Imprecision and the representation of belief

Giacomo Molinari

September 18, 2022

1 Overview

This essay is about the way three people respond to a formal result connecting binary (i.e. all-or-nothing) beliefs and precise credences. This result was outlined by Fitelson and Easwaran in a pair of papers (2015, 2016) as a response to the preface paradox, but has broader implications. Each of our three main characters interprets the result from a different philosophical perspective, and uses it for a different purpose. I will present a similar result connecting beliefs and imprecise credences, and then evaluate how it affects the views of each character. For at least some of their views, I will argue, the imprecise result is a significant improvement.

Here is the essay plan. In Section 2, I introduce some notation and describe the Preface paradox. I then present the formal result which was formulated in response to the paradox by Easwaran and Fitelson (2015). This result establishes a connection between binary and gradational doxastic models, and is centered around a representation of the former by means of the latter. In Section 3, our three main characters are introduced. I will outline how each of them interprets the representation, and for what purpose. Section 4 is a brief summary of the imprecise probabilistic notions that are used in Section 5 to develop a representation binary beliefs by means of imprecise probabilities, extending Fitelson and Easwaran's result. Section 6 discusses the impact of this new representation on our main characters. Section 7 sums up the main conclusions.

2 The precise representation

This section introduces Fitelson and Easwaran's (2015) representation of binary doxastic models by means of precise ones. I follow their exposition in starting from the Preface paradox, which gives one motivation for the representation. In the next section, we will see that this is not the only one. Before we start, however, a bit of terminology is in order.

2.1 Set-up and notation

I use the term *doxastic states* to refer to the epistemic features of agents that we are trying to model. I use the term *full belief* for a binary model of an agent's doxastic state, and *credence* for a gradational model. The particular imprecise model I will discuss is that of lower previsions, and I will call such a model an *imprecise credence*.

I assume the agent's beliefs to be defined over a set of propositions, denoted by $\mathcal{F}$, which is closed under negation.[1] Propositions are themselves interpreted as sets of situations, or "possible worlds", which capture the subjective possibilities of the agent. The agent is certain that exactly one situation in a finite set $\mathcal{W}$ will obtain; or equivalently, that exactly one possible world is the actual one. A world w is an element of a proposition A iff A is true when w is the case. I will also write $A(w)$ to denote the following function from $\mathcal{W}$ to $\{0, 1\}$:

$$A(w) = \begin{cases} 1 & \text{if } w \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

I use the term *random variable* to refer to functions, such as the one above, which map worlds onto real numbers. Thus $X(w)$ is the value of a random variable X at world w.

Full beliefs are functions $b : \mathcal{F} \to \{0, 1\}$ such that:

$$b(A) = \begin{cases} 1 & \text{if } A \text{ is believed,} \\ 0 & \text{if } A \text{ not believed.} \end{cases} \tag{2}$$

Similarly, precise credences correspond to functions $cr : \mathcal{F} \to [0, 1]$, with 1 representing the highest degree of belief, and 0 the lowest.

2.2 A paradox of full belief

A researcher has just written a long book on some difficult subject. Clearly, she believes each claim in her book to be true (or she would have removed it from the book). At the same time, she is aware that all other books of this length treating such a complex subject have some false claims in them, and she sees no reason why hers should be any different. Thus, she warns the reader that there is at least one false claim ahead by adding a claim to this effect in the book's preface.

This set up is known as the *Preface paradox*. The paradox arises because the researcher believes each claim in the book to be true, including the preface claim that at least one claim is false. Therefore, her beliefs are inconsistent. Easwaran and Fitelson (2015) describe this paradox as an example of conflict between two *prima facie* reasonable epistemic norms:

[1] In other words, I assume that every agent who has a doxastic attitude towards A also has a doxastic attitude towards its negation $\neg A$.

- **Deductive Consistency (DC)**
  Agents should have beliefs that are (classically) deductively consistent, meaning that the propositions they believe must form a deductively consistent set of propositions.

- **Evidential Norm (EN)**
  Agents should have beliefs that are supported by the total evidence available to them.

In this case, the researcher's evidence (all previous books of this kind had some mistakes) supports her belief that at least one claim in her book is false. Yet, together with her initial belief that each claim is true, this leads her to a violation of deductive consistency.

Switching to gradational models of doxastic states is a known solution to the paradox. In these models, (DC) is substituted by the following norm, first introduced by Ramsey (1931) and de Finetti (1974):

- **Credal Coherence (CC)** Agents should have credences that respect the following conditions:

$$cr(A) \geq 0 \text{ for all } A \in \mathcal{F} \tag{3}$$

$$cr(\mathcal{W}) = 1 \tag{4}$$

$$\text{if } A \cap B = \emptyset, \text{ then } cr(A) + cr(B) = cr(A \vee B) \tag{5}$$

  Since these are just (a formulation of) the axioms of probability, (CC) is the requirement that the agent's credences be probability functions.

Credal Coherence does not conflict with the Evidential Norm, so gradational representations of this kind are immune from the Preface. We can model Dr. Truthlove's high confidence in each of her book's initial claims by a credence p that assigns to each claim A_i, $i = 1, ..., n$ a high degree of belief. Then, so long as the book has enough claims, it will be possible for her to have a high credence that there is a false claim in the book while remaining coherent. For example, if $p(A_i) = 0.99$ for all i and $n > 100$, she can coherently have $p(F) = 0.99$ where F is the proposition that at least one of the A_i is false.

For this reason, the Preface is often appealed to as a reason to adopt gradational models and norms of rationality over their binary counterparts (Christensen 2004). How can the defender of full belief respond to this challenge? In the remainder of this section, I look at Fitelson and Easwaran's response, which builds up to a representation of full beliefs by means of probabilities.

2.3 A coherence norm for full beliefs

Fitelson and Easwaran (2015) suggest that supporters of full belief should abandon (DC) in favour of a weaker norm, something like (CC). To derive this norm, they start from a set of assumptions that are commonly used to justify (CC) for gradational models, and go under the label of *epistemic utility theory*. First, they assume that a doxastic state's accuracy the the main source of epistemic utility, and second, that rational agents seek to maximise their epistemic utility.

They propose to measure the accuracy of full beliefs with a function:

$$S(b, w) = \sum_{A \in \mathcal{F}} s(b(A), A(w)) \tag{6}$$

where $s : \{0, 1\} \times \{0, 1\} \to \mathbb{R}$ is defined as:

$$s(b(A), A(w)) = \begin{cases} R & \text{if } b(A) = 1 \text{ and } A(w) = 1, \\ -W & \text{if } b(A) = 1 \text{ and } A(w) = 0, \\ 0 & \text{if } b(A) = 0. \end{cases} \tag{7}$$

for some positive constants $R, W \in \mathbb{R}$. This definition simply says that the agent obtains a positive amount of accuracy R (i.e. she gains epistemic utility) whenever she believes a proposition that ends up being true. The agent loses an amount W of accuracy (i.e. she loses epistemic utility) whenever she believes a proposition that ends up being false. Finally, the agent neither gains nor loses accuracy by failing to believe a proposition. The only assumption Easwaran (2016) makes on the values R and W is that $R < W$, so as to avoid giving an agent who believes both A and $\neg A$ positive utility.[2] The meaning of these values for the representation will become clear later in this section.

Once a measure of accuracy has been defined, they capture the assumption that rational agents seek to maximise their accuracy by imposing the following rationality requirement:

- **Weak Accuracy-Dominance Avoidance (WADA)**
  For a belief $b : \mathcal{F} \to \{0, 1\}$ to be rational, it is necessary that it is not *weakly accuracy-dominated* by some other belief $b' : \mathcal{F} \to \{0, 1\}$. This means that there must not exist another belief $b' : \mathcal{F} \to \{0, 1\}$ such that:

  $$S(b, w) \leq S(b', w) \text{ for all } w \in \mathcal{W}, \tag{8}$$

  with strict inequality for at least one $w \in \mathcal{W}$.

