Consequences of Assigning Non-Measurable Sets Imprecise Probabilities

Final draft, forthcoming in *Mind*

Joshua Thong

Australian National University, Australia

joshuathong.zj@gmail.com

December 22, 2023

Abstract

This paper is a discussion note on Isaacs et al. 2022, who have claimed to offer a new motivation for imprecise probabilities, based on the mathematical phenomenon of *non-measurability*. In this note, I clarify some consequences of their proposal. In particular, I show that if their proposal is applied to a bounded 3-dimensional space, then they have to *reject* at least one of the following:

- If A is at most as probable as B and B is at most as probable as C, then A is at most as probable as C.
- Let $A \cap C = B \cap C = \emptyset$. A is at most as probable as B iff $(A \cup C)$ is at most as probable as $(B \cup C)$.

But rejecting either statement seems unattractive.

1 Introduction

Consider a spinner about to be spun around in a circle $\mathbb{C}$. Following Isaacs et al., let's '[s]uppose that the probabilities for [such] a spinner are fair, in that any rotation of a given set of points must have the same probability as that set of points' (Isaacs et al. 2022, p. 895). More precisely, for all subsets $A \subseteq \mathbb{C}$, if A is assigned a probability, then τA is also assigned the same probability for all rotations τ of A around $\mathbb{C}$, where τA is the set of points derived by applying rotation τ on A. Call this constraint *rotational symmetry*. Can every subset of $\mathbb{C}$ be assigned a probability that's rotationally symmetric? Orthodox probability theory answers in the negative. There will inevitably be some subsets of $\mathbb{C}$ that are assigned no rotationally symmetric probabilities, on pain of contradiction. An example of such a set is the Vitali set, whose construction is due to Vitali 1905.[1]

While the primary example employed by Isaacs et al. is that of a spinner about to be spun around $\mathbb{C}$, the primary example employed in this note is that of a unit cube $\mathbb{D}$, from which a point is about to be randomly chosen. And in line with their supposition of rotational

[1]Due to space constraints, I'll have to assume on the part of the reader knowledge of what the Vitali set is. However, the reader is free to refer to Isaacs et al. 2022 for a simple exposition on the Vitali set.

symmetry, let's suppose that the probabilities for randomly choosing a point in $\mathbb{D}$ are fair, in that any rigid motion of a given set of points in $\mathbb{D}$ must have the same probability as that set of points, where translations, rotations, and reflections are all rigid motions. More precisely, let P be a probability function[2] representing this scenario. Then, as per our supposition of rigid-motion symmetry, for all subsets A in the domain of P, $P(A) = P(\tau A)$ for all rigid motions τ of A in $\mathbb{D}$. Can every subset of $\mathbb{D}$ be assigned a probability by P, i.e. can P be defined on the powerset of $\mathbb{D}$, denoted as $\wp\mathbb{D}$? Orthodox probability theory answers in the negative again. There will inevitably be some subsets of $\mathbb{D}$ that P assigns no probabilities to, on pain of contradiction. Let's call these subsets *non-P-measurable* sets. Because of the existence of non-P-measurable sets, P cannot be defined on $\wp\mathbb{D}$.[3]

To see why there are non-P-measurable sets, assume for reductio that P is defined on $\wp\mathbb{D}$. Let $\mathcal{B} \subseteq \mathbb{D}$ be a ball (a sphere with a solid interior) such that two copies of $\mathcal{B}$ can fit nicely within $\mathbb{D}$, i.e. twice the volume of $\mathcal{B}$ is less than the volume of $\mathbb{D}$, which is 1. Banach and Tarski 1924 proved the following theorem:

Banach-Tarski Paradox. *A ball $\mathcal{B}$ is such that $\mathcal{B} = A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m$ (for some finite n and m where $n + m \geq 5$), where all the sets here are pairwise disjoint, and where there are rigid motions $\rho_1, \ldots, \rho_n$ and $\tau_1, \ldots, \tau_m$ such that $\mathcal{B}_1 = \rho_1 A_1 \cup \ldots \cup \rho_n A_n$ is a disjoint decomposition of one copy of $\mathcal{B}$ while $\mathcal{B}_2 = \tau_1 C_1 \cup \ldots \cup \tau_m C_m$ is a disjoint decomposition of another copy of $\mathcal{B}$.*[4]

Put simply, the **Banach-Tarski Paradox** states that a ball $\mathcal{B}$ can be decomposed into a finite number of parts, which can then be translated, rotated, and reassembled into *two* exact duplicates of the original ball $\mathcal{B}$. Such a theorem has consequences for probability theory.

