Time-slice Bayesianism as a potential solution to the problem of dilation and reflection for imprecise probabilities

Marc Fischer[1]

[1]Laboratoire Georges Friedel (CNRS UMR 5307), Centre SPIN, Mines Saint-Etienne, France

Abstract

One of the main objections against an imprecise probabilistic framework is the apparent absurdity of dilation when seemingly irrelevant evidence makes your belief in a proposition much less certain than it intuitively ought to be. In this work, after critically analysing an argument by White and refined by Topey, as well as responses by imprecise probabilists, I argue that one way to greatly alleviate the tension this type of case poses is to adopt a form of 'time-slice' Bayesianism. In the form I envision it, it means that our degrees of belief in A at time t_i are no longer *ontologically* defined as the result of updating our degrees of belief at time t_{i-1} with the evidence $E_{i-1,i}$ we obtained in between, but as a function of our total evidence available at time t_i and a fundamental prior set of credences. I explain why this move, which forces us to regard all probabilities as conditional probabilities *outside time*, greatly diminishes the *intuitive* appeal of dilation-based counterexamples to the soundness of imprecise Bayesianism.

Keywords. imprecise probability, dilation, synchronic, diachronic, reflection, indifference

1. Introduction

Suppose I show you two coins: A) a fair coin and B) one with an unknown bias. I further tell you that you must choose one of the two coins and that I will give you £1,000 if it lands on heads fewer than $k = 30$ times after being tossed $n = 40$ times. Which coin would you choose? According to orthodox objective Bayesianism, you should apply the Principle of Indifference (POI) to coin B and conclude that $p_A(\text{heads}) = p_B(\text{heads}) = 0.5$ [34], leaving you with no preference. However, similar psychological experiments, such as those that led to the famous Ellsberg paradox, show that most people well-versed in probabilistic reasoning would prefer coin A over coin B [29, 36]. For the fair coin, the probability of obtaining at most $k - 1 = 29$ heads is given by

$$p_A(n_{\text{heads}} < k) = \sum_{i=0}^{k-1} \binom{n}{i} p^i (1-p)^{n-i} \approx 99.88\%.$$

However, if we model $p_B(\text{heads})$ for coin B using the symmetrical interval $[0.1, 0.9]$, in the worst case, we find that

$$p_B(n_{\text{heads}} < k) \approx 0.57\%.$$

If we seek to maximise our worst possible expected outcome, it makes sense to prefer coin A over coin B. This line of reasoning has convinced many philosophers and practitioners to adopt a favourable attitude towards such an imprecise form of Bayesianism known as imprecise probabilism (IP) [32, 33]. However, not everyone has been persuaded [2, 6, 7]. A powerful objection to imprecise probabilities (IP) is the phenomenon of dilation: you start with a belief in A in a narrow probability interval, such as [21.2%, 23.1%], consider a piece of evidence about which you feel very uncertain (and that should not matter), and end up with a very large probability interval, such as [1.3%, 98.6%]. Dilation seems particularly vicious in situations where it happens regardless of the outcome of an event you plan on using as evidence. In Section 2 and Section 3, I will critically analyse a very influential and widely quoted dilation-based argument against IP developed by White and refined by Topey, along with the reactions from proponents of imprecise probabilities. Thereafter, in Section 4, I will present time-slice Bayesianism, discuss how it can help further undermine the strength of White's and Topey's objections as well as dilation in general, and argue why this view is worthy of consideration by the IP research community.

2. White's main argument and reactions in the IP community

While the phenomenon of dilation has been well known to imprecise probabilists for quite some time [12, 28], it has come to be seen by opponents of this framework as a valid ground for rejecting the whole approach [31, 34].

To begin with, I shall define imprecise Bayesianism in a way very similar to Joyce [13]. A *credal set* C is a set of probability distributions $p(\cdot)$ over (ω, F), that is the measurable space with a σ-algebra F that represents all possible events and propositions. In imprecise Bayesianism, an agent uses a *credal set* rather than a single probability distribution. The update of one's beliefs in event $A \in F$ upon seeing evidence $B \in F$ is done by applying Bayes' theorem to each element of the credal set:

$$p(A \mid B) = \frac{p(B \mid A) \cdot p(A)}{\int p(B \mid A) \cdot p(A)\, dA}, \quad \forall p \in C.$$

The result is also a credal set, where the probabilities over A form a new set adjusted by the evidence B. In what follows, I will use the convention that $C(A) = [a, b]$ with $0 \leq a \leq b \leq 1$ if and only if $a = inf_{p(\cdot)\in C}\big(p(A)\big)$ and $b = sup_{p(\cdot)\in C}\big(p(A)\big)$ [1]. The *Committee Metaphor* defended by Joyce [13] (and that was first developped by other authors such as [14, 17, 27]) illustrates imprecise Bayesianism by depicting the agent as a committee, where each member has their own probability distribution over an event. New evidence leads each member to update their probability, and the result is a new *credal set*, reflecting the uncertainty and differing perspectives of the members.

For the present discussion, I have chosen to present White's argument as follows (noting White's example of dilation can already be found in [12, p. 252]). Suppose Hubert has two coins: a perfectly fair coin that can produce either "heads" or "tails" and a mystery coin of unknown bias x, which can produce either "grey" or "black". The two coins are tossed simultaneously and independently. After each trial, Hubert only tells Frédéric which of the biconditionals {Heads ⇔ Grey} or {Tails ⇔ Grey} is true.