In other words, b is not rational if there is some other b' that is at least as accurate as b, and more accurate than b in some possible world. An alternative norm requires the inequality to be strict on every $w \in \mathcal{W}$, i.e. b is not rational if there is some b' strictly more accurate than it in every world. This is known as Strong Accuracy-Dominance

[2] This is mostly a matter of convention, see (Easwaran 2016, Appendix B.4).

Avoidance (SADA). I will focus on (WADA) in this essay for reasons of space, but Easwaran and Fitelson's (2016) representation can be built around either principle.[3]

Easwaran (2016) takes the above requirement *as a definition* of coherence for full beliefs. More precisely, he defines:

Definition 2.1 (Strong coherence)**.** A full belief $b : \mathcal{F} \to \{0, 1\}$ is *strongly coherent* iff it avoids weak accuracy-domination (WADA).

Then he puts forward the following norm:

- **Strong Belief Coherence (SBC)** Agents should, at any time t, have full beliefs that are strongly coherent.

2.4 The precise representation

After defining an analogue of (CC) for full belief, it remains to be shown that adopting this norm instead of (DC) saves one from the Preface paradox. To this purpose, Easwaran (2016) introduces his representation of full-belief by means of (precise) credences. The representation consists of the following definition:[4]

Definition 2.2 (Precise representation)**.** Let $b : \mathcal{F} \to \{0, 1\}$ be a full belief. We say that b is represented by a probability function (i.e. a coherent, precise credence) $p : \mathcal{F} \to \mathbb{R}$ iff b maximises the expected score under p.

It is useful to define the expectation of p as a function

$$Exp_p(X) = \sum_{w \in \mathcal{W}} p(\{w\})X(w) \tag{9}$$

where $X : \mathcal{W} \to \mathbb{R}$ is a random variable. Note that, if we think of the event $A \in \mathcal{F}$ as a random variable, then $Exp_p(A) = p(A)$. So Exp_p is really an extension of p to the set of all random variables $X : \mathcal{W} \to \mathbb{R}$. With this notation, we can say that b is represented by p iff b maximises

$$Exp_p(S(b, \cdot)) \tag{10}$$

where the value of $S(b, \cdot)$ at each world $w \in \mathcal{W}$ is the score of b at that world.

From this definition can be derived two main results:

Theorem 1. *Let $b : \mathcal{F} \to \{0, 1\}$ be a full belief. If b is represented by a probability p, then b is weakly coherent.*

[3] Note that it's easier to respect (SADA) than (WADA), i.e. (SADA) is the weaker norm.

[4] Actually Easwaran (2016, p.828) writes that b is represented by p when b maximises the expected score under p, *and* the values of b are related to those of p as in Theorem 3. I separate these two characterisations for clarity, and the content of Theorem 3 is that they are equivalent.

Theorem 2. *Let $b : \mathcal{F} \to \{0,1\}$ be a full belief. If b is represented by a probability p such that $p(\{w\}) > 0$ for all $w \in W$, then b is strongly coherent.*

These two theorems show that there are some coherent full beliefs available to the researcher in the preface scenario. To find them we can simply pick a coherent credence, such as the one in our previous example, with $p(A_i) = 0.99$ for each claim A_i, $i = 1, ...n > 100$; and $p(F) = 0.99$ for the claim that at least one A_i is false. Then the researcher need only pick a belief b represented by this probability, and she will be certain that this belief is coherent.[5] Still, this solution leaves one question open. Recall that the preface paradox arose as a conflict between Deductive Consistency (DC) and the Evidential Norm (EN). Are we sure that demanding beliefs to be coherent doesn't conflict with demanding that they reflect the agent's total evidence?

Fitelson and Easwaran (2015) answer by arguing that (EN) implies coherence. The upshot is that for any given situation, the set of full beliefs that reflect an agent's evidence will always be a subset of the set of all coherent beliefs: thus demanding that the agent be coherent will never prevent her from respecting the evidential norm. To show this, they start from the assumption that degrees of evidential support can be encoded in a probability function. Then they put forward the following necessary condition for satisfying the evidential norm (EN):

- **Necessary condition for satisfying the Evidential Norm**
  Let $b : \mathcal{F} \to \{0,1\}$ be a full belief. Then b satisfies (EN) only if there exists some probability function p such that, for a reasonable threshold $r \in [0,1]$ and for all $A \in \mathcal{F}$ we have:
  - ▷ If $b(A) = 1$, then $p(A) \geq r$,
  - ▷ If $b(A) = 0$, then $p(A) \leq r$.

The idea is that, under the assumption that the agent's evidence is encoded in a probability p, the agent's beliefs must align with that probability in order to respect (EN). The following result connects this evidential condition with the representation of belief by means of precise credences:

Theorem 3 (Easwaran 2016)**.** *Let $p : \mathcal{F} \to \mathbb{R}$ be a probability function, and let $b : \mathcal{F} \to \{0,1\}$ be a full belief. Then b maximises expected score under p iff for all $A \in \mathcal{F}$*

- *If $p(A) > \frac{W}{R+W}$ then $b(A) = 1$,*
- *If $p(A) < \frac{W}{R+W}$ then $b(A) = 0$,*
- *If $p(A) = \frac{W}{R+W}$ then $b(A)$ can be either 0 or 1.*

[5] Since the set of full beliefs over $\mathcal{F}$ is finite, at least one of them will always maximise expected score under p.

Putting together this theorem with the previous two, we now see that if we manage to encode the evidence in a probability p, and interpret the two score values W, R as defining a reasonable threshold for satisfying (EN) with regards to this probability, then the set of full beliefs b represented by p (that is, the full beliefs which maximise expected score under p) will contain all full beliefs that respect (EN).

3 Three perspectives

Enter our three main characters. Each of them reads the representation presented in the previous section from a different philosophical perspective, and uses it to advance a different philosophical project. Since I want to put forward a new representation, it will be useful to discuss these perspectives in more detail, so that I can assess my result based on how it affects each character's views.

(i) *Dr. Truthlove.*[6] Dr. Truthlove is a supporter of full belief. As her name suggests, she is especially concerned with having true beliefs, and so is happy to accept Belief Coherence as a rationality requirement, given its accuracy-based justification. The representation allows Dr. Truthlove to respond to the preface paradox in the same way probabilists do, while still maintaining that full beliefs are the real thing: probabilities are simply used as witnesses for the coherence of the full beliefs they represent.[7]

There are two main advantages Dr. Truthlove derives from using probabilities as representations. First, they give her a way to find coherent full beliefs: pick a probability on $\mathcal{F}$, and then maximise expected utility. Secondly, and more importantly, even if Dr. Truthlove does not herself recognize credences as valid epistemic models, other people do. These people have developed methods to address many of the epistemological problems that she faces. So Dr. Truthlove can hope to use the representation to adapt these methods and solutions to her binary framework.

(ii) *Dr. Locke.*[8] Dr. Locke is Dr. Truthlove's arch-rival. Although he shares Dr. Truthlove's desire for accuracy, he is a supporter of gradational rather than binary doxastic models. Indeed, he supports the Lockean thesis: this is the claim that an agent has a full belief attitude towards a proposition iff they have a credence in that proposition which is higher than a certain (possibly context-dependent) threshold. In other words, Dr. Locke thinks that to believe A *just is* to have a high credence in A. The precise representation (and particularly, Theorem 3) allows him to interpret all belief-talk (by traditional epistemologists, for example, but also by most users of natural language) as

[6]This is the perspective from which the result is presented in Easwaran (2016).

[7]Easwaran writes that "[t]he probability function is just a tool for showing that certain doxastic states are coherent. I used a probability function because it allowed me to adopt the tools of expected value and decision theory, but if some other sort of function had the same sort of connection to coherence, it would be acceptable as well" (2016, pp.830-831).