Since we have supposed for reductio that P is defined on $\wp\mathbb{D}$, $P(\mathcal{B})$ is defined. By the **Banach-Tarski Paradox**, $P(\mathcal{B}) = P(A_1) + \ldots + P(A_n) + P(C_1) + \ldots + P(C_m)$. Now, recall that we supposed that P is rigid-motion symmetric too. So, $P(A_1)+\ldots+P(A_n)+P(C_1)+\ldots+P(C_m) = P(\rho_1 A_1)+\ldots+P(\rho_n A_n)+P(\tau_1 C_1)+\ldots+P(\tau_m C_m)$. But since $\mathcal{B}_1 = \rho_1 A_1 \cup \ldots \cup \rho_n A_n$ and $\mathcal{B}_2 = \tau_1 C_1 \cup \ldots \cup \tau_m C_m$, $P(\rho_1 A_1)+\ldots+P(\rho_n A_n)+P(\tau_1 C_1)+\ldots+P(\tau_m C_m) = P(\mathcal{B}_1)+P(\mathcal{B}_2)$. Therefore, $P(\mathcal{B}) = P(\mathcal{B}_1) + P(\mathcal{B}_2)$. Because $\mathcal{B}_1$ and $\mathcal{B}_2$ are exact duplicates of $\mathcal{B}$ and P is rigid-motion symmetric, $P(\mathcal{B}) = P(\mathcal{B}_1) = P(\mathcal{B}_2)$. This means that $P(\mathcal{B}) = 2 \cdot P(\mathcal{B})$, which in turn implies that $P(\mathcal{B}) = 0$. Let $d \subseteq \mathcal{B}$ be a cube of length $\frac{1}{k}$ for some $k \in \mathbb{N}$. Since $d \subseteq \mathcal{B}$, $P(d) = 0$. Note that any translation of d is assigned the same probability as d by P. Since d is a cube of length $\frac{1}{k}$ and $\mathbb{D}$ is a unit cube, k^3 copies of d make up $\mathbb{D}$. This means that $P(\mathbb{D}) = k^3 \cdot P(d) = 0$. Contradiction, as $P(\mathbb{D}) = 1$. Therefore, P cannot be defined on $\wp\mathbb{D}$, i.e. there are some subsets of $\mathbb{D}$ that are assigned no probabilities by P. These are the non-P-measurable sets.

Because of the existence of non-P-measurable sets, P can only be defined on a σ-algebra $\mathcal{F} \subsetneq \wp\mathbb{D}$.[5] It is important that $\mathcal{F}$ leaves some subsets of $\mathbb{D}$ out, on pain of contradiction. Some examples of subsets that $\mathcal{F}$ leaves out are the various decompositions of $\mathcal{B}$, i.e.

[2]A function P is a *probability function* iff (i) $P(A) \geq 0$ for all events A in the domain of P, (ii) $P(\Omega) = 1$, and (iii) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ if all the A_i's are pairwise disjoint and $P(A_i)$ is defined for all A_i's.

[3]The phrase 'non-P-measurable' should make it clear that some subset $A \subseteq \mathbb{D}$ isn't assigned a probability by P. However, there still exists another probability function $P^{'}$ that assigns A a probability. Pick an arbitrary member $x \in A$ and define $P^{'}(\{x\}) = 1$. So, for all subsets B of $\mathbb{D}$, if $x \in B$, then $P^{'}(B) = 1$ and $P^{'}(B) = 0$ otherwise. It is clear that $P^{'}(A) = 1$. In fact, $P^{'}$ assigns every subset of $\mathbb{D}$ a probability. So, although A is non-P-measurable, it is still $P^{'}$-measurable.

[4]This statement of the **Banach-Tarski Paradox** is due to Pruss 2014. Also, note that, strictly speaking, the word 'paradox' is a misnomer, as the **Banach-Tarski Paradox** is in fact a *theorem*. But because this theorem is known as the **Banach-Tarski Paradox** in the literature, I'll keep its name. The reason why this theorem is called a paradox is because it strikes one as counter-intuitive and odd. How can a ball be decomposed and then reassembled into two exact copies of itself?

[5]$\mathcal{F}$ is a *σ-algebra* iff (i) $\Omega \in \mathcal{F}$, (ii) if $A \in \mathcal{F}$, then $\neg A \in \mathcal{F}$, and (iii) if $A_i \in \mathcal{F}$ for all $i \in \mathbb{N}$, then $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$.

$A_1, \ldots, A_n, C_1, \ldots, C_m$. Furthermore, note that this proof of the existence of non-P-measurable sets only requires **Finite Additivity**: if $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.[6]

Why care about non-P-measurable sets in the first place? Due to space constraints, I cannot provide a detailed answer.[7] But according to Isaacs et al., there are at least two main reasons. First, not all non-P-measurable sets are created equal. There are important distinctions differentiating non-P-measurable sets, which influence our decision-theoretic intuitions. These distinctions are altogether lost if no probabilities whatsoever are assigned to non-P-measurable sets. More about this in a while. Second, assuming that rational credences are probability functions, an agent 'should not be doomed to failure' if she wants to assign a rational credence to the proposition that the chosen point is in some non-P-measurable set (Isaacs et al. 2022, p. 900).