White's main argument [2] can then be summarised as follows:

1. If Frédéric follows IP, he ought to represent his ignorance about the mystery coin through a large probability interval, such as $C(\text{grey}) = [0.01\%, 99.99\%]$ and $C_{Heads} = 0.5$.
2. Given his ignorance about the mystery coin, learning that {Heads ⇔ Grey} does not favour heads over tails or vice versa.
3. $C(\text{Heads}) = \{0.5\}$ should thus remain the same after being updated by the evidence.
4. However, according to IP, Frédéric must now believe *a posteriori* that $C^+(\text{heads}) = C^+(\text{grey}) = [0.01\%, 99.99\%]$, where the $+$ sign represents posterior probabilities.
5. It is irrational to believe that $C^+(\text{heads}) = C^+(\text{grey}) = [0.01\%, 99.99\%]$, as the proposition {Heads ⇔ Grey} should be evidentially inert, and we know that the first coin is fair.

Premise 1 is the standard imprecise probabilistic view. Premise 2 appears to be uncontroversial, and Premise 4 is a straightforward consequence of applying Bayes' theorem to all probability distributions p in Frédéric's credal set C. The conclusion in Point 5 logically follows from the premises. However, defenders of imprecise probabilities question Premise 3 and its logical link to Premise 2.

Titelbaum and Hart [10] argued that even with precise probabilities, conditioning on a bijection can be deeply counterintuitive, even when the two events or propositions are stochastically independent. For example, if the mystery coin is replaced by a race involving six horses we *know nothing about* and the POI is applied to compute the probability that Speedy will win, we must believe that

$$c(\text{Heads} \mid \{\text{Heads} \Leftrightarrow \text{Speedy}\}) = c(\text{Speedy} \mid \{\text{Heads} \Leftrightarrow \text{Speedy}\}) = \frac{1}{6}, \quad (1)$$

even though we know nothing about Speedy's performance and are certain the coin is fair.

Joyce defended IP by rejecting Premise 3 and argued that it begs the question: if Frédéric knows almost nothing about the relative proportion of grey outcomes, he will also know almost nothing about the (typical) relative proportions of {Grey, Heads} to {Black, Tails} if we consider a very large number of independent tossings of the two coins. Assuming Frédéric ought to believe the outcome ratio equals 1 upon learning that {Heads ⇔ Grey} presupposes that he must have precise degrees of belief and apply the Principle of Indifference to the mystery coin [13]. The literature on imprecise probabilities offers several similar responses, all suggesting that White's argument is ultimately circular [2].

3. Topey's Argument and Reactions in the Literature

Topey significantly improved White's argument by further developing one of his ideas involving van Fraassen's principle of reflection [31].

1. If Frédéric follows IP, he believes at time $t = 0$ that $C(\text{grey}) = [0.01\%, 99.99\%]$ and $C(\text{Heads}) = \{0.5\}$.
2. Frédéric already knows that at time t, he will either learn that {Heads ⇔ Grey} or {Tails ⇔ Grey} is true.
3. If he learns that {Heads ⇔ Grey}, he will end up believing that $C^+(\text{Heads}) = C^+(\text{Grey}) = [0.01\%, 99.99\%]$.

[1] We are assuming here that A is an event or proposition over a discrete partition

[2] White also used several other arguments, including the one involving dilation that was later refined and defended by Topey.

4. If he learns that {Tails ⇔ Grey}, he will end up believing that $C^+(\text{Heads}) = C^+(\text{Tails}) = C^+(\text{Grey}) = [0.01\%, 99.99\%]$.
5. According to the principle of reflection, if Frédéric knows in advance that he will end up believing $C^+(\text{Heads}) = [0.01\%, 99.99\%]$ at time t no matter what happens, he ought already to believe at time $t = 0$ that $C(\text{Heads}) = [0.01\%, 99.99\%]$ instead of $C(\text{Heads}) = \{0.5\}$.
6. This result is clearly absurd and violates Lewis's Principal Principle, which states that our credences ought to align with known physical probabilities.

Premises 1, 2, 3, and 4 are undeniably true. Premise 6 too seems to be unassailable, as denying it would sever the link between our probabilistic beliefs and the real world.

The only possible move for an imprecise probabilist is to question Premise 5, which goes beyond what van Fraassen referred to as the "General Reflection Principle" [8]:

> **General Reflection Principle**. My current opinion about event E must lie in the range spanned by the possible opinions I may come to hold about E at a later time t, as far as my present opinion is concerned.

In his response to White, Joyce acknowledged that while it is true that $C^+(\text{Heads}) = [0.01\%, 99.99\%]$, there are crucial differences between the situations that should prevent us from applying reflection in this way. Following the *committee metaphor* (see Section 2), he argued that if {Heads ⇔ Grey} is true, the committee members with a strong prior belief in Grey would now believe that $p(\text{Heads})$ is quite high, whereas those who initially thought Grey was unlikely would also believe that Heads is unlikely. Conversely, if {Tails ⇔ Grey} turns out to be true, the committee members with a strong prior belief in Grey would now believe that $p(\text{Heads})$ is very low, whereas those who initially thought Grey was unlikely would now think that Heads is likely. In his response, Topey plausibly argued that the committee metaphor is a psychologically unrealistic description of an agent's beliefs. Since the committee members do not actually exist, he contended that this difference should have no bearing on the validity of reflection.

Defenders of imprecise credences have generally argued that Topey's Reflection Principle (TRP) can plausibly be replaced by another reflection principle that no longer leads to the counterintuitive conclusion. For instance, Moss argued that TRP (which she called "identity reflection") ought to be replaced by the principle of "pointwise reflection": if we know that at time t_1 $C^+(Heads)$ will be included in the interval $[a, b]$, then our current credal set $C(Heads)$ ought to be a *subset* of that interval [22]. Molinari [20, 21] further developped Moss's approach in his work about deference principles. However, Topey is not happy with such a move [3] and argued that if we know that $C^+(A)$ will equal $[a; b]$ for a given event A, we ought to believe precisely that right now, even if our posterior credences regarding other events might differ. This is because anything less would render the Reflection Principle inapplicable to uncontroversial cases involving precise probabilities. The strength of Topey's objection will be evaluated in Subsection 4.3.