[8]This reading of the result is presented in a recent paper by Kevin Dorst (2019).

shorthand for talk about credences and thresholds.[9]

(iii) *Ms. Truthlove.*[10] Like her sister, Ms. Truthlove supports binary beliefs and values truth. However, being a judge, she is particularly interested in believing only what is supported by the evidence. Under the assumption that evidence is modeled by probabilities, Ms. Truthlove can use the representation to reach the important conclusion that belief coherence does not conflict with the evidential norm.

Ms. Truthlove only cares about probabilities insofar as they model evidence. If we found some other framework to model evidence, and from it construct the same representation, she would be just as happy to use that to derive her desired conclusion, which is a claim about the relationship between two norms for full belief.

The remainder of the essay is devoted to spelling out an imprecise representation of full beliefs which, I argue, improves on the precise one for both Dr. Truthlove and Ms. Truthlove. Dr. Locke is a lost cause: since he assumes precise credences are the real thing, any representation involving imprecise credences will add little to his view. All I can say is that it might be possible to build a reworked version of Dr. Locke's view for my representation. Whatever the appeal of the Lockean thesis may be for the supporter of precise beliefs, we can expect a defender of imprecise beliefs to find an imprecise version of the thesis similarly appealing.

4 Imprecise credence and coherence

It's time to introduce imprecise credences, the model by which I want to represent full beliefs. There is a whole family of imprecise frameworks for modeling doxastic states (Walley 2000). Here I will focus on lower previsions, so I will use the terms "imprecise credences" and "lower previsions" interchangeably. A good way to think of imprecise credences is by comparison to precise ones. If a precise credence answers the question: "How confident are you that A is true?" with a precise number $p(A) \in [0, 1]$, an imprecise credence answers the same question with lower and upper bounds $\underline{P}(A), \overline{P}(A) \in [0, 1]$ on your confidence.

For technical reasons, it is easier to discuss imprecise credences as functions defined not just on the events $\mathcal{F}$, but on a linear space $\mathcal{K}(\mathcal{W})$ of random variables $X : \mathcal{W} \to \mathbb{R}$, with $\mathcal{F} \subseteq \mathcal{K}(\mathcal{W})$. Our assumption that $\underline{P}$ is defined on a linear space allows us to define the upper bounds of your credence on each $X \in \mathcal{K}(\mathcal{W})$ as the lower bounds of your credence on $-X$.[11] More formally, given a lower prevision $\underline{P} : \mathcal{K}(\mathcal{X}) \to \mathbb{R}$, we define its

[9] I will look at how the precise representation can be used to argue for the Lockean thesis in more detail in Section 6.

[10] This view articulates the central conclusion of Fitelson and Easwaran's (2015)

[11] This is again something that has been justified only by appeal to pragmatic considerations, but we will adopt this convention anyway, as it simplifies the exposition by allowing us to work with lower bounds only.

conjugate upper prevision by:

$$\overline{P}(X) = -\underline{P}(-X) \tag{11}$$

Note that the expectation Exp_p of a probability function $p : \mathcal{F} \to \mathbb{R}$ is also a function defined on $\mathcal{K}(\mathcal{W})$, and indeed it can be seen as a lower prevision with $\underline{P}(X) = \overline{P}(X)$ on every X. We refer to the lower previsions obtained by taking the expectation of some probability function as *linear previsions*. It is sometimes helpful to think of the lower prevision $\underline{P}$ as a "credal committee", that is to identify $\underline{P}$ with the following set of *dominating linear previsions*:

$$\mathcal{M}(\underline{P}) = \{p : \mathcal{K}(\mathcal{W}) \to \mathbb{R} \mid p \text{ is a linear prevision, and } p(X) \geq \underline{P}(X) \text{ for all } X \in \mathcal{K}(\mathcal{W})\} \tag{12}$$

where each element of the set is the precise credence (and expectation) of a "member" of the committee. Intuitively, the opinions of the imprecise agent result from amalgamation of the opinions of the committee's members (Joyce 2010).[12]

We have seen in the previous section that coherence for precise credences corresponds to being a probability, and that this requirement can be justified by appealing to epistemic utility. An analogous requirement of coherence can be specified for lower/upper previsions, although this norm has not yet been justified by means of epistemic utility considerations. The justifications we do have are pragmatic, involving the betting behaviour associated with lower previsions. While this does not invalidate the results presented in following sections, we will see that they would be improved if we had access to an accuracy-based justification of imprecise coherence. Coherent lower previsions are defined as follows (Walley 1991 p. 75):

Definition 4.1 (Coherent lower prevision (on a linear space))**.** Let $\underline{P}$ be a lower prevision defined on a linear space of random variables $\mathcal{K}(\mathcal{W})$. Then $\underline{P}$ is *coherent* iff

$$\underline{P}(X) \geq \inf X \quad \text{for all } X \in \mathcal{K}(\mathcal{W}), \tag{13}$$

$$\underline{P}(\lambda X) = \lambda \underline{P}(X) \quad \text{for all } X \in \mathcal{K}(\mathcal{W}), \tag{14}$$

$$\underline{P}(X+Y) \geq \underline{P}(X) + \underline{P}(Y) \quad \text{for all } X, Y \in \mathcal{K}(\mathcal{W}). \tag{15}$$

5 An imprecise representation

This section contains my proposal for an alternative representation of belief by means of imprecise credences. The representation will be defined in decision-theoretic terms: the reason behind this choice is discussed in the first part of the section. After presenting my results, I will argue that this imprecise representation is helpful for at least some of the characters introduced in Section 2.

[12]This analogy is formalised by Theorem 4 in Appendix A.1.

5.1 The decision-theoretic approach

Fitelson and Easwaran (2015) mention that their representation result can be given a decision-theoretic reading. To do so we must think of the different full beliefs available to an agent as different gambles the agent may accept, each resulting in utility $S(b, w)$ at world w. More formally, we define the gamble:

$$X_b = \text{ gamble that results in utility } S(b, w) \text{ when } w \text{ is the case.} \tag{16}$$

And denote by $\mathcal{X}_{\mathcal{F}}$ the set of all such gambles, each corresponding to a full belief over $\mathcal{F}$. We will write $X_b(w)$ to denote the value of gamble X_b at world $w \in W$.[13]

Consider an agent whose credence is a probability p. If she were forced to accept one gamble from those in $\mathcal{X}_{\mathcal{F}}$, which one would she pick? The answer given by decision theorists is that rational agents obey the following decision rule:

- **Maximise expected utility (MEU)**
  An agent with precise credence p, when forced to accept a gamble among those in $\mathcal{X}_{\mathcal{F}}$, must choose the X_b that maximises her expected utility Exp_p.

So the agent will choose the gamble(s) $X_b \in \mathcal{X}_{\mathcal{F}}$ such that:

$$Exp_p(X_b) \geq Exp_p(X_b') \text{ for all } b'. \tag{17}$$

But by the definition of X_b, we have that the above condition holds iff b maximises expected score according to p. Together with the results of Section 2 this leaves us with the equivalences summarised by the diagram in Figure 1.

This decision-theoretic approach is interesting for us because, in the imprecise setting, a definition of representation in terms of expected utility maximisation (such as Definition 2.2) is not available. Intuitively, this is because expected utility is not uniquely defined for imprecise credences: for any gambles X, Y, the lower/upper previsions $\underline{P}, \overline{P}$ define lower and upper bounds on the expected utility of X and Y. But unless $\underline{P}(X) \geq \overline{P}(Y)$ (or vice versa), it's not obvious which of the two has higher expected utility.[14]

On the other hand, agents with imprecise credences still need to make decisions. An extensive literature exists debating the merits and demerits of various imprecise decision rules. For this reason, I give a definition schema for my imprecise representation in terms of the choices an imprecise agent would make. Depending on which rule we take to be regulating the agent's choices, the schema will produce a different definition of representation.