So, for the two reasons given above, it should be at least possible to assign a probability to a non-P-measurable set, while maintaining rigid-motion symmetry. But since no precise probabilities can accomplish this task, only an *imprecise* probability (and consequently, an imprecise credence) can be assigned to such a set. Isaacs et al. deem a combination of these two reasons to be a new motivation and argument for allowing probabilities to be imprecise. Hence, according to them, *heterodox* Bayesianism is right, according to which an agent may have imprecise credences.

Up to now, all is good and fine. I agree with Isaacs et al. that there is motivation to assign a probability to at least one non-P-measurable set while maintaining rigid-motion symmetry, and since only an imprecise probability can accomplish this task, this in turn is a motivation for allowing probabilities to be imprecise. If they were to stop here, then I wouldn't have written this note. But it seems that they want to go *further*. In fact, Isaacs et al. argue that 'if imprecise credences and imprecise chances are allowed, then *no propositions need be left out*' of the probability calculus (Isaacs et al. 2022, p. 895, emphasis added). That is, according to them, *every* proposition can be assigned a probability (precise or imprecise). So, it is possible to assign an imprecise probability to not just at least one non-P-measurable set, but to *all* of them. Making this further claim is where things go awry.

In this note, I clarify the consequences of Isaacs et al.'s proposal. In particular, let $A \preceq B$ represent the statement 'A is at most as probable as B', where A and B are events. I will show that if their proposal is applied to a bounded 3-dimensional space, e.g. $\mathbb{D}$, then they have to *reject* at least one of the following:

- If $A \preceq B$ and $B \preceq C$, then $A \preceq C$.
- Let $A \cap C = B \cap C = \emptyset$. $A \preceq B$ iff $A \cup C \preceq B \cup C$.

If Isaacs et al. wish to maintain both constraints, as well as rigid-motion symmetry, then, contrary to their claim, there are some propositions bound to be left out of the probability calculus, even if probabilities are allowed to be imprecise. That is, no probabilities (precise or imprecise) can be assigned to these propositions. These claims will be proven in §3. But before §3, §2 will be devoted to *briefly* spelling out their way of assigning non-P-measurable sets imprecise probabilities.

[6]That only **Finite Additivity** is required to prove that non-P-measurable sets exist is philosophically significant. This is because in order to prove that there are subsets of $\mathbb{C}$ that are not assigned any rotationally symmetric probabilities, **Countable Additivity** is required instead: $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ if all the A_i's are pairwise disjoint and $P(A_i)$ is defined for all A_i's. But **Countable Additivity** is controversial. See de Finetti 1937 for discussion. Because of the controversy surrounding **Countable Additivity**, I decided to use the scenario of a point about to be randomly chosen in a unit cube as my primary example, instead of Isaacs et al.'s primary example of a spinner about to be spun around in a circle.

[7]However, the reader is free to refer to Isaacs et al. 2022 for a detailed answer.

It is also noteworthy that Isaacs et al. reject the usual interpretation of imprecise probabilities, according to which an imprecise probability is a supervaluation over a *set* of probability functions. Now, the reader may think that at least one of the two constraints stated above has to be rejected *no matter what* once probabilities are allowed to be imprecise. This is false, as I will show in §4 that, under the usual interpretation of imprecise probabilities, both constraints are *implied*. I will conclude in §5. All proofs are found in the Appendix.

2 Assigning imprecise probabilities

Not all non-P-measurable sets are created equal. To see why, note that *every* subset of $\mathbb{D}$ has a P*-outer measure* and a P*-inner measure*. Hereafter, let $\overline{P}$ denote P-outer measure and $\underline{P}$ denote P-inner measure. What is $\overline{P}(A)$ for an $A \subseteq \mathbb{D}$ then? Put simply, $\overline{P}(A)$ is the greatest lower bound (G.L.B.) of the probabilities of A's P-measurable supersets, i.e. $\overline{P}(A) = \inf\{P(B) : (P(B) \in \mathbb{R}) \wedge (A \subseteq B)\}$. Similarly, $\underline{P}(A)$ is the least upper bound (L.U.B.) of the probabilities of A's P-measurable subsets, i.e. $\underline{P}(A) = \sup\{P(B) : (P(B) \in \mathbb{R}) \wedge (B \subseteq A)\}$. The reader can understand $\overline{P}(A)$ as an approximation of A's probability 'from above' and $\underline{P}(A)$ as an approximation of A's probability 'from below'.

