis can at best provide a partial account of what it takes for a body of evidence to rule out a credence assignment as irrational. Of course, one could insist that we do have some information about chances which allows us to rule out the relevant credence assignments, but such an idea would have to be worked out in a lot more detail before it could be made plausible. Alternatively, one could simply deny my claim that these credence assignments would be irrational. However, as we'll soon discover, that response would merely

strengthen my objection.[10]

Going forward, I will assume that the evidence grounding thesis holds, so that a rational agent's credal state should include all and only those credence functions that are compatible with her total evidence. I will also assume that this notion of compatibility is an objective one, so that there is always a fact of the matter about which credence functions are compatible with a given body of evidence. However, I will not assume any particular understanding of compatibility, such as those provided by White's chance grounding thesis or my revised formulation. As we'll see, these assumptions spell trouble for the imprecise Bayesian. I will therefore revisit them in Section 6, to see whether they can be given up.

4   Local Belief Inertia

In certain cases, evidentially-motivated imprecise Bayesianism makes inductive learning impossible. Joyce already recognizes this, but I will argue that the implications are more wide-ranging and therefore more problematic than has been appreciated so far.[11] To illustrate the phenomenon, consider an example adapted from (Joyce [2010], p. 290).

> Unknown Bias
>
> A coin of unknown bias is about to be flipped. What is your credence $C(H_1)$ that the outcome of the first flip will be heads? And after having observed n flips, what is your credence that the coin will come up heads on the $(n+1)$th flip?

As in the Symmetrical Biases example discussed earlier, each $c \in C$ is here given as the expected value of a corresponding probability density function, f_c, over the possible chance hypotheses. We are not provided with any evidence that bears on the

---

[10]   Another case where it's not immediately clear how to apply the revised chance grounding thesis is propositions about past events. On what I take to be the standard view, such propositions have an objective chance of either 1 or 0, depending on whether they occurred or not (see for instance (Schaffer [2007])). So for a proposition P about an event that is known to be in the past, the only chance hypotheses left open by the evidence are (at most) 0 and 1. However, in certain cases, this will be enough to give us maximal imprecision. If we have no knowledge of what the chance of P was prior to the event's occurring (or not occurring), then it seems that any way of distributing credence across these two chance hypotheses will be compatible with our evidence, and hence that the credal state will include a credence function c with $c(P) = x$ for each $x \in [0, 1]$. Indeed, if we accept Levi's ([1980], chapter 9) credal convexity requirement, then whenever the credal state includes 0 and 1, it will also include everything in between. A further worry, which I will set aside here, is whether we can have any non-trivial objective chances if determinism is true.

[11]   Joyce is of course not the first to recognize this. See for instance Walley's ([1991], p. 93) classic monograph for a discussion of how certain types of imprecise probability have difficulties with inductive learning.

question of whether the first outcome will be heads, and hence our evidence cannot rule out any pdfs as incompatible. In turn, this means that no value of $c(H_1)$ can be ruled out, and therefore that our overall credal state with respect to this proposition will be maximally imprecise: $C(H_1) = (0, 1)$.[12] However, this starting point renders inductive learning impossible, in the following sense. Suppose that you observe the coin being flipped a thousand times, and see 500 heads and 500 tails. This looks like incredibly strong evidence that the coin is very, very close to fair, and would seem to justify concentrating your credence on some fairly narrow interval around 0.5. However, although each element of the credal state will indeed move toward the midpoint, there will always remain elements on each extreme. Indeed, for any finite sequence of outcomes and for any $x \in (0, 1)$, there will be a credence function $c \in C$ which assigns a value of x to the proposition that the next outcome will be heads, conditional on that sequence. Thus your credence that the next outcome will be heads will remain maximally imprecise, no matter how many observations you make.

Bradley ([2015]) calls this the problem of belief inertia. I will refer to it as local belief inertia, as it pertains to a limited class of beliefs, namely those about the outcomes of future coin flips. This is a troubling implication, but Joyce ([2010], p. 291) is willing to accept it:

> if you really know *nothing* about the [...] coin's bias, then you also really know *nothing* about how your opinions about [H_{n+1}] should change in light of frequency data. [...] You cannot learn anything in cases of pronounced ignorance simply because a prerequisite for learning is to have prior views about how potential data should alter your beliefs, but you have no determinate views on these matters at all.

Nevertheless, he suggests a potential way out for imprecise Bayesians who don't share his evidentialist commitments. The underlying idea is that we should be allowed to rule out those probability density functions that are especially biased in certain ways. Some pdfs are equal to zero for entire subintervals (a, b), which means that they could never learn that the true chance of heads lies within (a, b). Perhaps we want to rule out all such pdfs, and only consider those that assign a non-zero value to every subinterval (a, b). Similarly, some pdfs will be extremely biased toward chance hypotheses that are very close to one of the endpoints, with the result that the corresponding credence functions will be virtually certain that the outcome will be heads, or virtually certain that the outcome will be tails, all on the basis of

---

[12] Joyce ([2010], p. 290) thinks we should understand maximal imprecision here to mean the open set $(0, 1)$ rather than the closed set $[0, 1]$, but it's not obvious on what basis we might rule out the two extremal probability assignments. At any rate, my objection won't turn on which of these is correct, as we'll see shortly.

no evidence whatsoever. Again, perhaps we want to rule these out, and require that each $c \in C$ assigns a value to H_1 within some interval (c_-, c^+), with $c_- > 0$ and $c^+ < 1$.