4. How time-slice Bayesianism undermines the strength of dilation

4.1. Introducing time-slice Bayesianism. The phrase "time-slice Bayesianism" originates from Hedden's paper titled "Time-Slice Rationality" [11] and Cassel's rejoinder "Time-Slice Epistemology for Bayesians" [5]. In these works, they propose replacing the diachronic norms of Bayesian rationality—according to which rational beliefs at time t_i depend on both one's beliefs at time t_{i-1} and the new evidence $\mathrm{E}_{\mathrm{i-1,i}}$ received in between—with synchronic norms. These synchronic norms depend solely on beliefs and total evidence available at time t_i, without referencing the past.

As an illustration, consider the following example expressed diachronically: Cédric begins the day with the belief that the probability of rain is $p_{\mathrm{t_0}}(\text{Rain}) = 50\%$. After receiving the evidence $E_1 = \{\text{low humidity}\}$ at time t_1, he updates his belief to $p_{\mathrm{t_1}}(\text{Rain}) = p_{\mathrm{t_0}}(\text{Rain} \mid E_1) = 10.9\%$. A bit later, at time t_2, the evidence $E_2 = \{\text{heavy cloud formation}\}$ leads to another update: $p_{\mathrm{t_2}}(\text{Rain}) = p_{\mathrm{t_1}}(\text{Rain} \mid E_2) = 52.4\%$.

Time-slice Bayesianism is the view that there is no *irreducible* diachronic norm, meaning that any (valid) diachronic statement about probabilities and degrees of belief can ultimately be reduced to a synchronic statement. Moreover, it holds that we ought to reject any statement that cannot be reduced in this manner. For instance, within his own framework Meacham [18] would define $p(\text{Rain})$ at time t_2 based on the evidence E_2 and Frédéric's degrees of belief cr at t_2 regarding his beliefs at t_1:

$$cr_{t_2}(\text{Rain}) = \sum_i cr_{t_2}\Big(cr_{t_1} = p_i\Big) p_i(\text{Rain} \mid E_2),$$

where $\{p_i\}$ represents the set of possible probability functions Frédéric might have had at t_1. The challenge with this account, as noted by Cassel [5], is that people rarely hold precise degrees of belief about their own prior degrees of belief.

Hedden [11], on the other hand, proposes that the probability at t_2 depends solely on the *total evidence* at that time and a unique rational prior probability function

[3] Personal correspondance.

p_0:

$$cr_{t_2}(\text{Rain}) = p_0(\text{Rain} \mid E_1 \wedge E_2) = \frac{p_0(\text{Rain} \wedge E_1 \wedge E_2)}{p_0(E_1 \wedge E_2)}.$$

Cassel critiques this approach, highlighting that Hedden's definition of "evidence" aligns with epistemological internalism, where evidence supervenes on mental states. This makes it fallible, as it may depend on the reliability of memory and human senses. Consequently, evidence E_1 from t_1 might logically contradict evidence E_2 from t_2, leaving $cr_{t_2}(\text{Rain})$ undefined.

Adopting an externalist perspective on evidence and truth, I believe that evidence corresponds to empirical facts in the real world, rendering Cassel's objection irrelevant. However, Hedden's claim that there is a unique prior probability distribution that every rational agent should hold is far more troubling, as I shall explain in Subsection 4.4. Thus, I propose the following definition:

Definition: Time-slice imprecise Bayesianism

> The credence at time t of a rational agent depends solely on the total evidence E_{total} at that time and a *reasonable* primordial credal set C_0 corresponding to the absence of evidence. For any event A:
>
> $$C_t(A) = \Big\{cr_t(A) = p_0(A \mid E_{\text{total}}) = \frac{p_0(A \wedge E_{\text{total}})}{p_0(E_{\text{total}})},\ p_0 \in C_0\Big\}. \quad (2)$$

It is noteworthy that while time-slice Bayesianism does not consider that time is a fundamental irreducible variable, Bayesian conditionalisation can still be meaningfully applied between t_1 and t_2. For the special case where the credal set contains only a single probability distribution, it naturally follows from the definition of $p_{t_1}(Rain)$ and $p_{t_2}(Rain)$ that

$$p_{t_2}(\text{Rain}) = p_{t_1}(\text{Rain} \mid E_2) \cdot p_{t_1}(E_2),$$

provided we do not consider the extreme situations where a piece of evidence makes a proposition or its negation logically certain. Adopting time-slice Bayesianism is thus fully compatible with the continued use of diachronic updating formulas that have proven incredibly useful in many fields of research.

4.2. Impact on White's argument. If time, as well as the past and future, are no longer irreducible quantities in determining the rationality of belief, White's main argument can be reformulated as follows:

1. If Frédéric follows the imprecise probabilist (IP) approach, he ought to represent his ignorance about the mystery coin using a large probability interval, such as $C(\text{Grey}) = [0.01\%; 99.99\%]$.
2. Given his ignorance about the mystery coin, the knowledge of {Heads ⇔ Grey} should not favour heads over tails or vice versa.
3. Consequently, he should believe that
$$C_0(\text{Heads}|\{\text{Heads} \Leftrightarrow \text{Grey}\}) = \{0.5\}.$$
4. However, according to the IP framework, Frédéric must also believe that
$$C_0(\text{Heads}|\{\text{Heads} \Leftrightarrow \text{Grey}\}) = C_0(\text{Grey}|\{\text{Heads} \Leftrightarrow \text{Grey}\}) = [0.01\%; 99.99\%]. \quad (3)$$
5. It is irrational to believe that knowing {Heads ⇔ Grey} could have any impact on $C(\text{Heads})$.