[13]It is no coincidence that our notation for gambles is the same as that of random variables: every random variable X can be thought of as a gamble that pays $X(w)$ when w is the case.

[14]Contrast this with a linear prevision, where $\underline{P}(X) = \overline{P}(X)$ and $\underline{P}(Y) = \overline{P}(Y)$.

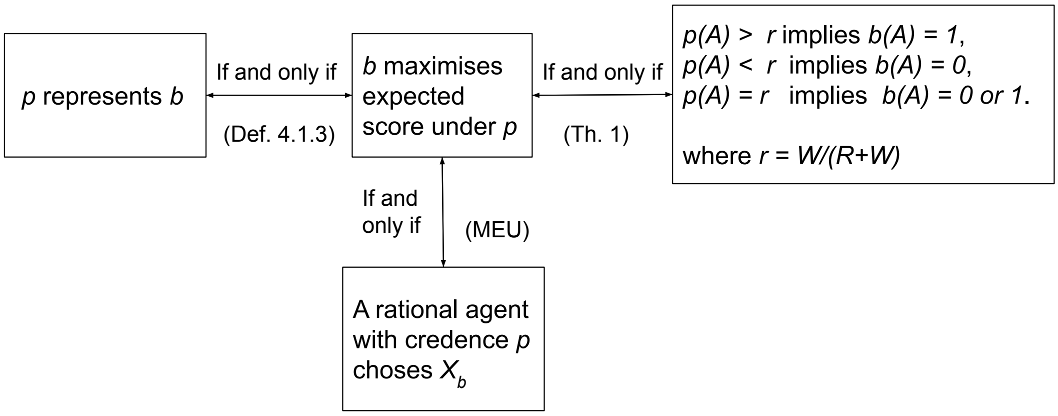

Figure 1: Equivalences derived from the precise representation of full beliefs.

Definition 5.1 (Imprecise representation)**.** A full belief $b : \mathcal{F} \to \{0, 1\}$ is represented by a coherent imprecise credence $\underline{P}$ according to a given decision rule iff the gamble X_b is *optimum in* $\mathcal{X}_\mathcal{F}$ *for* $\underline{P}$ according to that rule.[15]

Where we say X_b is *optimum in* $X_\mathcal{F}$ *for* $\underline{P}$ according to a given decision rule iff the rule allows a rational agent with imprecise credence $\underline{P}$ to choose option X_b among the $\mathcal{X}_\mathcal{F}$.

As mentioned earlier, if we restrict ourselves to precise credences, the consensus is that rational agents decide using (MEU). The definition above is in this case equivalent to the precise representation of Section 2. But as we move to imprecise credences, a number of reasonable decision rules are available.[16] Each of them, when plugged in the definition above, will produce a different representation. In this essay I will focus on the following rule:

- **Maximality**
  Consider an agent with imprecise credence $\underline{P}$ defined on a linear space $K(\mathcal{W}) \supseteq (\mathcal{F} \cup \mathcal{X}^*)$, where:

  $$\mathcal{X}^* = \mathcal{X}_\mathcal{F} \cup \{(X_b - X_{b'}) : X_b, X_{b'} \in \mathcal{X}_\mathcal{F}\}.$$

  When forced to accept a gamble among those in $\mathcal{X}_\mathcal{F}$, this agent must pick an option $X_b \in X_\mathcal{F}$ such that the following two conditions hold:

  (i) There is no $X_{b'} \in \mathcal{X}_\mathcal{F}$ such that $X_b'(w) \geq X_b(w)$ for all $w \in \mathcal{W}$, and $X_{b'}(w) > X_b(w)$ for some $w \in \mathcal{W}$.

[15]The domain of $\underline{P}$ is assumed to be the one specified in the decision rule's definition.
[16]For a good summary of IP decision rules and their properties, see (Troffaes 2007).

(ii) $\overline{P}(X_b - X_{b'}) \geq 0$ for all $X_{b'} \in \mathcal{X}_{\mathcal{F}}$.

Here is the rule's intuitive meaning. Condition (i) simply requires optima to not be dominated, whereas condition (ii) is best understood by considering what happens when it fails. In this case we have $\overline{P}(X_b - X_{b'}) < 0$, or equivalently, $\underline{P}(X_{b'} - X_b) > 0$, meaning that the agent prefers giving away X_b in exchange for $X_{b'}$ to the status quo, as the lower prevision for this exchange is greater than zero. Equivalently, thinking of the imprecise credence as a "credal committee", each member (i.e. each dominating linear prevision) agrees that $X_{b'}$ is better than X_b. So choosing X_b when $X_{b'}$ is available is clearly irrational in this scenario.

Compared to the other alternatives in the literature, Maximality gives the simplest and most useful representation. In particular, from the first condition of Maximality, it's easy to show the following analogue of Theorem 2 holds:

Proposition 1. *Let $b : \mathcal{F} \to \{0, 1\}$ a full belief. If b is represented by some coherent imprecise credence $\underline{P}$ using the Maximality definition, then b is a strongly coherent belief.*

Proof. (Appendix A.3) □

Interestingly, the converse also holds:

Proposition 2. *Let $b : \mathcal{F} \to \{0, 1\}$ a full belief. If b is strongly coherent, then b is represented by some coherent imprecise credence $\underline{P}$ using the Maximality definition.*

Proof. (Appendix A.3) □

The two results above are the reason why I will focus on the Maximality representation for the rest of the essay (see Appendix A.2 for a comparison with other IP decision rules). So from now on, by "imprecise representation", I refer to the representation obtained from Maximality.

One last thing must be mentioned before we can evaluate this imprecise representation from the perspective of our three characters. It has to do with the domain of the representing credences. In the Maximality definition, I require that $\underline{P}$ be defined not only on the set of events $\mathcal{F}$, but also on a larger domain of gambles. This ensures that the conditions imposed by the rule make sense, i.e. that they don't apply $\underline{P}$ outside of its domain. Such care was not needed in the precise case, since if we start from a probability over $\mathcal{W}$, its expectation will be uniquely defined over all gambles in $\mathcal{K}(\mathcal{W})$. In other words, there is a unique way to extend a precise probability from events to gambles on the same possibility space. On the other hand, lower previsions defined on $\mathcal{F}$ can be extended to $\mathcal{K}(\mathcal{W})$ in multiple ways. The most popular is called *natural extension*:

Definition 5.2 (Natural extension(Walley 1991, p. 122))**.** Let $\underline{P}$ be a lower prevision defined on some set $\mathcal{G}$ of gambles over a possibility space $\mathcal{W}$. Let $\mathcal{L}(\mathcal{W})$ be the set of all gambles over $\mathcal{W}$. Then the *natural extension of* $\underline{P}$ to $\mathcal{L}(\mathcal{W})$ is the function $\underline{E} : \mathcal{L}(\mathcal{W}) \to \mathbb{R}$ defined by:

$$\underline{E}(X) = \sup\{\alpha : X - \alpha \geq \sum_{j=1}^{n} \lambda_j(\underline{P}(X_j) - X_j) \text{ for some } n \geq 0, X_j \in \mathcal{G}, \lambda_j \geq 0, \alpha \in \mathbb{R}\} \tag{18}$$

Taking the natural extension of a coherent lower prevision $\underline{P}$ is the minimal way of extending it while maintaining coherence: if $\underline{E}$ is the natural extension of $\underline{P}$, and $\underline{E}'$ is a different, coherent extension, we always have $\underline{E} \leq \underline{E}'$ on $\mathcal{L}(\mathcal{W})$. In other words, coherence *requires* an agent with lower prevision $\underline{P}$ on $\mathcal{G}$ to have a lower prevision of *at least* $\underline{E}$ on $\mathcal{L}(\mathcal{W})$.[17]

If we argued that natural extensions are to be preferred over all other coherent extensions, then we could simply write a new definition of imprecise representation, saying that a coherent credence $\underline{P}$ represents a belief b whenever its natural extension $\underline{E}$ (whose domain is certainly large enough) represents b in the sense of the previous definition. I have not done so because I don't think such an argument can be given. That is, I don't think natural extensions are *in general* preferable to other coherent extensions, although an argument can be given for preferring them in some contexts.