With these two restrictions in place, the spread of our credence is meant to shrink as we make more observations, so that after having seen 500 heads and 500 tails, it is centred rather narrowly around 0.5, thereby making inductive learning possible again. While recognizing this as an available strategy, Joyce does not endorse it himself, as it is contrary to the evidentialist underpinnings of his view. In any case, the strategy doesn't do the trick. Even if we could find a satisfactory motivation, it would not deliver the result Joyce claims it does, as the following theorem shows:

Theorem 1. Let the random variable X be the coin's bias for heads, and let the random variable Y_n be number of heads in the first n flips. For a given n, a given y_n, a given interval (c_-, c^+) with $c_- > 0$ and $c^+ < 1$, and a given $c_0 \in (c_-, c^+)$, there is a pdf, f_X, such that

1. $E[X] \in (c_-, c^+)$,

2. $E[X \mid Y_n = y_n] = c_0$, and

3. $\int_a^b f_X(x)\, dx > 0$ for every $a, b \in [0, 1]$ with $a < b$.

The first and third conditions are the two constraints that Joyce suggested we impose. The first ensures that the pdf is not extremely biased toward chance hypotheses that are very close to one of the endpoints, and the third ensures that it is non-zero for every subinterval (a, b) of the unit interval. The second condition corresponds to the claim that we still don't have inductive learning, in the sense that no matter what sequence of outcomes is observed, for every $c_0 \in (c_-, c^+)$, there will be a pdf whose expectation conditional on that sequence is c_0.

Proof. Consider the class of beta distributions. First, we will pick a distribution from this class whose parameters α and β are such that the first two conditions are satisfied. Now, the expectation and the conditional expectation of a beta distribution are respectively given as

$$E[X] = \frac{\alpha}{\alpha + \beta}, \text{ and } E[X \mid Y_n = y_n] = \frac{\alpha + y_n}{\alpha + \beta + n}.$$

The first two conditions now give us the following constraints on α and β:

$$c_- < \frac{\alpha}{\alpha + \beta} < c^+, \text{ and } \frac{\alpha + y_n}{\alpha + \beta + n} = c_0.$$

The first of these constraints gives us that

$$\frac{c_-}{1 - c_-}\beta < \alpha < \frac{c^+}{1 - c^+}\beta.$$

The second constraint allows us to express α as

$$\alpha = \frac{c_0(\beta + n) - y_n}{1 - c_0}.$$

Putting the two together, we get

$$\beta > \frac{(1 - c_-)(y_n - c_0 n)}{c_0 - c_-} \text{ and } \beta > \frac{(1 - c^+)(y_n - c_0 n)}{c_0 - c^+}.$$

As we can make β arbitrarily large, it is clear that for any given set of values for n, y_n, c_-, c^+ and c_0, we can find a value for β such that the two inequalities above hold. We have thus found a beta distribution that satisfies the first two conditions. Finally, we show that the third condition is met. The pdf of a beta distribution is given as

$$f_X(x) = \frac{1}{B(\alpha, \beta)}x^{\alpha-1}(1 - x)^{\beta-1},$$

where the beta function B is a normalization constant. As is evident from this expression, we will have $f_X(x) > 0$ for each $x \in (0, 1)$, which in turn implies that $\int_a^b f_X(x)\,dx > 0$ for every $a, b \in [0, 1]$ with $a < b$. Moreover, this holds for any values of the parameters α and β. Therefore every beta distribution satisfies the third condition, and our proof is done. □

What this shows is that all the work is being done by the choice of the initial interval. Although many credence functions will be able to move outside the interval in response to evidence, for every value inside the interval, there will always be a a credence function that takes that value no matter what sequence of outcomes has been observed. Thus the set of prior credence values will be a subset of the set of posterior credence values. The intuitive reason for this is that we can always find an initial probability density function which is sufficiently biased in some particular way to deliver the desired posterior credence value.

There are therefore two separate things going on in the unknown bias case, both of which might be thought worrisome: the problem of maximal imprecision, and the problem of belief inertia. As the result shows, Joyce's proposed fix addresses the former but not the latter, and our beliefs can therefore be inert without being maximally imprecise.[13] Granted, having a set of posterior credence values that always

---

[13] In turn, this explains why it doesn't matter whether we understand maximal imprecision to mean

includes the set of prior credence values as a subset is a less severe form of belief inertia than having a set of posterior credence values that is always identical to the set of prior credence values. However, even this weaker form of belief inertia means that no matter how much evidence the agent receives, she cannot converge on the correct answer with any greater precision than is already given in her prior credal state.

Now, Theorem 1 only shows that one particular set of constraints is insufficient to make inductive learning possible in the unknown bias case. Thus some other set of constraints could well be up to the job. For example, consider the set of beta distributions with parameters α and β such that $\beta/m \leq \alpha \leq m\beta$ for some given number m. If we let the credal state contain one credence function for each of these distributions, inductive learning will be possible.

It may be objected that we should regard belief inertia, made all the more pressing by Theorem 1, not as a problem for imprecise Bayesianism, but rather as a problem for an extreme form of evidentialism.[14] Suppose that a precise Bayesian says that all credences that satisfy the first and third conditions are permissible to adopt as one's precise credences. Theorem 1 would then tell us that it is permissible to change your credence by an arbitrarily small amount in response to any evidence. Although hardcore subjectivists would be happy to accept this conclusion, most others would presumably want to say that this constitutes a failure to respond appropriately to the evidence. Therefore, whatever it is that a precise moderate subjectivist would say to rule out such credence functions as irrational, the imprecise Bayesian could use the same account to explain why those credence functions should not be included in the imprecise credal state.