Let us consider a very large number of independent tossings of the two coins. If $p(\text{Grey})$ happens to be very small, the long-term ratio

$$\frac{N(\text{Grey} \wedge \text{Heads})}{N(\{\text{Grey} \wedge \text{Heads}\} \cup \{\text{Black} \wedge \text{Tails}\})}$$

will also typically be very small as there are then much more Black ∧ Tails combinations than Grey ∧ Heads. Conversely, if $p(\text{Grey})$ happens to be very large, it will typically also be very large. Thus, Premise 5 clearly begs the question against the imprecise probabilist, who rejects the principle of indifference and does not agree that we ought to treat an unknown coin as if it were symmetrical and would typically land on one side half of the time.

While there are already strong responses to what I call "White's main argument," I believe that adopting a timeless perspective further undermines it, as this standpoint makes its flaws even more apparent.

4.3. Impact on Topey's argument. Here is what I believe to be an accurate timeless formulation of Topey's argument:

1. If Frédéric follows the imprecise probabilist (IP) approach, he ought to represent his ignorance about the mystery coin through a large probability interval, such as $C_0(\text{Grey}) = [0.01\%; 99.99\%]$.
2. Let E_1 = {Heads ⇔ Grey} and E_2 = {Tails ⇔ Grey} Frédéric also believes that
$$C_0(\text{Heads}|E_1) = C_0(\text{Grey}|E_1) = [0.01\%; 99.99\%]. \quad (4)$$
and at the same time that
$$C_0(\text{Heads}|E_2) = C_0(\text{Grey}|E_2) = [0.01\%; 99.99\%]. \quad (5)$$
3. Since the events E_1 and E_2 are complementary, Frédéric ought to believe that if $C_0(\text{Heads} \mid E_1) = C_0(\text{Heads} \mid E_2)$, then they must also equal the unconditional $C_0(\text{Heads})$.

4. Frédéric ought, therefore, to believe that $C_0(\text{Heads}) = [0.01\%; 99.99\%]$ instead of $C_0(\text{Heads}) = \{0.5\}$.
5. This is clearly an absurd result that violates Lewis' Principal Principle, according to which our credence ought to match all known physical probabilities.

It is truly remarkable that Topey's argument about the problems with the temporal evolution of the agent's beliefs have now been reduced to problems with the primordial credal set C_0. Premise 3 , which involves the notion of complementarity, seems to be a fair translation of Topey's Reflection Principle into a tenseless language. But what could motivate it? One option might be to ground this intuition in gambling scenarios. Consider that Hubert proposes Frédéric to give him £0.8 if Heads, but he must give me £0.2 if Tails. If this is all he knows, he will surely accept the bet, as the expected reward is

$$E(R) = p(\text{Heads}) \cdot 0.8 - p(\text{Tails}) \cdot 0.2 = 0.3 > 0.$$

If, however, he now knows either that $E_1 = \{\text{Heads} \Leftrightarrow \text{Grey}\}$ or $E_2 = \{\text{Tails} \Leftrightarrow \text{Black}\}$, it is *no longer certain* that he will accept the bet, as the expected reward could be negative and this is intuitively absurd as learning one of these events does not favour Heads over Tails or vice-versa. Nevertheless, apart from the prior commitment to the notion that an unknown coin ought to be treated like a fair coin, it is hard to see what justifies such a psychological intuition. While some people might still be moved by it, its appeal seems to be considerably weaker than for the temporal scenario first envisioned as its question-begging nature is much more transparent.

The best way to salvage Topey's argument appears thus to be to deny the feasibility of time-slice Bayesianism and affirm the universal validity of his version of the Reflection Principle as irreducible to timeless language.

Topey defended it by arguing that if one requires one's credences about other events to also be equal before applying reflection, the principle would quickly become useless, as it could no longer be applied to uncontroversial cases involving precise probabilities where rejecting it would clearly be irrational. Here is the type of scenario he had in mind:

> Fred knows that $Pr(A) = 0.3$, $Pr(B) = 0.7$, $Pr(C) = 0.7$, $Pr(A \mid A \leftrightarrow B) = Pr(B \mid A \leftrightarrow B) = 0.5$, $Pr(A \mid A \leftrightarrow C) = Pr(C \mid A \leftrightarrow C) = 0.5$, $Pr(B \mid A \leftrightarrow C) > 0.7$, and $Pr(C \mid A \leftrightarrow B) > 0.7$. If he learns that $A \leftrightarrow B$ or that $A \leftrightarrow C$, he will believe that $Pr_+(A) = 0.5$. If he is told that he'll either learn $A \leftrightarrow B$ or $A \leftrightarrow C$ two hours later, he must, by Reflection, already believe that $Pr(A) = Pr_+(A) = 0.5$. However, in the first case, $Pr_+(A) = Pr_+(B)$, whereas in the second case, $Pr_+(A) = Pr_+(C)$. Thus, differences in the posterior credences regarding other events are simply irrelevant to the application of the principle: whenever the posterior credences regarding A are equal to one another, it ought to be applied regardless of the other posterior credences regarding B and C.

However, Topey is wrong to assume that the *only way* to salvage that conclusion is by assuming the universal validity of a reflection principle involving credences. Indeed, Briggs [4] demonstrated that the conclusions of the (precise Bayesian) Reflection Principle can be derived from Bayes' theorem, thereby rendering the former dispensable. If he knows that either $\{A \leftrightarrow B\}$ or $\{A \leftrightarrow C\}$ is true, Fred immediately knows via Bayes' theorem that $Pr^+(A) = 0.5$, without having to wait two hours for the precise information. Consequently, there seems to be no good grounds for adopting an irreducible reflection principle that would undermine the use of imprecise probabilities. It is worth noting that we could also defend Moss' identity reflection principle in a similar fashion in a diachronic framework.