If $\overline{P}(A) = \underline{P}(A)$, then A is P-measurable, with $P(A) = \overline{P}(A) = \underline{P}(A)$. If, however, $\overline{P}(A) \neq \underline{P}(A)$, then A is non-P-measurable, i.e. $P(A)$ is *undefined*. At this point, the reader should note that $\overline{P}$ and $\underline{P}$ are *not* probability functions, as both functions are not even finitely additive. In fact, it is this failure of **Finite Additivity** that allows $\overline{P}$ and $\underline{P}$ to be defined on $\wp\mathbb{D}$. The reader should also note that while there are non-P-measurable sets A such that $\overline{P}(A) = 0.75$ and $\underline{P}(A) = 0.25$, there are also non-P-measurable sets B such that $\overline{P}(B) = 0.0001$ and $\underline{P}(B) = 0$, etc. As alluded to in §1, these are the important distinctions differentiating non-P-measurable sets. 'These distinctions are lost if we regard [non-P-measurable sets] simply as credence gaps, as orthodoxy would have it' (Isaacs et al. 2022, p. 899). For all rigid motions τ, if $\tau A \subseteq \mathbb{D}$, then $\overline{P}(A) = \overline{P}(\tau A)$ and $\underline{P}(A) = \underline{P}(\tau A)$, i.e. $\overline{P}$ and $\underline{P}$ are rigid-motion symmetric.

According to Isaacs et al., $\overline{P}$ and $\underline{P}$ provide natural bounds for assigning imprecise probabilities. Hereafter, let P_{im} be an *imprecise* probability function, i.e. P_{im} assigns some sets imprecise probabilities. In general, for all $A \in \wp\mathbb{D}$,

$$P_{im}(A) = \begin{cases} P(A) & \text{if } A \text{ is } P\text{-measurable.} \\ \left[\underline{P}(A), \overline{P}(A)\right] & \text{if } A \text{ is non-}P\text{-measurable.} \end{cases} \tag{1}$$

Since $\overline{P}$ and $\underline{P}$ are rigid-motion symmetric, P_{im} is rigid-motion symmetric too, i.e. for all subsets A in the domain of P_{im}, $P_{im}(A) = P_{im}(\tau A)$ for all rigid motions τ of A in $\mathbb{D}$.

3 Qualitative probability: probability as an ordering

Hereafter, for any two events A and B, let $A \preceq B$ represent the statement 'A is at most as probable as B'. Then, $A \succeq B$ represents the statement 'A is at least as probable as B'. Also, let $\mathcal{F}(\preceq)$ be the domain of $\preceq$. That is, if $A \preceq B$ or $A \succeq B$, then $A \in \mathcal{F}(\preceq)$ and $B \in \mathcal{F}(\preceq)$. But if either $A \notin \mathcal{F}(\preceq)$ or $B \notin \mathcal{F}(\preceq)$, then $A \not\preceq B$ and $A \not\succeq B$. In order to denote the particular ordering induced by P_{im}, as defined in the previous section, I'll use the notation '$\preceq_{\mathbb{D}}$'. It is obvious that $\mathcal{F}(\preceq_{\mathbb{D}}) \subseteq \wp\mathbb{D}$.

From $\preceq$, other relations can be defined. As usual in the literature, let Ω be a sample space and $\mathcal{F}(\preceq)$ be a σ-algebra on Ω.

Definition 1. *Let* $A \in \mathcal{F}(\preceq)$ *and* $B \in \mathcal{F}(\preceq)$.

1. $A \sim B \coloneqq (A \preceq B) \wedge (B \preceq A)$. *This represents the statement 'A is equiprobable to B'.*

2. $A \prec B := (A \preceq B) \wedge (B \not\preceq A)$. *This represents the statement 'A is less probable than B'.*

The domain of P_{im} is $\wp\mathbb{D}$. This function P_{im} induces a certain ordering of the events in $\wp\mathbb{D}$. Here's what I mean. Consider any two arbitrary events A and B in $\wp\mathbb{D}$. In general,

- If $\overline{P}(A) \leq \underline{P}(B)$, then $A \preceq_{\mathbb{D}} B$.
- If $\overline{P}(A) < \underline{P}(B)$, then $A \prec_{\mathbb{D}} B$.
- If $\overline{P}(A) = \overline{P}(B)$ and $\underline{P}(A) = \underline{P}(B)$, i.e. $P_{im}(A) = P_{im}(B)$, then $A \sim_{\mathbb{D}} B$.