I agree that belief inertia is not an objection to imprecise Bayesianism as such: it becomes an objection only when that framework is combined with Joyce's brand of evidentialism. Nevertheless, I do believe the problem is worse for imprecise Bayesianism than it is for precise Bayesianism. On the imprecise evidentialist view, you are epistemically required to include all credence functions that are compatible with your evidence in your credal state. If we take Joyce's line and don't impose any further conditions, this means that, in the unknown bias case, you are epistemically required to adopt a credal state that is both maximally imprecise and inert. If we instead are sympathetic to the two further constraints, it means that you are epistemically required to adopt a credal state that will always include the initial interval from which you started as a subset. By contrast, on the precise evidentialist view, you are merely epistemically permitted to adopt one such credence function as your own. Of course, we may well think it's epistemically impermissible to adopt such credence functions. But a view on which we are epistemically required to include

---

(0, 1) or [0, 1]. Belief inertia will arise regardless of which of the two we choose.

[14] I'm grateful to an anonymous referee for drawing my attention to this point.

them in our credal state seems significantly more implausible.

A further difference is that any fixed beta distribution will eventually be pushed toward the correct distribution. Thus any precise credence function will eventually give us the right answer, even though this convergence may be exceedingly slow for some of them. By contrast, Theorem 1 shows that the initial interval (c_-, c^+) will always remain a subset of the imprecise Bayesian's posterior credal state. Therefore, belief inertia would again seem to be more of a problem for the imprecise view than for the precise view.

Finally, it's not at all obvious what principle a precise Bayesian might appeal to in explaining why the credence functions that intuitively strike us as insufficiently responsive to the evidence are indeed irrational. Existing principles provide constraints that are either too weak (for instance the principal principle or the reflection principle) or too strong (for instance the principle of indifference). It may well be possible to formulate an adequate principle, but to my knowledge this has not yet been done.

At any rate, Joyce is willing to accept local belief inertia in the unknown bias case, and his reasons for doing so may strike one as quite plausible. When one's evidence is so extremely impoverished, it might make sense to say that one doesn't even know which hypotheses would be supported by subsequent observations. This case is a fairly contrived toy example, and one might hope that such cases are the exception and not the rule in our everyday epistemic lives. So a natural next step is to ask how common these cases are. If it turns out that they are exceedingly common—as I will argue that they in fact are—then we ought to reject evidentially-motivated imprecise Bayesianism, even if we were initially inclined to accept particular instances of belief inertia.

5   From Local to Global Belief Inertia

I will argue that belief inertia is in fact very widespread. My strategy for establishing this conclusion will be to first argue that an imprecise Bayesian who respects the evidence grounding thesis must have a particular prior credal state, and second to show that any agent who starts out with this prior credal state and updates by imprecise conditionalization will have inert beliefs for a wide range of propositions.

In order for the Bayesian machinery—whether precise or imprecise—to get going, we must first have priors in place. In the precise case, priors are given by the credence function an agent adopts before she receives any evidence whatsoever. Similarly, in the imprecise case, priors are given by the set of credence functions an agent adopts as her credal state before she receives any evidence whatsoever.

The question of which constraints to impose on prior credence functions is a familiar and long-standing topic of dispute within precise Bayesianism. Hardcore sub-

jectivists hold that any probabilistic prior credence function is permissible, whereas objectivists wish to narrow down the number of permissible prior credence functions to a single one. In between these two extremes, we find a spectrum of moderate views. These more measured proposals suggest that we add some constraints beyond probabilism, without thereby going all the way to full-blown objectivism.

The same question may of course be asked of imprecise Bayesianism as well. In this context, our concern is with which constraints to impose on the set of prior credence functions. Hardcore subjectivists hold that any set of probabilistic prior credence functions is permissible, whereas objectivists will wish to narrow down the number of permissible sets of prior credence functions to a single one. In between these two extremes, we again find a spectrum of moderate views.

For an imprecise Bayesian who is motivated by evidential concerns, the answer to the question of priors should be straightforward. By the evidence grounding thesis, our credal state at a given time should include all and only those credence functions that are compatible with our evidence at that time. In particular, this means that our prior credal state should include all and only those credence functions that are compatible with the empty body of evidence. Thus, in order to determine which prior credal states are permissible, we must determine which credence functions are compatible with the empty body of evidence. As you'll recall, I assumed that the relevant notion of compatibility is an objective one. This means that there will be a unique set of all and only those credence functions that are compatible with the empty body of evidence.[15] Which credence functions are these?

In light of our earlier examples, we can rule out some credence functions from the prior credal state. In particular, we can rule out those that don't satisfy the principal principle. If we were to learn only that the chance of P is x, then any credence function that does not assign a value of x to P will be incompatible with our evidence. And given that the credal state is updated by conditionalizing each of its elements on all of the evidence received, it follows that we must have $c(P|\text{ch}(P) = x) = x$ for each c in the prior credal state C_0. Along these lines, some may also wish to add other deference principles.

Now, one way of coming to know the objective chance of some event seems to be via inference from observed physical symmetries.[16] If that's right, it would appear to give us a further type of constraint on credence functions in the prior credal state. More specifically, if some proposition Symm about physical symmetries entails that $\text{ch}(P) = x$, then all credence functions c in the prior credal state should be such that $c(\text{ch}(P) = x \mid \text{Symm}) = 1$. Given that we've accepted the principal principle, this means that we also get that $c(P \mid \text{Symm}) = x$. Now, what sort of things do we have

---

[15] This objectivism may strike you as implausible or undesirable. In the next section, we will consider whether an imprecise Bayesian can give it up without also giving up their evidentialist commitment.