4.4. On the uniqueness of the primordial credence function. As I explained in 4.1, if we want to define the credences of an agent at time t timelessly, we need both the total evidence available to him or her and a reasonable primordial credence function C_0, corresponding to the absence of any evidence. What can now be said about the uniqueness of C_0 and the consequences for the agent's beliefs?

To answer this question, it is useful to consider the example mentioned in the introduction, where we have both a fair coin and an unknown coin, and we want to engage in bets related to the event $E = \{N(\text{Heads}) \leq 30\}$. Following the binomial distribution, for the fair coin we have $p_{\text{Fair}}(E) \approx 99.88\%$, whereas for the unknown coin, if we allow for the possibility that it is biased but not extremely so ($p_{\text{Unknown}}(\text{Heads}) \in [0.1; 0.9]$), all we can say is that $p_{\text{Unknown}}(E) \in [0.15\%; 1 - 6.66 \times 10^{-16} \approx 1]$. Let us now consider that I ask you to choose one of these two bets:

- **Bet 1**: If E comes true for the fair coin, you get £1,000,000; if $\overline{E}$, you give me £10.
- **Bet 2**: If E comes true for the unknown coin, you get £1,000,000; if $\overline{E}$, you give me £1.

The expected reward of **Bet 1** is

$$R_1 = 1,000,000 \times p_{\text{Fair}}(E) - 10 \times p_{\text{Fair}}(\overline{E}) \approx 998,799.88\,£.$$

The expected reward of **Bet 2** is

$$R_2 = 1,000,000 \times p_{\text{Unknown}}(E) - 1 \times p_{\text{Unknown}}(\overline{E})\,.$$

Given that $p_{\text{Unknown}}(\text{Heads}) \in [0.1; 0.9]$, it follows that $R_2 \in [1,490.02; 1,000,000]$.

If, however, we adopt **precise impermissive Bayesianism** (see [19] for a thorough analysis), there is only one number that a rational agent can assign to $p_{\text{Unknown}}(\text{Heads})$. Since this number must fulfil both the symmetry and non-arbitrariness constraints, it follows that $p_{\text{Unknown}}(\text{Heads}) = 0.5$ and that

$$R_2 = 998,799.988 > R_1.$$

Therefore, any agent that prefers **Bet 1** over **Bet 2** is *irrational.*

A possible strategy the precise Bayesian might use to avoid this conclusion would be to no longer apply the POI directly to the binary outcome Head/Tail but to the continuous probability $\theta = p(Head) \sim U(0,1)$. Given n independent flips, let X be the number of heads observed. The probability of observing at most $k_{\max}$ heads is given by:

$$P(E \mid \theta) = \sum_{k=0}^{k_{\max}} \binom{n}{k} \theta^k (1-\theta)^{n-k}. \tag{6}$$

Our goal is to compute the expected probability over all possible values of θ:

$$P(E) = \mathbb{E}_\theta[P(E \mid \theta)] = \int_0^1 \sum_{k=0}^{k_{\max}} \binom{n}{k} \theta^k (1-\theta)^{n-k} d\theta. \tag{7}$$

Each term in the sum involves the integral:

$$I_k = \int_0^1 \theta^k (1-\theta)^{n-k} d\theta. \tag{8}$$

This is recognized as the Beta function:

$$I_k = B(k+1, n-k+1) = \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)}. \tag{9}$$

Using $\Gamma(k+1) = k!$, we obtain:

$$I_k = \frac{k!(n-k)!}{(n+1)!}. \tag{10}$$

Thus, summing over k:

$$P(E) = \sum_{k=0}^{k_{\max}} \binom{n}{k} \frac{k!(n-k)!}{(n+1)!}. \tag{11}$$

Using $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, we simplify:

$$P(E) = \sum_{k=0}^{k_{\max}} \frac{n!}{(n+1)!}. \tag{12}$$

Since $\frac{n!}{(n+1)!} = \frac{1}{n+1}$, this results in:

$$P(E) = \frac{k_{\max}+1}{n+1}. \tag{13}$$

The expected payoff of a bet depends on whether the event E occurs. We consider a bet where we receive a reward W if the event E occurs and incur a loss L otherwise. The expected winnings, given an unknown probability θ for heads, can be expressed as the expectation over all possible outcomes of the random experiment. If the probability of E under a given θ is denoted as $P_\theta(E)$, the possible payoffs are:

- **If** E **occurs:** we receive the reward W.
- **If** E **does not occur:** we incur the loss L.

The probabilities for these cases are:

$$P_\theta(E) = P(E \mid \theta), \quad P_\theta(\overline{E}) = 1 - P_\theta(E). \tag{14}$$

Thus, the conditional expected winnings under a fixed θ are given by:

$$\mathbb{E}[\text{Winnings} \mid \theta] = P_\theta(E)W + (1 - P_\theta(E))L. \tag{15}$$

Since θ itself is a random variable (assumed to follow $\theta \sim U(0,1)$), we must compute the expectation over all possible values of θ:

$$\mathbb{E}_\theta[\text{Winnings}] = \int_0^1 \mathbb{E}[\text{Winnings} \mid \theta] f(\theta) d\theta. \tag{16}$$

Given that $\theta \sim U(0,1)$, its probability density function is $f(\theta) = 1$ for $\theta \in [0,1]$, so the expectation simplifies to:

$$\mathbb{E}_\theta[\text{Winnings}] = \int_0^1 \left(P_\theta(E)W + (1 - P_\theta(E))L\right) d\theta. \tag{17}$$

Using the linearity of integration, we separate the terms:

$$\mathbb{E}_\theta[\text{Winnings}] = W \int_0^1 P_\theta(E) d\theta + L \int_0^1 (1 - P_\theta(E)) d\theta. \tag{18}$$