At this point, it is important to note that $\preceq_{\mathbb{D}}$ is not a total ordering; it is only a *partial* ordering, i.e. there exist events A and B in $\mathcal{F}(\preceq_{\mathbb{D}})$ such that $A \not\preceq_{\mathbb{D}} B$ and $A \not\succeq_{\mathbb{D}} B$. This is because there are events A and B in $\wp\mathbb{D}$ such that $P_{im}(A) \subseteq P_{im}(B)$, e.g. $P_{im}(A) = [0.1, 0.3]$ and $P_{im}(B) = [0, 0.5]$, etc. There is no straightforward comparison between these two events. So, even though it is true that if $A \preceq_{\mathbb{D}} B$ or $A \succeq_{\mathbb{D}} B$, then $A \in \mathcal{F}(\preceq_{\mathbb{D}})$ and $B \in \mathcal{F}(\preceq_{\mathbb{D}})$, it is *false* that if $A \in \mathcal{F}(\preceq_{\mathbb{D}})$ and $B \in \mathcal{F}(\preceq_{\mathbb{D}})$, then either $A \preceq_{\mathbb{D}} B$ or $A \succeq_{\mathbb{D}} B$.

Call $\preceq$ a *qualitative probability* just in case $\preceq$ satisfies the following axioms. For all $A, B, C \in \mathcal{F}(\preceq)$,

Non-Negativity. $\emptyset \preceq A$.

Non-Triviality. $\emptyset \prec \Omega$.

Transitivity. *If* $A \preceq B$ *and* $B \preceq C$, *then* $A \preceq C$.

Qualitative Additivity. *Let* $A \cap C = B \cap C = \emptyset$. $A \preceq B$ *iff* $A \cup C \preceq B \cup C$.

Call the triple $\langle \Omega, \mathcal{F}(\preceq), \preceq \rangle$ a *qualitative probability space* just in case (i) Ω is a set, (ii) $\mathcal{F}(\preceq)$ is a σ-algebra on Ω, and (iii) $\preceq$ is a qualitative probability. I hope that the axioms above are intuitive enough for the reader to understand. So, I will not explain any of them in detail.

The reader should note that I did not assume the following as an axiom on qualitative probabilities.

Totality. *For all* A *and* B *in* $\mathcal{F}(\preceq)$, *either* $A \preceq B$ *or* $A \succeq B$.

So, it is possible that there exist A and B in $\mathcal{F}(\preceq)$ such that $A \not\preceq B$ and $A \not\succeq B$. After all, $\preceq_{\mathbb{D}}$ is only a partial ordering.

Call a qualitative probability $\preceq$ *rigid-motion symmetric* just in case for all $A \in \mathcal{F}(\preceq)$, $A \sim \tau A$ for all rigid motions τ of A in Ω. Then, since P_{im} is rigid-motion symmetric and $\preceq_{\mathbb{D}}$ is the particular ordering induced by P_{im}, $\preceq_{\mathbb{D}}$ is rigid-motion symmetric too, i.e. for all $A \in \mathcal{F}(\preceq_{\mathbb{D}})$, $A \sim_{\mathbb{D}} \tau A$ for all rigid motions τ of A in $\mathbb{D}$. Furthermore, since $P_{im}(\emptyset) = 0$ and $P_{im}(\mathbb{D}) = 1$, $\emptyset \prec_{\mathbb{D}} \mathbb{D}$, satisfying **Non-Triviality**. And because for all $A \in \wp\mathbb{D}$, $0 = \overline{P}(\emptyset) \leq \underline{P}(A)$, $\emptyset \preceq_{\mathbb{D}} A$, satisfying **Non-Negativity**. The stage is now set and I can prove that $\preceq_{\mathbb{D}}$ is not a rigid-motion symmetric qualitative probability.

Theorem 1. $\mathcal{F}(\preceq_{\mathbb{D}}) = \wp\mathbb{D}$.

Theorem 2. $\preceq_{\mathbb{D}}$ *is not a rigid-motion symmetric qualitative probability.*

Why does **Theorem 2** matter? After all, Isaacs et al. have already specified a way to assign every subset of $\mathbb{D}$ probabilities (precise or imprecise), i.e. through P_{im}. Isn't that all that matters? Why care further about the ordering induced by P_{im}? Well, **Theorem 2** matters because it highlights a tension in their project. As explained above, $\preceq_{\mathbb{D}}$ is rigid-motion symmetric and satisfies **Non-Triviality** and **Non-Negativity**. So, by **Theorem 2**, $\preceq_{\mathbb{D}}$ must admit some violations of either **Transitivity** or **Qualitative Additivity**. But

Transitivity and **Qualitative Additivity** seem to be natural axioms on any probabilistic ordering of events. Indeed, these two axioms are common across all major axiomatizations of qualitative probability, e.g. Keynes 1921, de Finetti 1937, Koopman 1940, Luce 1968, Domotor 1969, Krantz et al. 1971, Fine 1973, etc. Furthermore, Icard 2016 has provided qualitative versions of Dutch-book arguments for either axiom.