[16] I'm grateful to Pablo Zendejas Medina and an anonymous referee for emphasizing this.

to include in Symm in order for the inference to be correct? In the case of a coin flip, we presumably have to include things like the coin's having homogenous density together with facts about the manner in which it is flipped.[17] But given that we are trying to give *a priori* constraints on credence functions, it seems that this cannot be sufficient. We must also know that, say, the size of the coin or the time of the day are irrelevant to the chance of heads, and similarly for a wide range of other factors. Far-fetched as these possibilities may be, it nevertheless seems that we cannot rule them out *a priori*.

I will return to a discussion of the role of physical symmetries shortly. For the moment, it suffices to note that symmetry considerations, just like the principal principle and other deference principles, can only constrain conditional prior credence assignments, leaving the whole range of unconditional prior credence assignments open. Are there any legitimate constraints on unconditional prior credence assignments? Some endorse the regularity principle, which requires credence functions to assign credence 0 only to propositions that are in some sense (usually doxastically) impossible. So perhaps we should demand that all credence functions in the prior credal state be regular.[18]

So far, I've surveyed a few familiar constraints on credence functions. The thought is that if we add enough of these, we may be able to avoid many instances of belief inertia. However, this strategy faces a dilemma: on the one hand, adding more constraints means that we are more likely to successfully solve the problem. On the other, the more constraints we add, the more it looks like we're going beyond our evidence, in much the same way that the principle of indifference would have us do. Given that Joyce endorsed imprecise Bayesianism for the very reason that it allowed us to avoid having to go beyond the evidence in this manner, this would be especially problematic. Let us therefore assume that the only constraints we can impose on the credence functions in our prior credal state are the principal principle and other deference principles, constraints given by symmetry considerations, and possibly also the regularity principle. This gives us the following result. The evidence grounding thesis, together with an objective understanding of compatibility, imply:

> Maximally Imprecise Priors
> For any contingent proposition P, a rational agent's prior credence $C_0(P)$ in that proposition will be maximally imprecise.[19]

---

[17] See Strevens (1998) for one account of how this works in more detail.

[18] For reasons given by Easwaran ([2014]), Hájek ([unpublished]), and others, I'm skeptical of regularity as a normative requirement on credence functions, but for present purposes I'm happy to grant it.

[19] Where 'maximally imprecise' means either $C_0(P) = (0, 1)$ or $C_0(P) = [0, 1]$, depending on whether or not we accept the regularity principle.

Why does this follow? Take an arbitrary contingent proposition P. If we accept the regularity principle, the extremal credence assignments 0 and 1 are of course ruled out. The principal principle and other deference principles only constrain conditional credence assignments. For example, the principal principle requires each c in the prior credal state C_0 to satisfy $c(P \mid \mathrm{ch}(P) = x) = x$, where $\mathrm{ch}(P) = x$ is the proposition that the objective chance of P is x. Other deference principles have the same form, with $\mathrm{ch}(\cdot)$ replaced by some other probability function one should defer to. By the law of total probability for continuous variables, we have that

$$c(P) = \int_0^1 c(P \mid \mathrm{ch}(P) = x) \cdot f_c(x) \, dx,$$

where $f_c(x)$ is the pdf over possible chance hypotheses that is associated with c. By the principal principle, it follows for all values of x that $c(P \mid \mathrm{ch}(P) = x$, which in turn means that

$$c(P) = \int_{-\infty}^{\infty} x f_c(x) \, dx.$$

This means that the value of $c(P)$ is effectively determined by the pdf $f_c(x)$. Therefore, if we are to use the principal principle to rule out some assignments of unconditional credence in P, we have to do so by ruling out, *a priori*, some pdfs over chance hypotheses. Given the constraints we have accepted on the prior credal state, the only way of doing this[20] would be via symmetry considerations. However, in order to do so we would first have to rule out certain credence assignments over the various possible symmetry propositions. As we have no means of doing so, it follows that neither the principal principle nor symmetry considerations allow us to rule out any values for $c(P)$. Any other deference principles will have the same formal structure as the principal principle, and the corresponding conclusions therefore hold for them as well. We thus get maximally imprecise priors.

Next, we will examine how an agent with maximally imprecise priors might reduce their imprecision. Before doing that, however, I'd like to address a worry you might have about the inference to Maximally Imprecise Priors above. I have been speaking of prior credal states as if they were just like posterior credal states, the only difference being that they're not based on any evidence. But of course, the notion of a prior credal state is a fiction: there is no point in time at which an actual agent adopts it as her state of belief. And given that my formulation of the evidence grounding thesis makes it clear that it is meant to govern credal states at particular points in time, we have no reason to think that it also applies to prior credal states.