The second integral simplifies as follows:

$$\int_0^1 (1 - P_\theta(E)) d\theta = 1 - \int_0^1 P_\theta(E) d\theta. \tag{19}$$

Thus, we obtain:

$$\mathbb{E}_\theta[\text{Winnings}] = W\mathbb{E}_\theta[P_\theta(E)] + L(1 - \mathbb{E}_\theta[P_\theta(E)]). \tag{20}$$

Since the expectation of $P_\theta(E)$ over all possible values of θ is precisely $P(E)$, we obtain the desired formula:

$$\mathbb{E}_\theta[\text{Winnings}] = P(E)W + (1 - P(E))L. \tag{21}$$

This formula shows that the expected winnings are determined by the weighted sum of the winnings W and losses L, where the weights are given by the probability of the event E. If $P(E)$ is derived from an uncertainty distribution such as $\theta \sim U(0,1)$, it is replaced by the

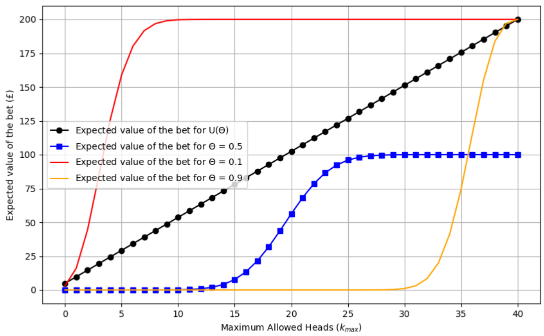

Figure 1. *Comparison of the two bets when modelling our ignorance about θ through a uniform prior*

expected probability over all possible values of θ, leading to a different expected winnings calculation than in the case of a fixed θ.

Let us consider that $k_{max} = 30$, $n = 0$, and the two following bets:

- **Bet 1**: If E comes true for the fair coin, you get £ 100.
- **Bet 2**: If E comes true for the unknown coin, you get £ 200.

Figure 1 compares the expected wins of the two bets if we model our uncertainty through a uniform prior over the values of θ or through the lower and upper probability for $p(Head)$. What we can see is that for $k_{max} = 30$, if we were to use the uniform prior, we would be forced to favour Bet 2 over Bet 1. This is deeply problematic because, while the value of Bet 1 is entirely based on our knowledge of the fair coin, the higher expected return of Bet 2 is purely a result of our complete ignorance. This could be highly misleading if the unknown coin happens to be heavily biased towards Heads ($\theta = 0.9$).

I therefore strongly doubt that precise impermissive Bayesianism could ever become the consensus view. The reason is that, for many people, the intuition that it is wrong—if not downright absurd—to call anyone refusing to choose Bet 2 **irrational** is stronger than (or at least as strong as) the intuitions relied upon to argue that degrees of belief ought to be precise numbers every rational agent ought to agree upon. This is illustrated by the fact that, even though precise impermissive Bayesianism has been around for quite a long time (at least since Laplace's first mention of the principle of insufficient reason in 1812 [16]), it has never been able to convince the majority of well-informed individuals working on the foundations of probability. As the arguments in its favour have become more sophisticated, so too have the counterarguments.

But what about **imprecise Bayesian impermissivism** now? It is tempting to require that the unique primordial imprecise credence function C_0 also satisfy the constraints of symmetry and non-arbitrariness, in addition to making a clear distinction from the case of the known fair coin. This would then lead us straight to $C_0(\text{Heads}) = (0, 1)$[4]. It is, however, well known that such a vacuous prior would lead to the problem of belief inertia [22]: regardless of the amount of total evidence at our disposal, we would always believe that $C(\text{Heads}) = (0, 1)$, even if we were to observe that after 1 million tosses, the relative frequency $f(\text{Heads})$ closely oscillates around 0.50. This is clearly not an acceptable behaviour for any theory of induction worthy of the name. In his defence of time-slice rationality, Hedden [11] believes that C_0 should be unique, but he is at the same time quite open to the possibility of using imprecise instead of precise probabilities. He argues that, instead of using formal criteria such as those leading to the principle of indifference, we ought to resort to "*substantive (i.e., non-formal) constraints on rational credences which single out a uniquely rational prior*" and that, if the desired properties such as simplicity, naturalness, and explanatoriness conflict with one another, they ought to be weighted against one another. The combination of all such weightings would then lead to a uniquely defined mushy credence function C_0. In other words, we would have $C_0(\text{Heads}) = [a, b]$ with $0 < a < b < 1$, where a and b take on precise values, even though we might not be able, at least for now, to identify what they are. The problem I have with that view is that there is no guarantee that unambiguous a and b exist for a single individual based on his or her own intuitions, let alone for a very large group of individuals or even experts. So far, this appears to be nothing more than a *statement of faith*.

Consequently, I prefer to consider the view of **moderate permissible imprecise Bayesianism**. By permissible, I mean denying that there is a unique imprecise primordial credence function C_0, while at the same time denying that "anything goes". In the case of an unknown coin, believing that $C_0(\text{Heads}) = [0.499, 0.501]$ is clearly unacceptable; however, there is no good reason to think that we ought to believe that $C_0(\text{Heads}) = [0.01, 0.99]$ instead of $C_0(\text{Heads}) = [0.001, 0.999]$. While this type of *vagueness* [35] may be seen as problematic, it is an unavoidable fact of life in many scientific disciplines. In the field of the physics of divided solids, for example, a powder can often be operationally defined as an ensemble of grains with moderate interactions between the particles that display a range of phenomena such as low apparent density (the property of 'sponginess'), high compressibility (the property of 'fluffiness'), and high flowability (the property of 'fluidness') [1]. Given that, a bunch of

[4] If we assume that $p(\text{Heads})$ cannot exactly equal 0 or 1.

grains with extremely weak interactions would not be a powder, and neither would an ensemble of grains so tightly connected that they behave like a single solid and do not display any flowability. There will, however, be borderline cases where, if we say that A is a powder, and the average strength of the interactions between the particles of B equals 99.99% of those of A, it would be very problematic (all other things being equal) to say that B is not a powder and to identify a stark boundary between powders and non-powders.