Given how natural **Transitivity** and **Qualitative Additivity** are and the qualitative Dutch-book arguments for them, Isaacs et al. should give a good reason for rejecting at least one of these axioms. At the very least, a rejection of one of these axioms must be clearly flagged and indicated. Hence, this note was written for the sake of clarifying the consequences of their proposal. While their proposal is definitely interesting, any adequate consideration of their proposal must take into account the seemingly unattractive consequence of allowing violations of either **Transitivity** or **Qualitative Additivity**.

Before closing this section, I'll clarify what follows if it is insisted that any probabilistic ordering $\preceq$ of events concerning a point about to be randomly chosen in $\mathbb{D}$ be a rigid-motion symmetric qualitative probability. Such an ordering $\preceq$ isn't defined on $\wp\mathbb{D}$, i.e. $\mathcal{F}(\preceq) \subsetneq \wp\mathbb{D}$.[8] So, there is some subset $A \subseteq \mathbb{D}$ such that $A \notin \mathcal{F}(\preceq)$, which implies that $A \not\preceq A$. This means that no probability (precise or imprecise) that is rigid-motion symmetric can be assigned to A. For if A were assigned any probability (precise or imprecise) that is rigid-motion symmetric at all, then $A \preceq A$. But $A \not\preceq A$. Therefore, contrary to Isaacs et al.'s claim, some propositions are bound to be left out of the probability calculus, e.g. the proposition that the randomly chosen point is in A. In terms of credences, this means that an agent S is 'doomed to failure' if she wants to assign a credence to the proposition that the randomly chosen point is in A (Isaacs et al. 2022, p. 900). Allowing credences to be imprecise will not help here, contrary to their claim.

4 The usual interpretation of imprecise probabilities

At this point, the reader may wonder whether it is even possible for an ordering $\preceq$ of events induced by an arbitrary imprecise probability function P_{im} to be a qualitative probability in the first place?[9] After all, for all we know, as long as $\preceq$ is induced by P_{im}, it may be *impossible* for $\preceq$ to be a qualitative probability. If so, then my critique of Isaacs et al.'s project is unfair.

In this section, I show that under the usual interpretation of imprecise probabilities, according to which an imprecise probability is a supervaluation over a *set* of probability functions, it is *possible* to define $\preceq$ such that $\preceq$ is a qualitative probability. Here's such a definition. Let $\mathcal{G}$ be a family of *precise* probability functions, the supervaluation over which P_{im} is derived from.

Definition 2. *$A \preceq B$ iff for all probability functions $P \in \mathcal{G}$, $P(A) \leq P(B)$.*

Theorem 3. *$\preceq$, as defined in **Definition 2**, is a qualitative probability.*

So, it is possible, under the usual interpretation of imprecise probabilities, to define $\preceq$ such that $\preceq$ is a qualitative probability. Therefore, it is noteworthy that Isaacs et al. reject the usual interpretation of imprecise probabilities.

> While our reasoning is friendly towards imprecise credences, it is *unfriendly* towards the standard way of interpreting them. We reject the idea of a credal committee ... Precise probability functions—even en masse—can't do what needs doing. The imprecise credences we advocate are *nothing* like sets of precise probabilities. (Isaacs et al. 2022, p. 905, emphasis added)

[8]The proof of this claim is similar to the proof of **Theorem 2**.

[9]Note that P_{im} here is an *arbitrary* imprecise probability function, and not just the function defined in §2.

Their rejection of the usual interpretation of imprecise probabilities makes more salient my demonstration that $\preceq_{\mathbb{D}}$ must admit some violations of either **Transitivity** or **Qualitative Additivity**. After all, their insistence that imprecise probabilities are nothing like sets of precise probabilities may be what compels them to embrace violations of either axiom.

5 Conclusion

In this note, I showed that Isaacs et al. need to embrace violations of either **Transitivity** or **Qualitative Additivity**, should they insist on their proposal to assign every subset of $\mathbb{D}$ a probability (precise or imprecise), while maintaining rigid-motion symmetry. But rejecting either axiom seems unattractive, given how natural they are as axioms on qualitative probabilities and the qualitative Dutch-book arguments favouring them. Any careful consideration of their proposal as it stands must take into account the consequence of rejecting either **Transitivity** or **Qualitative Additivity**. If the reader is comfortable with rejecting at least one of these axioms, then she can go ahead and embrace their proposal.[10]

[10]I'm grateful to Alan Hájek, Michael Nielsen, Nathaniel Gan, Yoaav Isaacs and an anonymous referee for their invaluable comments on this note. This work was supported by an Australian Government Research Training Program (RTP) Scholarship.