If the prior credal state is a fiction, what kind of a fiction is it? Titelbaum ([un-

---

[20] Other than the uninteresting case of the regularity principle ruling out discontinuous pdfs that concentrate everything on the endpoints 0 and 1.

published], p. 110) suggests that we think of priors as encoding an agent's ultimate evidential standards.[21] Her ultimate evidential standards determine how she interprets the information she receives. In the precise case, an agent whose credence function at t_1 is c_1 will regard a piece of evidence E_i as favouring a proposition P if and only if $c_1(P|E_i) > c_1(P)$. So her credence function c_1 gives us her evidential standards at t_1. Of course, her evidential standards in this sense will change over time as she obtains more information. It may be that in between t_1 and t_2 she receives a piece of evidence E_2 such that $c_2(P|E_i) < c_2(P)$. If she does, at t_2 she will no longer regard E_i as favouring P. In order to say something about how she is disposed to evaluate total bodies of evidence, we must turn to her prior credence function, which encodes her ultimate evidential standards. If an agent with prior credence function c_0 has total evidence E, she will again regard that evidence as favouring P if and only if $c_0(P|E) > c_0(P)$. In the same way, we can think of a prior credal state as encoding the ultimate evidential standards of an imprecise agent.[22]

Suppose that we have a sequence of credence functions $c_1, c_2, c_3, \ldots$, where each element c_i is generated by conditionalizing the preceding element c_{i-1} on all of the evidence obtained between t_{i-1} and t_i. We will then be able to find a prior credence function c_0 such that, for each c_i in the sequence, $c_i(\cdot) = c_0(\cdot|E_i)$, where E_i is the agent's total evidence at t_i. Because a credal state is just a set of credence functions, we will also be able to find a prior credal state C_0 such that the preceding claim holds of each of its elements.[23] This means that, in order to arrive at Joyce's judgements about particular cases, we must make assumptions about the prior credal state as well. Consider for instance the third urn example, where we don't even know what colours the marbles might have. If we are to be able to say that it is irrational to have a precise credence in B_3 (the proposition that a marble drawn at random from this urn will be black), we must also say that it is irrational to have a prior credal state C_0 such that there is an x such that $c(B_3|E) = x$ for each $c \in C_0$, where E is the (limited) evidence available to us (namely that the urn contains one hundred marbles of unknown colours, and that one will be drawn at random). Similarly, in

---

[21] This kind of view of priors is of course not original to Titelbaum. See for example Lewis ([1980], p. 288).

[22] In this case, we will have to say a bit more about what it means for an agent to regard a piece of evidence as favouring a proposition. Presumably a supervaluationist account, along the lines of the one we sketched for unconditional comparative judgements, will do: an agent with credal state C will regard a piece of evidence E_i as determinately favouring P if and only if $c(P|E_i) > c(P)$ for each $c \in C$.

[23] Now, c_i and E_i will not determine a unique c_0. There will be distinct c_0 and c_0' such that $c_i(\cdot) = c_0(\cdot \mid E_i)$ and $c_i(\cdot) = c_0'(\cdot \mid E_i)$. In the case of an imprecise Bayesian agent, this means that we cannot infer her prior credal state from her current credal state together with her current total body of evidence. However, given that we are for the moment assuming that the notion of compatibility is an objective one, the prior credal state C_0 should consist of all and only those credence functions that satisfy the relevant set of constraints, and hence that C_0 will be unique.

the unknown bias case, we must rule out as irrational any prior credal state which does not yield the verdict of maximal imprecision.

So although the prior credal state is in a certain sense fictitious, the evidence grounding thesis must still apply to it, if it is to apply to posterior credal states at all. Because of the intimate connection (via imprecise conditionalization on the total evidence) between the prior credal state and posterior credal states, any claims about the latter will imply claims about the former. Therefore, if the evidence grounding thesis is to constrain an agent's posterior credal states, it must also constrain her ultimate evidential standards, namely her prior credal state. Thus the argument for Maximally Imprecise Priors still stands.

In order to determine how widespread belief inertia is, we must now consider how an agent with maximally imprecise priors might reduce her imprecision with respect to some particular proposition. One obvious way for her to do so is through learning the truth of that proposition. If she learns that P, then all credence functions in her posterior credal state will agree that $c(P) = 1$. Given that we required all credence functions in the prior credal state to satisfy the principal principle, another way for the agent to reduce her imprecision with respect to P is to learn something about the chance of P. If she learns that $\text{ch}(P) = x$, then all credence functions in her posterior credal state will agree that $c(P) = x$. Similarly, if she learns that the chance of P lies within some interval $[a, b]$, then all of them will assign a value to P that lies somewhere in that interval.[24] And if we take other deference principles on board as well, those will yield analogous cases.

Although knowledge of objective chance is a staple of probability toy examples, how often do we come by such knowledge in real life? The question is all the more pressing for the imprecise Bayesian. As the unknown bias case illustrated, if an imprecise Bayesian starts out with no information about the objective chance of some class of events, she cannot use observed outcomes of events in this class to narrow down her credence. By contrast, precise Bayesians can use such information to obtain a posterior credence that will eventually be within an epsilon of the objective chance value.

As discussed earlier, we do have one other way of obtaining information about objective chance, namely via inference from physical symmetries. Now, the question is: how often are we in a position to conditionalize on propositions about such symmetries? First, and most obviously, the principle will only be able to constrain credences in propositions for which the relevant physical symmetries are present.

---

[24] I have not explained how the update works when an agent learns that the chance of P lies within some interval $[a, b]$. One way of doing this is to set each pdf f_c to equal zero everywhere outside of that interval and then normalize it, so that $\int_a^b f_c(x)\,dx = 1$. Although I don't believe much of my argument turns on it, there are other ways of doing this as well. I'm grateful to an anonymous referee for drawing my attention to this.

Thus even if we are happy to say that the proposition that my friend Jakob will have *phaal* curry for dinner tonight, or the proposition that the next raven to be observed will be black have non-trivial objective chances, there are presumably no physical symmetries to rely on here. Hence the principle has limited applicability.