So where does this leave us? How troubling is it that our representing imprecise credences must be defined on a larger domain than the beliefs they represent? The answer will be different depending on the character who is evaluating the representation. From some of their perspectives, the domain issue is irrelevant. From other perspectives the domain issue is relevant; but in these cases, I will argue, some justification can be given for preferring natural extensions.

6 Three perspectives revisited

It's time to compare the imprecise representation presented in the previous section with the one given by Easwaran and Fitelson. Since the precise representation was assessed from three perspectives, corresponding to our three main characters, the comparison will also be threefold.

6.1 Dr. Truthlove

Let's start with Dr. Truthlove. Recall that she is a defender of binary belief models, and that she is interested in the precise representation for two reasons: first, because it allows

[17]If we interpret the lower prevision of a gamble as an agent's supremum buying price for that gamble, then $\underline{E}(X)$ is the price at which the agent *must be willing to buy* X; the agent could also be willing to buy X for a higher price, but we cannot infer this from $\underline{P}$ alone. The same holds, *mutatis mutandis*, if we interpret lower previsions as lower bounds to confidence.

her to respond to the Preface paradox without seeing credences as anything beyond a mathematical tool; and secondly, because this tool has a rich epistemological tradition built around it, whose methods she may hope to adapt to the binary framework.

Our imprecise representation can be thought of in a similar way. Dr. Truthlove can use imprecise credences to save her from the Preface in the same way as precise ones (indeed, every probability is also a lower prevision). In fact, from her perspective, the imprecise representation improves on the precise one. This is because, although all full beliefs that have a precise representation are coherent, it's not the case that being coherent implies having a precise representation. Easwaran (2016, Appendix F) shows this with an example: for some set of propositions, and some values of W and R, there are coherent full beliefs which are not represented by any precise probability. This is bad news for Dr. Truthlove. If she relies on the representation to test for coherence, she is ruling out some perfectly good full beliefs, just because they are not represented by a probability. But if reference to probabilities is a mere tool to represent coherent beliefs, and some coherent beliefs cannot be represented in this way, then the tool should be improved. And one way to improve it is to represent beliefs as lower previsions.[18]

Note also that Dr. Truthlove can happily ignore the domain problem mentioned at the end of the last section. So what if an agent doesn't actually have an opinion about gambles of the form $(X_b - X_{b'})$? Lower previsions are not reflective of any attitude on the agent's part: they are just witness of the coherence of her beliefs. Furthermore, since precise credences are special cases of imprecise ones, all the methods that were made available to the defender of binary models by the precise representation are still available to her under the imprecise one, together with the methods and techniques that are unique to imprecise probability models, and to Maximality as a decision rule. In this sense the imprecise representation extends, rather than substitutes, the precise one.

6.2 Dr. Locke

Dr. Locke's perspective is more problematic. This isn't surprising, since he supports precise gradational models. My imprecise representation does not involve precise credences, and so it has no significance for Dr. Locke. But we should say something more. It is an interesting question whether anything like the Lockean thesis can be maintained by supporter of imprecise models. Before we can answer it, it is a good idea to review how Dr. Locke might argue for this thesis, for example by following Dorst's (2019) argument:

1. Dorst models an agent's doxastic states by a pair (cr, b) made of a precise credence cr and a full belief b. He assumes that accuracy is the fundamental epistemic virtue,

[18]Easwaran (2016, Appendix G) conjectures an alternative solution, based on a weakening the notion of coherence. The idea is to give different weights to the scores of different propositions: then b is coherent* iff it avoids accuracy-domination under *all* weight assignments. A recent paper by Rothschild (2021) shows that this strategy works if we focus on avoiding strong accuracy domination.

measuring it on credences by a proper scoring rule, and on beliefs by a score of the kind we introduced in Section 2, with parameters W, R.

2. Since incoherent credences are accuracy-dominated, he restricts himself to the pairs (cr, b) where cr is a probability.

3. He then argues that, in order to promote accuracy by her own lights, an agent must have full beliefs b which maximise accuracy according to her credence cr.

4. Using Theorem 3, he deduces that the only pairs (cr, b) that are rational are those where cr represents b. This justifies the Lockean Thesis as a requirement for rationality:[19]

   - **Lockean Thesis**
     Assume an agent's doxastic state is represented by (cr, b) with cr coherent. Then for all $A \in \mathcal{F}$, if $b(A) = 1$ it must be $c(A) \geq \frac{W}{W+R}$, and if $b(A) = 0$ it must be $c(A) \leq \frac{W}{W+R}$.

   At this stage, the *"must"* in the Lockean thesis is interpreted normatively: if the agent's doxastic state violates the Lockean condition, it is irrational.

5. Dorst then introduces the following *Pragmatist premise*: it makes sense to posit a kind of doxastic state only insofar as this helps us with "the explanation, prediction, and rationalization of the dynamics of rational agents" (Dorst 2019, p.197).

6. Since he thinks the precise representation allows us to "explain, predict, and rationalize" with credences every doxastic state that could be modeled by full beliefs, Dorst concludes that we are justified in dropping full beliefs altogether, and reads the Lockean Thesis as a metaphysical reduction. At least for rational agents, what we call "belief" in a proposition *just is* credence above a certain threshold.[20]

It is not difficult to run an imprecise version of this argument. We start from pairs $(\underline{P}, b)$, where $\underline{P}$ is an imprecise credence. As in the imprecise case, we restrict ourselves to coherent credences, but now must do so on pragmatic grounds. Then the rest of the argument goes through: unless b is represented by $\underline{P}$, it is not optimal according to her imprecise credence and the Maximality decision rule. So rational agents interested in epistemic utility, and who decide using the Maximality rule, ought to avoid incoherent beliefs.

[19] A number of different formulations of the Lockean Thesis, and a more in-depth discussion of its relationship to the precise representation can be found in Rothschild (2021).

[20] Dorst (2019) goes on to generalise the result, for example allowing different thresholds to be used for different propositions. Here I focus on the simplest case of a single threshold.

The domain problem should also not be a significant obstacle for Dr. Locke. He assumes that the lower prevision is capturing genuine features of the agent's doxastic state: if an agent is rational and makes her decisions via Maximality, she must have some way (or perhaps multiple ways) of coherently extending $\underline{P}$ to compare gambles. This fact follows from Dr. Locke's assumptions, and is sufficient to run the argument —there is no need to specify the method by which this extension is actually carried out.

Even though we can replicate Dr. Locke's argument in the imprecise case, it's not clear that its conclusion can be formulated as a Lockean thesis. The metaphysical reduction/normative correspondence drawn by the argument is a reduction/correspondence of *global* (binary) models into *global* (gradational, possibly imprecise) models. But the Lockean thesis is formulated as a *local* statement, that is, a statement about individual beliefs and credences. In the precise case, it is the claim that believing a proposition A just is having a credence in A that is above a certain threshold. And in this form the thesis just does not hold in the imprecise case. The closest we can get to it is the following result:

Proposition 3. *If a coherent lower prevision* $\underline{P} : \mathcal{F} \cup \mathcal{X}^* \to \mathbb{R}$ *represents a full-belief* $b : \mathcal{F} \to \{0, 1\}$ *in the sense of Definition 5.1, then:*

- *If* $\underline{P}(A) > \frac{W}{R+W}$ *then* $b(A) = 1$.
- *If* $\overline{P}(A) < \frac{W}{R+W}$ *then* $b(A) = 0$.
- *If* $\underline{P}(A) \leq \frac{W}{R+W} \leq \overline{P}(A)$ *then* $b(A)$ *can be either 0 or 1.*

Proof. (Appendix A.4) □

Does this count as a variant of the Lockean thesis? The main difference from the precise case is that Proposition 3 lacks a converse. Even if b respects the three conditions with regards to some coherent $\underline{P}$, it may still be incoherent, and therefore not represented by any coherent lower prevision.[21] So knowing that an agent's beliefs are locally aligned with a coherent imprecise credence is not enough to guarantee that they are represented by that credence. Representability is a global property, which requires the coherence of the agent's beliefs as a whole.