Another feature of the moderate permissible imprecise Bayesianism I am proposing would be to use Moss's notion of global constraint [22]. Unlike local constraints valid for every precise probability distribution $p_0 \in C_0$, global constraints can only be defined with respect to the credal set C_0 as a whole. Such global constraints might, for example, include avoiding the phenomenon of belief inertia or cheap evidence.

A problem that remains to be handled is that, while a unique C_0 leads to a credence function C that is unambiguous given the same total body of evidence E, it is no longer the case when we allow for the existence of a range of reasonable C_0. The worry is that the same agent, who fails to learn anything new between t_1 and t_2, might nevertheless have a credal set C that keeps fluctuating, and such behaviour seems to be intuitively problematic to many people. There are two potential solutions to this problem. The first, proposed by Hedden himself as an alternative to the uniqueness of the prior, would be to add the constraint that *"at each particular time, you ought not have a disposition or policy of abandoning your current credences in favour of others"* (in which case "*partial* time-slice rationality" might be a more appropriate phrase). The second would be to bite the bullet and consider that the range of reasonable primordial credal sets C_0 (given all the constraints) might be narrow enough to guarantee that such fluctuations in one's beliefs would be small and harmless when it comes to losing money.

5. Conclusion

The phenomenon of dilation remains a significant objection to imprecise probability, and the work of defenders of imprecise Bayesianism, such as Moss [22], Pedersen and Wheeler [25], Kelly [15], and Bradley [3], is highly valuable in mitigating the strength of this objection. I have adopted several important ideas from these works. However, the role and importance of *time* and diachronicity in the appeal of these intuitions have not, to the best of my knowledge, ever been addressed systematically in the literature on imprecise probabilities.

I believe that the approach of time-slice Bayesianism, presented and defended here, further diminishes the intuitive appeal of the arguments put forward by White and Topey and fundamentally changes the nature of the question under discussion.

Instead of asking: "*Can imprecise probabilism still be reasonably used if there are situations where it leads to intuitively inadmissible changes in one's beliefs between t_1 and t_2 due to evidence E?*"

we are now asking two different questions, namely:

- *Is the credal set C, obtained by conditioning the primordial credal set C_0 on the total evidence E, truly problematic for some total evidence $E \in F$?*
- *If so, how can we constrain the reasonable range of C_0 in such a way that C exhibits the desired behaviour $\forall E \in F$ in addition to other desirable properties?*

Regarding the first question, and the thought experiment I have been dealing with, my answer is a resounding "No!". As I have been arguing all along, believing both that $C_0(Heads) = 0.5$ and $C_0(Heads|Heads \leftrightarrow Tails) = [0.01\%, 99.99\%]$ (and $C_0(Heads|Heads \leftrightarrow Grey) = [0.01\%, 99.99\%]$) at the same time is entirely unproblematic when it comes to rationality.

I am sympathetic to Norton's material theory of induction (see [23, 24]), which argues that it is a mistake to assume the existence of a *single universal* framework applicable to all real-world problems. Whether the framework is precise or imprecise, or whether it involves permissible or impermissible Bayesianism, the appropriate inductive technique is determined by the material properties of the specific problem at hand. Therefore, I believe the second question may be answered differently depending on the field of application. The framework presented here is general enough to accommodate this variability (see the seminal work of Pedersen and Wheeler for guidance on recognising and avoiding improperly constructed fundamental prior credal sets [25]). The aims of this paper are modest: it seeks only to show that adopting a timeless perspective is a *potentially* useful strategy for addressing objections from those who continue to believe that dilation demonstrates the inadmissibility of imprecise probabilistic models and remain unmoved by the counterarguments that have been presented by defenders of IP so far [5]. I thus contend that time-slice Bayesianism is a promising framework for future research within the IP community.

One promising research avenue is to explore how a robust notion of total evidence at time t - which includes not only empirical evidence but also the process of coming to believe in that evidence based on a method M, as discussed by [9] - can be realistically implemented. There is also a need to think more deeply about the grounds

[5] As Hart and Titelbaum [10] and I [7] have noted, intuitions can play a pivotal role in evaluating new probabilistic frameworks. However, they may prove unreliable, particularly when applied to complex or unconventional thought experiments. Moreover, intuitions often differ from person to person, which may explain why some individuals remain deeply troubled by the problem of dilation, while others believe that satisfactory solutions have already been found.

for taking such a timeless stance or for rejecting one. One positive argument could be the fact that many plausible physical theories do not involve time as an irreducible quantity [26] and that if we want our epistemology to be compatible with them, it should not involve irreducible temporal norms either. An interesting challenge has been very recently put forward by [30] (work in progress), in which she argues that it is not irrational to have different conflicting preferences at time t, but that this can lead to an irrational series of decisions that is "irreducibly diachronic" in that no single step can be singled out as the faulty one. It remains to be seen whether other conclusions can be drawn or whether her main thesis itself is compatible with the view proposed here that there is no irreducibly diachronic norm when it comes to having appropriate *degrees of belief*.

The present article should thus be seen as a conversation starter about views whose consequences need to be further evaluated.

Additional author information

Acknowledgements. I am deeply grateful to Brett Topey for our discussions regarding his critique of the IP framework, as well as to Jason Konek and Kevin Blackwell for our stimulating exchanges about IP and their encouraging thoughts and remarks.

Competing interests. No competing interests have to be declared for this purely theoretical study.