Second, in cases where the relevant physical symmetries do exist, we must also know that other factors are irrelevant to the objective chance, as mentioned earlier. From our everyday interactions with the world, as well as from physical theory, we know that the size of a coin and the time of the day are irrelevant to the chance of heads. But how might our imprecise Bayesian accommodate this datum? We know from before that she will have a maximally imprecise prior in any contingent proposition, and hence in any physical theory. So in order to make use of these physical symmetries, she must first narrow down the range of these credences, and assign higher credence to theories according to which the irrelevant factors are indeed irrelevant. But this brings us back to the same problem: how can the imprecise Bayesian reduce her imprecision with respect to these physical theories? Even if we think it's intelligible to think of physical theories as having objective chance of being true, it seems clear that we'll never be in a position to conditionalize on propositions about their objective chance. Furthermore, given that physical theories make claims that go beyond one's evidence, we cannot directly conditionalize a physical theory itself. Thus it would appear that, in practice, the imprecise Bayesian cannot use symmetry considerations to reduce her imprecision. I take it as a given that we do have some way of rationally narrowing down the range of possible objective chance values. We may not know their exact values, but we can nevertheless do a lot better than forever remaining maximally imprecise. The challenge for the evidentially-motivated imprecise Bayesian is to explain how this is possible within their framework.

As you will recall, I suggested that we might want to take on board deference principles other than the principal principle. So a further way of reducing one's imprecision with respect to some proposition would be to defer to a relevant expert. To do so, we must say a bit more about who counts as an expert. The first thing to note here is that if someone has arrived at a relatively precise credence in P through reasoning that is not justified by the lights of evidentially-motivated imprecise Bayesianism, she cannot plausibly count as an expert with respect to P. If the precision of her credence goes beyond her evidence in an unwarranted way, the same must hold of anyone who defers to her credence as well. This greatly limits the applicability of the deference principle. Therefore, we can only legitimately defer to experts in cases where those experts have conditionalized on P directly.[25] However, in order to do so we must not only know what the expert's credence in P is, but also that she is indeed an expert. And again, we don't seem to have a way of narrowing

---

[25] As well as in cases where the expert herself bases her credence on that of another expert, along a sequence of deferrals that must eventually end with someone who conditionalized on P directly.

down our initial, maximally imprecise credence that this person is an expert with respect to *P*.

Given that the constraints we accepted on prior conditional credence assignments have such limited practical applicability, we get the following result:

> Global Belief Inertia
> For any proposition *P*, a rational agent will have a maximally imprecise credence in *P* unless her evidence logically entails either *P* or its negation.

Even if we were willing to concede some instances of local belief inertia, such as in the unknown bias case, this conclusion should strike us as unacceptable. It invalidates a wide range of canonically rational comparative confidence judgements. Propositions that are known to be true are assigned a credence of 1, those that are known to be false are assigned a credence of 0, and all others are assigned a maximally imprecise credence. Although some comparative confidence judgements will remain intact—for instance, all credence functions will regard four heads in a row as more likely than five heads in a row—many others will not.[26] Surely a theory of inductive inference should do better.

Where does this leave us?

6   Responding to Global Belief Inertia

In a sense, Global Belief Inertia is hardly a surprising result in light of my strong assumptions. I assumed the evidence grounding thesis, which states that the credal state must contain all and only those credence functions that are compatible with the evidence. Moreover, I assumed that compatibility is an objective notion, so that there is always an agent-independent fact of the matter as to whether a particular credence function is compatible with a given body of evidence. Finally, I noted that compatibility must be very permissive (in the sense of typically counting a wide range of credence functions as compatible with any particular body of evidence), because otherwise we risk making the same mistake as the one we accused the principle of indifference of making. With all of these assumptions on board, it's almost a given that Global Belief Inertia follows. The question is whether we can motivate imprecise Bayesianism on the grounds that precise credences are often epistemically reckless because they force us to go beyond our evidence, without having the resulting view fall prey to Global Belief Inertia.

Some technical fixes may solve the problem. We saw that Joyce's suggestion for how to avoid belief inertia in the unknown bias case didn't do the job, but perhaps an

---

[26]   See (Rinard [2013]) for further discussion of the implications of maximal imprecision for comparative confidence judgements.

approach along similar lines could be made to work.[27] However, as Joyce concedes, such a proposal could not be justified in light of his evidentialist commitments. Similarly, we might try replacing imprecise conditionalization with some other update rule that allows us to move from maximal imprecision to some more precise credal state. One natural idea is to introduce a threshold, so that credence functions which assigned a value below that threshold to a proposition that we then go on to learn, get discarded from the posterior credal state: $C_1 = \{c(\cdot \mid E_1) : c \in C_0 \wedge c(E_1) > t\}$.[28] The threshold proposal comes with problems of its own: it violates the commutativity of evidence (the order in which we learn two pieces of evidence can make a difference for which credal state we end up with), and it may lead to cases where the credal state becomes the empty set. But again, the more fundamental problem is that it violates the evidentialist commitment. By discarding credence functions that don't meet the threshold, we go beyond the evidence.

In general, the dilemma for evidentially-motivated imprecise Bayesianism is that in order to avoid widespread belief inertia, we must either place stronger constraints on the uniquely rational prior credal state, or concede that there is a range of different permissible prior credal states. However, these two strategies expose the view to the same criticism that we made of objective and subjective precise Bayesianism: they allow agents to go beyond their evidence.