Whether these differences are important enough to thwart the Lockean project I cannot judge. I should point out, however, that all beliefs b in pairs (cr, b) that respect the precise Lockean Thesis end up being coherent. So perhaps, if our main interest is the reduction of one model to the other, it would not be unreasonable to *restrict the thesis*'s domain to pairs (cr, b) where b is coherent, especially since we *already restrict it* (point 2 of Dorst's argument) to pairs where cr is coherent. The main difference is that, while in the precise case both restrictions can be justified on accuracy grounds, we can only justify coherence of lower previsions from pragmatic considerations.

[21] For example, every b respects the conditions when $\underline{P}$ is the vacuous prevision, with $\underline{P}(A) = 0$, $\overline{P}(A) = 1$ on every A. This lower prevision is coherent (Appendix A.1).

6.3 Ms. Truthlove

Like her sister, Ms. Truthlove supports binary beliefs, yet her view is centered around the assumption that the evidential norm for beliefs can be expressed in terms of representing probabilities. More precisely, she assumes that the total evidence of the agent can be captured by a probability p over $\mathcal{F}$. Under this assumption, Ms. Truthlove can use the precise representation to prove that belief coherence does not conflict with (EN).

A problem for Ms. Truthlove's view is that, although respecting (EN) never forces one to be incoherent, there are coherent beliefs which do not correspond to any body of evidence. These are of course the coherent beliefs which lack a probabilistic representation. The fact that such beliefs exist is not as bad for Ms. Truthlove as it was for her sister, given that she is not interested in the representation as a way to establish coherence. Yet it is odd that some coherent beliefs would be *a priori* irrational, under any body of evidence. If we have no evidence at all, it seems reasonable to expect (EN) to impose no constraints whatsoever on beliefs.

There are two alternative diagnoses for this problem: either belief coherence is too weak, or probabilities over $\mathcal{F}$ are not sufficient to represent all possible evidential states. But belief coherence is justified on accuracy grounds, and Ms. Truthlove loves to be accurate. So she should modify her evidential assumption.

Let's look at a simple example.[22] In a game show, a glass urn hidden behind a curtain is filled with red and black balls. The contestant has no information about the number of balls in the urn, nor about the proportion of balls of each color. The game host randomly draws four balls from the urn, with replacement. They are all black. How does this evidence constrain the agent's belief that the next ball the host draws will be red?

Say we describe the initial lack of evidence by the uniform prior over black and red, $p(B) = P(R) = 1/2$, and then update p by conditionalization on every draw. Doing so brings the probability to $p(R) = 1/6$, $p(B) = 5/6$ after the four draws. Considering how little information the contestant has, this is suspiciously strong evidence: it commits her to believe the next ball is red with *the exact same confidence* than she has in a fair die rolling a 6, but with *strictly greater confidence* than she has in two fair dice rolls summing up to 6. It might not be irrational for the agent to reason in this way —but does her evidence really *require* all these comparative judgements? Consider also what happens if, after the four draws, the curtain is lifted, and the contestant sees that the urn contains exactly 1 red and 5 black balls. Then her evidential probability remains $p(R) = 1/6$, $p(B) = 5/6$, and so the evidential constraints on her belief remain unchanged, even though she has learned a great deal about the situation. Indeed, *now* we would be happy to say the evidence requires her to make the kind of comparative

[22]This example is inspired by Joyce (2010). A range of examples of this kind are commonly used to question the adequacy of precise probabilities as models of belief. But they can be used just as well to question their adequacy as models of total evidence.

judgements outlined above.

Imprecise probabilities can capture different evidential constraints before and after the curtain is lifted. When the curtain is down, the evidence is modeled by a lower prevision with high degree of imprecision, i.e $\overline{P}(R) - \underline{P}(R)$ is large. This will not require the agent to make any comparison between her confidence in R and rolling a 6 with a fair die. After the curtain is lifted, the lower prevision collapses into a precise probability $\overline{P}(R) = 1/6 = \underline{P}(R)$, and the evidence now requires all the comparative judgements we would expect. So although lower previsions are themselves not perfect models of evidence, they are more expressive than precise probabilities.[23] This is enough to improve Ms. Truthlove's response. We can rewrite the necessary condition for satisfying (EN) as:

- **Necessary condition for satisfying the Evidential Norm (NEN*)**: Let $b : \mathcal{F} \to 0, 1$ a full belief. Then b satisfies (EN) only if there is a coherent lower prevision $\underline{P} : \mathcal{F} \cup \mathcal{X}^* \to \mathbb{R}$ such that b is represented by $\underline{P}$ (for some values R, W). Equivalently, b satisfies (EN) only if b is coherent, and there is a coherent $\underline{P}$ such that:
  - If $\underline{P}(A) > \frac{W}{R+W}$, then $b(A) = 1$,
  - If $\overline{P}(A) < \frac{W}{R+W}$, then $b(A) = 0$.

Note that in cases of complete ignorance about A, where the evidence is captured by the vacuous prevision $\underline{P}(A) = 0, \overline{P}(A) = 1$, we have $\underline{P}(A) < \frac{W}{R+W} < \overline{P}(A)$ for any admissible R, W, and so every coherent b satisfies the above norm. In other words, our necessary condition puts no constraints on belief when the agent has no evidence, just as we wanted.

The domain of $\underline{P}$ in our evidential norm is not just $\mathcal{F}$, but $\mathcal{F} \cup \mathcal{X}^*$. I have argued that enlarging the domain is not a problem for Dr. Truthlove and Dr. Locke. But what about Ms. Truthlove? We know how to model evidence about $\mathcal{F}$ by imprecise credences, but where do the values of $\underline{P}$ on $\mathcal{X}^*$ come from? Here I think that, unless the agent somehow has direct evidence about the elements of $\mathcal{X}^*$, we should define $\underline{P}$ on $\mathcal{F}$ first, and then take its natural extension. This is because Ms. Truthlove is interested in the *constraints* imposed by the evidence on her beliefs. Taking the natural extension corresponds to extending the constraints on $\mathcal{F}$ in accord to coherence, but without making them any stronger than they need to be. Indeed, the class of beliefs represented by the natural

[23] Here is an example of the kind of evidence that cannot be captured by a lower prevision. Let $\mathcal{W} = \{w_1, w_2\}$ where $H = \{w_1\}$ is the event that a tossed coin comes up heads, and $T = \{w_2\}$ that it comes up tails. Say the agent learns the coin is biased towards tails, but does not know by what amount. It should be $\underline{P}(T - H) < \epsilon$ for all $\epsilon > 0$. This is because, if $\underline{P}(T - H) = \epsilon > 0$, then $\underline{P}$ is modeling the evidence that T is more likely than H of at least ϵ. But the agent knows of no such bound. So we must put $\underline{P}(T - H) = 0$. Yet this is also inadequate, as it fails to capture the evidence that T is more likely than H.

extension $\underline{E}$ is always larger than, or equal to, the class represented by any other coherent extension $\underline{E}'$, and thus seems more appropriate for our necessary condition.