References

[1] Yasuo Arai. "Powder fundamentals". In: *Chemistry of Powder Production*. Dordrecht: Springer Netherlands, 1996, pp. 24–89. DOI: 10.1007/978-94-009-1493-3_3.

[2] Seamus Bradley. "Dilation and Informativeness". In: *Reflections on the Foundations of Probability and Statistics: Essays in Honor of Teddy Seidenfeld*. Springer, 2022, pp. 125–141.

[3] Seamus Bradley and Katie Steele. "Uncertainty, learning, and the "problem" of dilation". In: *Erkenntnis* 79 (2014), pp. 1287–1303.

[4] Rachael Briggs. "Distorted reflection". In: *Philosophical Review* 118.1 (2009), pp. 59–85.

[5] Lisa Cassell. "Time-slice epistemology for Bayesians". In: *Inquiry* (2022), pp. 1–24.

[6] Marc Fischer. "On the principal principle and imprecise subjective Bayesianism: A reply to Christian Wallmann and Jon Williamson". In: *European Journal for Philosophy of Science* 11.2 (2021), p. 47.

[7] Marc Fischer. "On Imprecise Bayesianism in the Face of an Increasingly Larger Outcome Space: A Reply to John E. Wilcox". In: *Journal for General Philosophy of Science* 53.4 (2022), pp. 367–379.

[8] Bas C van Fraassen. "Conditionalization, a new argument for". In: *Topoi* 18.2 (1999), pp. 93–96.

[9] Ruobin Gong, Joseph B Kadane, Mark J Schervish, Teddy Seidenfeld, and Rafael B Stern. "Learning and total evidence with imprecise probabilities". In: *International Journal of Approximate Reasoning* 151 (2022), pp. 21–32.

[10] Casey Hart and Michael G Titelbaum. "Intuitive dilation?" In: *Thought: A Journal of Philosophy* 4.4 (2015), pp. 252–262.

[11] Brian Hedden. "Time-slice rationality". In: *Mind* 124.494 (2015), pp. 449–491.

[12] Timothy Herron, Teddy Seidenfeld, and Larry Wasserman. "The extent of dilation of sets of probabilities and the asymptotics of robust bayesian inference". In: *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*. Vol. 1994. 1. Cambridge University Press. 1994, pp. 250–259.

[13] James M Joyce. "A defense of imprecise credences in inference and decision making". In: *Philosophical perspectives* 24 (2010), pp. 281–323.

[14] Joseph B Kadane. *Bayesian methods and ethics in a clinical trial design*. John Wiley & Sons, 2011.

[15] Kevin T Kelly. *Ockham's Razor, Hume's Problem, Ellsberg's Paradox, Dilation, and Optimal Truth Conduciveness*. 2008.

[16] Pierre Simon Laplace. *Théorie analytique des probabilités*. Courcier, 1820.

[17] Isaac Levi. "Conflict and social agency". In: *The Journal of Philosophy* 79.5 (1982), pp. 231–247.

[18] Christopher JG Meacham. "Unravelling the tangled web: Continuity, internalism, uniqueness and self-locating belief". In: *Oxford studies in epistemology* 3 (2010), pp. 86–125.

[19] Christopher JG Meacham. "Impermissive bayesianism". In: *Erkenntnis* 79 (2014), pp. 1185–1217.

[20] Giacomo Molinari. "Trust the evidence: two deference principles for imprecise probabilities". In: *International Symposium on Imprecise Probability: Theories and Applications*. PMLR. 2023, pp. 356–366.

[21] Giacomo Molinari. "Deference Principles for Imprecise Credences". In: (2024).

[22] Sarah Moss. “Global constraints on imprecise credences: Solving reflection violations, belief inertia, and other puzzles”. In: *Philosophy and Phenomenological Research* 103.3 (2021), pp. 620–638.

[23] John D Norton. *The material theory of induction*. University of Calgary Press, 2021.

[24] William Peden. “Direct inference in the material theory of induction”. In: *Philosophy of Science* 86.4 (2019), pp. 672–695.

[25] Arthur Paul Pedersen and Gregory Wheeler. “Demystifying dilation”. In: *Erkenntnis* 79 (2014), pp. 1305–1342.

[26] Simon Saunders. “How relativity contradicts presentism”. In: *Royal Institute of Philosophy Supplements* 50 (2002), pp. 277–292.

[27] Teddy Seidenfeld, Joseph B Kadane, and Mark J Schervish. “On the shared preferences of two Bayesian decision makers”. In: *The Journal of Philosophy* 86.5 (1989), pp. 225–244.

[28] Teddy Seidenfeld and Larry Wasserman. “Dilation for sets of probabilities”. In: *The Annals of Statistics* 21.3 (1993), pp. 1139–1154.

[29] Michael J Smithson. “Human judgment and imprecise probabilities”. In: *Red* 30 (1997), p. 60.

[30] Johanna Thoma. “Some Mistakes Are Irreducibly Diachronic”. In: (2024).

[31] Brett Topey. “Coin flips, credences and the reflection principle”. In: *Analysis* 72.3 (2012), pp. 478–488.

[32] Peter Walley. “Towards a unified theory of imprecise probability”. In: *International Journal of Approximate Reasoning* 24 (2000), pp. 125–148.

[33] GREGORY Wheeler. “A gentle approach to imprecise probability”. In: *Essays in Honor of Teddy Seidenfeld, Theory and Decision Library A, Dordrecht. Springer* (2021).

[34] Roger White. “Evidential symmetry and mushy credence”. In: *Oxford studies in epistemology* 3 (2010), pp. 161–186.

[35] Timothy Williamson. *Vagueness*. Routledge, 2002.

[36] Carlo Zappia. “Leonard savage, the Ellsberg paradox, and the debate on subjective probabilities: evidence from the archives”. In: *Journal of the History of Economic Thought* 43.2 (2021), pp. 169–192.