6 Appendix

*Proof of **Theorem 1**.* It is obvious that $\mathcal{F}(\preceq_{\mathbb{D}}) \subseteq \wp\mathbb{D}$. So, all that's left to be proven is $\wp\mathbb{D} \subseteq \mathcal{F}(\preceq_{\mathbb{D}})$. Let A be an arbitrary event in $\wp\mathbb{D}$. Since P_{im} is defined on $\wp\mathbb{D}$, $P_{im}(A)$ is defined, which implies that $P_{im}(A) = P_{im}(A)$. Because $\preceq_{\mathbb{D}}$ is an ordering induced by P_{im} and $P_{im}(A) = P_{im}(A)$, $A \preceq_{\mathbb{D}} A$. This means that $A \in \mathcal{F}(\preceq_{\mathbb{D}})$. Therefore, $\wp\mathbb{D} \subseteq \mathcal{F}(\preceq_{\mathbb{D}})$. Combining with the fact that $\mathcal{F}(\preceq_{\mathbb{D}}) \subseteq \wp\mathbb{D}$, $\mathcal{F}(\preceq_{\mathbb{D}}) = \wp\mathbb{D}$. □

Lemma 1. *If $A \subseteq B$, then $A \preceq B$.*

Proof. By **Non-Negativity**, $\emptyset \preceq B - A$. Apply **Qualitative Additivity** to get $A = \emptyset \cup A \preceq (B - A) \cup A = B$. □

Lemma 2. *Let $A \cap C \sim B \cap C \sim \emptyset$. $A \preceq B$ iff $A \cup C \preceq B \cup C$.*

Proof. Assume that $A \cap C \sim B \cap C \sim \emptyset$. ($\Rightarrow$) Apply **Qualitative Additivity** to $A \cap C \sim \emptyset$ to get $A = (A \cap C) \cup (A - C) \sim \emptyset \cup (A - C) = A - C$. Through similar steps, $B \sim B - C$. Since $A - C \sim A \preceq B \sim B - C$, **Transitivity** can be applied to get $A - C \preceq B - C$. Then, apply **Qualitative Additivity** to get $A \cup C = (A - C) \cup C \preceq (B - C) \cup C = B \cup C$.

($\Leftarrow$) This direction of the lemma can be proven through similar steps. □

Lemma 3. *Let $A \cap B = C \cap D = \emptyset$. If $A \preceq C$ and $B \preceq D$, then $A \cup B \preceq C \cup D$.*

Proof. See Krantz et al. 1971, pp. 211-212 for the proof. □

*Proof of **Theorem 2**.* Assume for reductio that $\preceq_{\mathbb{D}}$ is a rigid-motion symmetric qualitative probability. By the **Banach-Tarski Paradox**, $\mathcal{B} = A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m$. Then, since $\mathcal{F}(\preceq_{\mathbb{D}}) = \wp\mathbb{D}$ by **Theorem 1**, $\mathcal{B} \sim_{\mathbb{D}} A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m$. Because $\preceq_{\mathbb{D}}$ is rigid-motion symmetric, **Lemma 3** can be applied to get $A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m \sim_{\mathbb{D}} \rho_1 A_1 \cup \ldots \cup \rho_n A_n \cup \tau_1 C_1 \cup \ldots \cup \tau_m C_m$. But since $\mathcal{B}_1 = \rho_1 A_1 \cup \ldots \cup \rho_n A_n$ and $\mathcal{B}_2 = \tau_1 C_1 \cup \ldots \cup \tau_m C_m$, $A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m \sim_{\mathbb{D}} \mathcal{B}_1 \cup \mathcal{B}_2$. Apply **Transitivity** to $\mathcal{B} \sim_{\mathbb{D}} A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m$ and $A_1 \cup \ldots \cup A_n \cup C_1 \cup \ldots \cup C_m \sim_{\mathbb{D}} \mathcal{B}_1 \cup \mathcal{B}_2$ to get $\mathcal{B} \sim_{\mathbb{D}} \mathcal{B}_1 \cup \mathcal{B}_2$. Because $\mathcal{B}_1$ and $\mathcal{B}_2$ are exact duplicates of $\mathcal{B}$ and $\preceq_{\mathbb{D}}$ is rigid-motion symmetric, $\mathcal{B} \sim_{\mathbb{D}} \mathcal{B}_1 \sim_{\mathbb{D}} \mathcal{B}_2$. Apply **Transitivity** to $\mathcal{B}_1 \sim_{\mathbb{D}} \mathcal{B}$ and $\mathcal{B} \sim_{\mathbb{D}} \mathcal{B}_1 \cup \mathcal{B}_2$ to get $\mathcal{B}_1 \sim_{\mathbb{D}} \mathcal{B}_1 \cup \mathcal{B}_2$. Then, **Qualitative Additivity** can be applied to get $\emptyset \sim_{\mathbb{D}} \mathcal{B}_2$. Since $\mathcal{B} \sim_{\mathbb{D}} \mathcal{B}_1 \sim_{\mathbb{D}} \mathcal{B}_2$ and $\mathcal{B}_2 \sim_{\mathbb{D}} \emptyset$, by **Transitivity**, $\mathcal{B} \sim_{\mathbb{D}} \emptyset$. Let $d \subseteq \mathcal{B}$ be a cube of length $\frac{1}{k}$ for some $k \in \mathbb{N}$. Because $d \subseteq \mathcal{B}$ and $\mathcal{B} \sim_{\mathbb{D}} \emptyset$, $d \sim_{\mathbb{D}} \emptyset$, by **Lemma 1**. And because $\preceq_{\mathbb{D}}$ is rigid-motion symmetric, any translation ϕ of d is equiprobable to d, i.e. $\phi d \sim_{\mathbb{D}} d$. Apply **Transitivity** to $\phi d \sim_{\mathbb{D}} d$ and $d \sim_{\mathbb{D}} \emptyset$ to get $\phi d \sim_{\mathbb{D}} \emptyset$. Note that because $\phi d \sim_{\mathbb{D}} \emptyset$, $\phi d \cap d \sim_{\mathbb{D}} \emptyset$ by **Lemma 1** since $\phi d \cap d \subseteq \phi d$. Since $\phi d \cap d \sim_{\mathbb{D}} \emptyset$ and $\emptyset \cap d \sim_{\mathbb{D}} \emptyset$, **Lemma 2** can be applied to $\phi d \sim_{\mathbb{D}} \emptyset$ to get $\phi d \cup d \sim_{\mathbb{D}} \emptyset \cup d = d$. Then, **Transitivity** can be applied to $\phi d \cup d \sim_{\mathbb{D}} d$ and $d \sim_{\mathbb{D}} \emptyset$ to get $\phi d \cup d \sim_{\mathbb{D}} \emptyset$. This step can be repeated multiple times to get $\mathbb{D} \sim_{\mathbb{D}} \emptyset$, as k^3 copies of d make up $\mathbb{D}$. Contradiction, as $\emptyset \prec_{\mathbb{D}} \mathbb{D}$, by **Non-Triviality**. □