7 Conclusion

The representation of full beliefs by means of imprecise credences has some significant advantages over its precise counterpart. The main formal improvement is the existence of a converse result (Proposition 2) which guarantees that all strongly coherent full beliefs are represented by some coherent lower prevision. To a defender of full belief (Dr. Truthlove) interested in establishing and studying belief coherence, the converse result is a great improvement, and so is the access to the methods of imprecise probability theory. It is harder to establish whether a Lockean (Dr. Locke) aiming to reduce full beliefs to his model of choice would be interested any imprecise representation, since this perspective assumes a commitment to the representing model. Still, one may wish to use the representation in a Lockean argument for imprecise credences, perhaps with the hope of some day having access to an accuracy-based justification of imprecise coherence. Finally, an evidentialist (Ms. Truthlove) should find the imprecise representation more broadly applicable than its precise counterpart, because of the greater expressive power of lower previsions as models of total evidence.

References

[1] David Christensen et al. *Putting logic in its place: Formal constraints on rational belief*. Oxford University Press on Demand, 2004.

[2] Bruno de Finetti. Theory of probability, vols. 1 & 2, 1974.

[3] Kevin Dorst. Lockeans maximize expected accuracy. *Mind*, 128(509):175–211, 2019.

[4] Kenny Easwaran. Dr. truthlove or: How i learned to stop worrying and love bayesian probabilities. *Noûs*, 50(4):816–853, 2016.

[5] Kenny Easwaran and Branden Fitelson. Accuracy, coherence, and evidence. *Oxford studies in epistemology*, 5:61–96, 2015.

[6] James M Joyce. A defense of imprecise credences in inference and decision making. *Philosophical perspectives*, 24:281–323, 2010.

[7] Frank Plumpton Ramsey. The foundations of mathematics and other logical essays. 1931.

[8] Daniel Rothschild. Lockean beliefs, dutch books, and scoring systems. *Erkenntnis*, pages 1–17, 2021.

[9] Matthias CM Troffaes. Decision making under uncertainty using imprecise probabilities. *International journal of approximate reasoning*, 45(1):17–29, 2007.

[10] Peter Walley. Statistical reasoning with imprecise probabilities. 1991.

[11] Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24(2-3):125–148, 2000.

A Appendix

A.1 A useful result

Let's start by stating a theorem which gives us a nice way to show that a lower prevision $\underline{P}$ is coherent from its corresponding set of dominating linear previsions:

Theorem 4 (Lower envelope theorem (Walley 1991))**.** *The following three statements about a lower prevision* $\underline{P} : \mathcal{K}(\mathcal{W}) \to \mathbb{R}$ *are equivalent:*

1. $\underline{P}$ *is a coherent lower prevision.*

2. $\underline{P}$ *is the lower envelope of a class of linear previsions defined over* $\mathcal{K}(\mathcal{W})$*.*

3. *For every* $X \in \mathcal{K}(W)$ *we have that* $\underline{P}(X) = p(X)$ *for some linear prevision* $p \in \mathcal{M}(\underline{P})$*.*

We can use this theorem to prove that the lower prevision defined by $\underline{P}(X) = \inf X$ for all $X \in \mathcal{K}(\mathcal{W})$ is coherent, since it is the lower envelope of the set of all linear previsions on $\mathcal{K}(\mathcal{W})$. To see this, consider any $X \in \mathcal{K}(\mathcal{W})$. X is a function from $\mathcal{W} \to \mathbb{R}$, and since $\mathcal{W}$ is finite, $\inf X = X(w)$ for some $w \in \mathcal{W}$. Then consider the linear prevision $p : \mathcal{K}(\mathcal{W}) \to \mathbb{R}$ corresponding to the expectation function of the probability that assigns $p(\{w\}) = 1$. We have $p(X) = \inf X = \underline{P}(X)$, and p clearly dominates $\underline{P}$, so $p \in \mathcal{M}(\underline{P})$. Hence $\underline{P}$ is coherent. This is known as the *vacuous prevision* (Walley, 1991).

A.2 Other IP representations

We can now look at the representations defined by IP decision rules other than Maximality. Here is a list of some of the more popular IP decision rules:

- **Γ-Maximin**:
  An agent with imprecise credence $\underline{P}$ defined on a linear space $K(\mathcal{W}) \supseteq (\mathcal{F} \cup \mathcal{X}_{\mathcal{F}})$, when forced to accept a gamble among those in $\mathcal{X}_{\mathcal{F}}$, must pick an option $X_b \in X_{\mathcal{F}}$ that maximises the *lower expected utility* $\underline{P}(X_b)$.

- **E-admissibility**:
  An agent with imprecise credence $\underline{P}$ defined on a linear space $K(\mathcal{W}) \supseteq (\mathcal{F} \cup \mathcal{X}_{\mathcal{F}})$, when forced to accept a gamble among those in $\mathcal{X}_{\mathcal{F}}$, must pick an option $X_b \in X_{\mathcal{F}}$ that maximises expected utility under some $p \in \mathcal{M}(\underline{P})$.

- **Maximality**
  Consider an agent with imprecise credence $\underline{P}$ defined on a linear space $K(\mathcal{W}) \supseteq (\mathcal{F} \cup \mathcal{X}^*)$, where:

$$\mathcal{X}^* = \mathcal{X}_{\mathcal{F}} \cup \{(X_b - X_{b'}) : X_b, X_{b'} \in \mathcal{X}_{\mathcal{F}}\}.$$

When forced to accept a gamble among those in $\mathcal{X}_\mathcal{F}$, this agent must pick an option $X_b \in X_\mathcal{F}$ such that the following two conditions hold:

(i) There is no $X_{b'} \in \mathcal{X}_\mathcal{F}$ such that $X_b'(w) \geq X_b(w)$ for all $w \in \mathcal{W}$, and $X_{b'}(w) > X_b(w)$ for some $w \in \mathcal{W}$.

(ii) $\overline{P}(X_b - X_{b'}) \geq 0$ for all $X_{b'} \in \mathcal{X}_\mathcal{F}$.

- **Interval Domination**
  An agent with imprecise credence $\underline{P}$ defined on a linear space $K(\mathcal{W}) \supseteq (\mathcal{F} \cup \mathcal{X}_\mathcal{F})$, when forced to accept a gamble among those in $\mathcal{X}_\mathcal{F}$, must pick an option $X_b \in X_\mathcal{F}$ such that $\overline{P}(X_b) \geq \underline{P}(X_{b'})$ for all $X_{b'} \in \mathcal{X}_\mathcal{F}$.

As mentioned in the main text, each of these will produce a different representation when plugged into 5.1. To see why I chose to run with Maximality, let's look at the representations produced by each rule.

A.2.1 E-admissibility

Consider first the representation based on E-admissibility. In order for b to be represented by an imprecise $\underline{P}$, its corresponding gamble X_b has to be an optimum for $\underline{P}$, Under E-admissibility, this is the case iff there is a precise probability that dominates $\underline{P}$ (i.e. an element of $\mathcal{M}(\underline{P})$) for which X_b maximises expected utility. So we have that all b represented by some imprecise $\underline{P}$ are also represented by some precise p. Furthermore, the vacuous prevision is a coherent imprecise belief, and its set $\mathcal{M}(\underline{P})$ is just the set of all linear previsions on the same domain. Restricting these linear previsions to $\mathcal{F}$, we obtain the set of all probabilities on $\mathcal{F}$. Therefore, each b that is represented by some precise credence will also be represented by the vacuous prevision if we adopt the E-admissibility definition. So if we use E-admissibility, the move to imprecise credences does not alter the class of representable beliefs, and we don't get an analogue of Proposition 2.