You might worry that the argument for Global Belief Inertia relied on a tacit assumption that the only way of spelling out the underlying evidentialism is via some connection to objective chance (as done, for example, by the chance grounding theses). Once we see that this leads to Global Belief Inertia, we should give up that view, but that doesn't mean we have to give up the evidentialism itself. Indeed, even in the absence of a detailed account of how evidence constrains credal states, it seems quite obvious that our current evidence does not support a precise credence in, say, the proposition that there will be four millimeters of precipitation in Paris on 3 April 2237. So the case for evidentially-motivated imprecision still stands.[29]

The claim is not merely that there is no unique precise credence that is best supported by the evidence. If it were, precise Bayesians could simply respond by saying that there are multiple precise credences, each of which one could rationally adopt in light of the evidence. Instead, the claim must be that, on its own, any precise credence would be an unjustified response to the evidence. Hence the evidence only supports imprecise credences. But does it support a unique imprecise credence, or are there multiple permissible imprecise credences? On the face of it, the claim that

[27] I mentioned one such idea in the context of the unknown bias case: let all the credence functions be based on beta distributions whose parameters are restricted in a particular way.

[28] This threshold rule is mentioned by Bradley and Steele ([2014]). A related method is the maximum likelihood rule given by Gilboa and Schmeidler ([1993]).

[29] I'm grateful to an anonymous referee for articulating this line of thought in a very helpful way.

it supports a unique imprecise credence looks quite implausible. At any rate, it is a claim that stands in need of further motivation. The revised chance grounding thesis gave us one possible explanation of this uniqueness. By including credence functions in the credal state on the basis of their consistency with what we know about objective chance, our criterion gives a clear-cut answer in every case, and hence uniqueness follows. But now that we've rejected the revised chance grounding thesis because of the widespread belief inertia it gave rise to, we no longer have any reason to suppose that the evidence will always support a unique credal state. In the absence of a more detailed account of evidential support for credal states, we should reject uniqueness.

Suppose therefore that we instead accept that our evidence supports multiple imprecise credences. On what grounds can we then say that it doesn't also support some precise credences? The intuition behind the thought that no precise credence is supported by the evidence also suggests that, for sufficiently small values of ϵ, no imprecise credence of $[x - \epsilon, x + \epsilon]$ is supported by the evidence, so the relevant distinction cannot merely be between precise and imprecise credences. What the intuition suggests is instead presumably that no credence that is too precise is supported by the evidence, whether this be perfect precision or only something close to it. But again, to say what qualifies as too precise, we need a more detailed account of evidential support for credal states.

At this point, my interlocutor might simply reiterate their original point, cast in a slightly new form. Yes, they will say, we don't know exactly which credences are too precise for our evidence. But even though we don't have a detailed account, it is still quite clear that some credences are too precise whereas others aren't. So the case for evidentially-motivated imprecision still stands. To give this idea a bit more flesh, consider an analogy with precise Bayesianism.[30] Unless they are thoroughly subjectivist, precise Bayesians hold that some prior credence functions are rational and others aren't. For example, stubborn priors that are moved an arbitrarily small amount even by large bodies of evidence may well be irrational. This cannot be explained by any evidence about objective chance, or indeed by any other kind of evidence, because by definition priors aren't based on any evidence. There are just facts about which of them are rational and which aren't. Furthermore, a credence function is supported by a body of evidence just in case it is the result of conditionalizing a rational prior on that body of evidence.[31] Now, imprecise Bayesians can say the same of their view. Some imprecise prior credal states are rational and others aren't. Again, this cannot be based on any evidence about objective chance, because prior credal states aren't based on any evidence. There are just facts about which of

---

[30] Again helpfully suggested to me by an anonymous referee.
[31] See (Williamson [2000], chapter 10) for an example of a view of this kind, cast in terms of evidential probability.

them are rational and which aren't. Furthermore, a credal state is supported by a body of evidence just in case it is the result of conditionalizing a rational prior credal state on that body of evidence.

I won't attempt to resolve this large dispute here, so let me just say two things in response. The first is simply that those who follow Joyce's line of argument is unlikely to be happy with this kind of position, given that it appears to be vulnerable to the same criticisms as those he raised for precise objective Bayesianism. Of course, imprecise Bayesians who don't share these commitments may well want to respond along these lines, which brings me to my second point: even if they can't give us an exact characterization of which imprecise priors are permissible, they should at least be able to show that none of the permissible priors give rise to widespread belief inertia. Before that has been done, it seems premature to think that the problem has been solved.

Before concluding, let me briefly explore some other tentative suggestions for where to go from here. If we wish to keep the formal framework as it is (namely imprecise probabilism and imprecise conditionalization, together with the supervaluationist understanding of credal states), then one option is to scale back our ambitions. Instead of saying that imprecise credences are rationally required in, say, the second and third urn cases, we only say that they are among the permissible options. This response constitutes a significant step in the direction of subjectivism. We can still place some constraints on the credence functions in the prior credal state (for example that they satisfy the principal principle). But instead of requiring that the prior credal state includes all and only those credence functions that satisfy the relevant constraints, we merely require that it includes only (but not necessarily all) credence functions that satisfy them. On this view, precise Bayesianism goes wrong not in that it forces us to go beyond our evidence (any view that avoids belief inertia will have to!), but rather because it forces us to go far beyond our evidence, when other more modest leaps are also available. How firm conclusions we want to draw from limited evidence is in part a matter of epistemic taste: some people will prefer to go out on a limb and assign relatively precise credences, whereas others are more cautious, and prefer to remain more non-committal. Both of these preferences are permissible, and we should therefore give agents some freedom in choosing their level of precision.