*Proof of **Theorem 3**.* The proof of **Theorem 3** is trivial but tedious. So, I will only prove that $\preceq$ satisfies **Transitivity**, but leave it to the reader to verify that $\preceq$ satisfies the rest of the axioms.

Assume that $A \preceq B$ and $B \preceq C$. These mean that for any arbitrary probability function $P \in \mathcal{G}$, $P(A) \leq P(B)$ and $P(B) \leq P(C)$. This in turn implies that for any arbitrary probability function $P \in \mathcal{G}$, $P(A) \leq P(C)$. By **Definition 2**, $A \preceq C$. □

References

Banach, Stefan, and Tarski Alfred 1924: 'Sur la décomposition des ensembles de points en parties respectivement congruentes'. *Fundamenta Mathematicae*, 6, pp. 244–77.

de Finetti, Bruno 1937: 'Foresight: Its Logical Laws, Its Subjective Sources'. In Kotz S. and Johnson N. L. 1992 (eds.), *Breakthroughs in Statistics*, pp. 134-74. New York: Springer.

Domotor, Zoltan 1969: 'Probabilistic Relational Structures and Their Applications'. Technical Report for the Institute for Mathematical Studies in the Social Sciences, 144, pp. 73-90.

Fine, Terrence 1973: *Theories of Probability: An Examination of Foundations*. New York: Academic Press.

Icard, Thomas 2016: 'Pragmatic Considerations on Comparative Probability'. *Philosophy of Science*, 83(3), pp. 348–70.

Isaacs, Yoaav, Hájek Alan, and Hawthorne John 2022: 'Non-Measurability, Imprecise Credences, and Imprecise Chances'. *Mind*, 131(523), pp. 894–918.

Keynes, John Maynard 1921: *A Treatise on Probability*. London: Macmillan.

Koopman, Bernard Osgood 1940: 'The Axioms and Algebra of Intuitive Probability'. *Annals of Mathematics*, 41(2), pp. 269-92.

Krantz, David, Luce Duncan, Suppes Patrick, and Tversky Amos 1971: *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York: Academic Press.

Luce, Duncan 1968: 'On the Numerical Representation of Qualitative Conditional Probability'. *Annals of Mathematical Statistics*, 39(2), pp. 481-91.

Pruss, Alexander 2014: 'Regular Probability Comparisons Imply the Banach–Tarski Paradox'. *Synthese*, 191, pp. 3525–40.

Vitali, Giuseppe 1905: *Sul problema della misura dei gruppi di punti di una retta*. Bologna: Tip. Gamberini e Parmeggiani.