A.2.2 Γ-Maximin

Moving on to Γ-Maximin. When $\underline{P}$ is coherent, we have that (Walley, 1991 p.76):

$$\underline{P}(X + Y) \leq \underline{P}(X) + \overline{P}(Y)$$

The condition for X_b being a Maximin optimum, that $\underline{P}(X_b) \geq \underline{P}(X_{b'})$ for all $X_{b'}$, allows us to derive:

$$\begin{aligned}
&\underline{P}(X_{b'}) - \underline{P}(X_b) \leq 0 \text{ for all } X_{b'} \in \mathcal{X}_\mathcal{F} \\
&\iff \underline{P}(X_{b'}) + \overline{P}(-X_b) \leq 0 \text{ for all } X_{b'} \in \mathcal{X}_\mathcal{F} \\
&\implies \underline{P}(X_{b'} - X_b) \leq 0 \text{ for all } X_{b'} \in \mathcal{X}_\mathcal{F} \\
&\iff \overline{P}(X_b - X_{b'}) \geq 0 \text{ for all } X_{b'} \in \mathcal{X}_\mathcal{F}
\end{aligned}$$

so being a Γ-Maximin optimum implies respecting the second condition of Maximality.

However, Γ-Maximin optima may violate the first condition of Maximality (Troffaes 2007). We can show this by a simple example. Say an agent with vacuous lower prevision $\underline{P}$ is forced to choose between two gambles X, Y where $X(w_1) = 0 = Y(w_1)$, $X(w_2) = 1$ and $Y(w_2) = 2$. Γ-maximin allows her to choose X, even though this option is weakly dominated. So in general, Γ-Maximin-optima are not safe from weak dominance.

I say that this happens *in general*, because it's not clear that it does happen when we are considering the specific set of gambles $X_{\mathcal{F}}$. At present I can neither show that Γ-maximin fails in this way, nor that it doesn't. Furthermore, one might argue that (WADA) is too strict of a requirement anyway, and that we should be content with avoiding strong domination, which Γ-maximin optima clearly do. I have chosen to leave these questions aside for this essay, focusing on the safer representation given by Maximality, but it's worth pointing out that Γ-maximin may provide an alternative, equally interesting representation.

A.2.3 Interval Domination

Finally, let's look at the representation based on Interval Domination. In order to be optimum for $\underline{P}$, X_b must be such that $\overline{P}(X_b) \geq \underline{P}(X_{b'})$ for all $X_{b'} \in \mathcal{X}_{\mathcal{F}}$. But this condition is weaker than the second condition of maximality, because for a coherent $\underline{P}$ we have that $\overline{P}(X + Y) \leq \overline{P}(X) + \overline{P}(Y)$ (Walley 1991, p.76), and so:

$$
\begin{aligned}
&\overline{P}(X_b - X_{b'}) \geq 0 \;\text{ for all } X_b' \in \mathcal{X}_{\mathcal{F}} \\
&\implies \overline{P}(X_b) + \overline{P}(-X_{b'}) \geq 0 \;\text{ for all } X_b' \in \mathcal{X}_{\mathcal{F}} \\
&\iff \overline{P}(X_b) - \underline{P}(X_{b'}) \geq 0 \;\text{ for all } X_b' \in \mathcal{X}_{\mathcal{F}} \\
&\iff \overline{P}(X_b) \geq \underline{P}(X_{b'}) \;\text{ for all } X_b' \in \mathcal{X}_{\mathcal{F}}.
\end{aligned}
$$

So each X_b that is optimum for $\underline{P}$ under Maximality will also be optimum for $\underline{P}$ under Interval Domination. This means that an analogue of Proposition 2 will also hold for Interval Domination. However, some X_b are optimal under Interval Domination but not under Maximality, and so by Proposition 2 they are not strongly coherent. Consider as an example the case of two possible worlds $\{w_1, w_2\}$, setting $R = 2$ and $W = 3$. Let $\underline{P}$ be the vacuous prevision, and consider the belief b such that $b(\{w_1\}) = 1$ and $b(A) = 0$ for all other events $A \in \mathcal{F}$. At best (when w_1 obtains) this belief has accuracy 2, and at worse (when w_2 obtains) it has accuracy -3. So the upper and lower prevision of X_b are 2 and -3, respectively. The belief whose accuracy has highest lower prevision is b', where $b'(\{w_1, w_2\}) = 1$ and $b'(A) = 0$ for all other events $A \in \mathcal{F}$, and this lower prevision is 2. Since $\overline{P}(X_b) = 2 \geq \underline{P}(X_{b'})$, X_b is optimum under Interval Domination, and so it is represented by $\underline{P}$. Yet clearly b is incoherent, as it is accuracy-dominated by b'. Thus there is no analogue of Proposition 1 for Interval Domination.

A.3 Proof of Propositions 1 and 2

Starting with Proposition 1. The first condition of Maximality says that, for X_b to be optimum in $\mathcal{X}_{\mathcal{F}}$, there must be no $X_{b'}$ such that $X_{b'}(w) \geq X_b(w)$ for all $w \in W$, and $X_{b'}(w) > X_b(w)$ for some $w \in \mathcal{W}$. By definition of X_b, this is equivalent to the requirement that there be no full belief b' such that $S(b', w) \geq S(b, w)$ for all $w \in \mathcal{W}$, with strict inequality for some w. But this is just the requirement of Weak Accuracy-Dominance Avoidance (WADA).

Now Proposition 2. Let b be a strongly coherent belief, and let $\underline{P}$ be the vacuous prevision. Since b is strongly coherent, for each each $X_{b'} \in \mathcal{X}_{\mathcal{F}}$, we have that either $(X_b - X_{b'})(w) = 0$ for all $w \in \mathcal{W}$, or $(X_b - X_{b'})(w) > 0$ on some $w \in \mathcal{W}$. So for all $X_{b'}$, $\sup(X_b - X_{b'}) = \overline{P}(X_b - X_{b'}) \geq 0$. Thus X_b respects both conditions (i-ii) for being a Maximality optimum, and so b is represented by $\underline{P}$.

A.4 Proof of Proposition 3

Assume $\underline{P} : \mathcal{F} \cup \mathcal{X}^* \to \mathbb{R}$ is coherent, and that $b : \mathcal{F} \to \{0, 1\}$ is represented by $\underline{P}$. Then Maximality requires that b is coherent, and that $\underline{P}(X_b - X_{b'}) \geq 0$ for all b'.

To prove the first point, assume that $\underline{P}(A) > \frac{W}{R+W}$. Then by Theorem 4 we have that $p(A) > \frac{W}{R+W}$ for all $p \in \mathcal{M}(\underline{P})$. Now assume by way of contradiction that $b(A) = 0$. Define the full belief b^* as equal to b on all events except A, where $b^*(A) = 1$. Then:

$$
\begin{aligned}
& p(A) > \frac{W}{R+W} \quad \text{for all } p \in \mathcal{M}(\underline{P}) \\
& \iff p(A)R - (1 - p(A))W > 0 \;\; \text{for all } p \in \mathcal{M}(\underline{P}) \\
& \iff \sum_{i=1}^{n} p(w_i)s(1, w_i(A)) > \sum_{i=1}^{n} p(w_i)s(0, w_i(A)) \;\; \text{for all } p \in \mathcal{M}(\underline{P}) \\
& \iff \sum_{i=1}^{n} \sum_{E \in \mathcal{F}} p(w_i)S(b^*(E), w_i(E)) > \sum_{i=1}^{n} \sum_{E \in \mathcal{F}} p(w_i)S(b(E), w_i(E)) \;\; \text{for all } p \in \mathcal{M}(\underline{P}) \\
& \iff p(X_{b^*}) > p(X_b) \;\; \text{for all } p \in \mathcal{M}(\underline{P}) \\
& \iff p(X_{b^*} - X_b) > 0 \;\; \text{for all } p \in \mathcal{M}(\underline{P})
\end{aligned}
$$

By Theorem 4 again, since $\underline{P}$ is coherent, we will have that for some $p \in \mathcal{M}(\underline{P})$:

$$\underline{P}(X_{b^*} - X_b) = p(X_{b^*} - X_b) > 0 \tag{19}$$

and thus $\overline{P}(X_b - X_{b^*}) < 0$, which contradicts our assumption that $\underline{P}$ represents b. So it must be $b(A) = 1$. The second point can be proven in the same way, while the third point trivially holds because it is a conditional statement whose consequent is true by definition of full belief.