Another option is to enrich the formal framework in a way that provides us with novel resources for dealing with belief inertia. For example, we might associate a weight with each credence function in the credal state and let the weight represent the credence functions degree of support in the evidence.[32] By letting the weights change in response to incoming evidence, inductive learning becomes possible again,

---

[32] See (Gärdenfors and Sahlin [1982]) for an approach along these lines.

even in cases where the spread of values assigned to a proposition by elements of the credal state remains unchanged. In a similar vein, Bradley ([unpublished]) suggests that we introduce a confidence relation over the set of an agent's probability judgements.[33] For example, after having observed 500 heads and 500 tails in the unknown bias case, we may be more confident in the judgement that the probability of heads is in $[0.48, 0.52]$ than we are in the judgement that it is in $[0.6, 1]$. Needless to say, the details of these proposals have to be worked out in much greater detail before we can assess them. Nevertheless, they look like promising options for imprecise Bayesians to explore in the future.

7 Conclusion

I have argued that evidentially motivated imprecise Bayesianism entails that, for any proposition, one's credence in that proposition must be maximally imprecise, unless one's evidence logically entails either that proposition or its negation. This means that the problem of belief inertia is not confined to a particular class of cases, but is instead completely general. I claimed that even if one is willing to accept certain instances of belief inertia, one should nevertheless reject any view which has this implication. After briefly looking at some responses, I tentatively suggested that the most promising options are either (i) to give up objectivism and concede that the choice of a prior credal state is largely subjective, or (ii) to enrich the formal framework with more structure.

Acknowledgements

Department of Philosophy, Logic and Scientific Method
London School of Economics
London, WC2A 2AE
United Kingdom
vallinder@gmail.com

---

[33] This approach is inspired by (Hill [2013]).

Bibliography

Bertrand, J. [1889]: *Calcul des probabilités*. Paris: Gauthier-Villars.

Bradley, R. [unpublished]: *Decision Theory with a Human Face*.

Bradley, S. [2015]: 'Imprecise Probabilities', in E. N. Zalta (*ed.*) *The Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/imprecise-probabilities/>.

Bradley, S. and Steele, K. [2014]: 'Uncertainty, Learning, and the "Problem" of Dilation', *Erkenntnis* **79**(6), pp. 1287–1303.

Easwaran, K. [2014]: 'Regularity and Hyperreal Credences', *Philosophical Review* **123**(1), pp. 1–41.

Gärdenfors, P. and Sahlin, N-E. [1982]: 'Unreliable Probabilities, Risk Taking, and Decision Making', *Synthese* **53**(3), pp. 361–386.

Gilboa, I. and Schmeidler, D. [1993]: 'Updating Ambiguous Beliefs', *Journal of Economic Theory* **59**, pp. 33–49.

Hájek, A. [2003]: 'What Conditional Probability Could Not Be', *Synthese* **137**(3), pp. 273–323.

———— [unpublished]: 'Staying Regular?'

Hill, B. [2013]: 'Confidence and decision', *Games and Economic Behavior* **82**, pp. 675–692.

Jaynes, E.T. [1973]: 'The Well Posed Problem', *Foundations of Physics* **4**(3), pp. 477–492.

Joyce, J. M. [2005]: 'How Probabilities Reflect Evidence', *Philosophical Perspectives* **19**, pp. 153–178.

———— (2010). 'A Defence of Imprecise Credences in Inference and Decision Making', *Philosophical Perspectives* **24**, pp. 281–323.

Levi, I. [1980]: *The Enterprise of Knowledge*, Cambridge, Mass.: MIT Press.

Lewis, D. [1980]: 'A Subjectivist's Guide to Objective Chance', in Jeffrey, R. C. (*ed.*), *Studies in Inductive Logic and Probability*, Volume II, Berkeley: University of California Press, pp. 263–293.

Meacham, C. J. G. and Weisberg, J. [2011]: 'Representation Theorems and the Foundations of Decision Theory', *Australasian Journal of Philosophy* **89**(4), pp. 641–663.

Rinard, S. [2013]: 'Against Radical Credal Imprecision', *Thought: A Journal of Philosophy* **2**(1), pp. 157–165.

Schaffer, J. [2007]: 'Deterministic chance?', *British Journal for the Philosophy of Science* **58**(2), pp. 113–140.

Strevens, M. [1998]: 'Inferring Probabilities from Symmetries', *Noûs* **32**(2), pp. 231–246.

Titelbaum, M. G. [unpublished]: *Fundamentals of Bayesian Epistemology*.

van Fraassen, B. C. [1989]: *Laws and Symmetry*, Oxford: Clarendon Press.

————— [1990]: 'Figures in a Probability Landscape', in Dunn, J. M. and Gupta, A. (*eds*.), *Truth or Consequences: Essays in Honor of Nuel Belnap*, Dordrecth: Kluwer.

Walley, P. [1991]: *Statistical Reasoning with Imprecise Probabilities*. Monographs on Statistics and Applied Probability, Vol 42. London: Chapman and Hall.

White, R. [2010]: 'Evidential Symmetry and Mushy Credence', *Oxford Studies in Epistemology* **3**, pp. 161–186.

Williamson, T. [2000]: *Knowledge and its Limits*. Oxford: Oxford University Press.