

Better Foundations for Subjective Probability

Sven Neth

Forthcoming in *Australasian Journal of Philosophy**

Abstract

How do we ascribe subjective probability? In decision theory, this question is often addressed by representation theorems, going back to Ramsey (1926), which tell us how to define or measure subjective probability by observable preferences. However, standard representation theorems make strong rationality assumptions, in particular expected utility maximization. How do we ascribe subjective probability to agents which do not satisfy these strong rationality assumptions? I present a representation theorem with weak rationality assumptions which can be used to define or measure subjective probability for partly irrational agents.

1 Introduction

In philosophy, psychology and economics, we often ascribe subjective probability or credence to people. For example, we might say that Ann's subjective probability that it will rain tomorrow is .3. What is the basis for such ascriptions of subjective probability?

If we want to find out Ann's subjective probability that it will rain tomorrow, one natural idea is to look at which bets Ann is willing to accept. If I offer Ann a bet which pays one dollar if it rains tomorrow, how much is

*This is a preprint of a paper whose final and definite form will be published in the *Australasian Journal of Philosophy*. See <https://doi.org/10.1080/00048402.2024.2329970>. Please cite published version.

this bet worth to her? There is a long tradition in decision theory inspired by this idea, going back to Ramsey (1926). In this tradition, we assume that an agent satisfies certain principles of rationality and then *define* or *measure* subjective probability in terms of preferences.¹

But what if Ann does not satisfy the rationality assumptions required by Ramsey? Ramsey assumes that the agent under consideration is an expected utility maximizer and there are good reasons to doubt that real-life agents maximize expected utility. Now perhaps this means that real-life agents do not have any subjective probabilities or that we cannot measure them. However, there is an alternative option: we can provide decision-theoretic foundations for subjective probability with weaker rationality assumptions.

I introduce a representation theorem building on Savage (1972) and Krantz et al. (1971) which connects subjective probability to preference with much weaker rationality assumptions than standard representation theorems. In particular, I allow agents to not maximize expected utility, to violate stochastic dominance and to consider most options incomparable. The key idea is to start with comparative probability judgments and to construct a unique probability function which represents these comparative judgments. My representation theorem has important philosophical upshots: it makes sense of how we can ascribe precise subjective probability to partly irrational agents and how decision theory can provide useful advice.

As suggested above, there are two ways of understanding the project of grounding ascriptions of subjective probability in preferences. First, one might attempt to define subjective probability in terms of preference. On this view, to say that Ann's subjective probability that it will rain tomorrow is .3 just means that Ann is willing to accept certain bets. So ascriptions of subjective probability are a 'representational device' to talk about something more fundamental: Ann's preferences. Let us call this view *constructivism*. Second, one might think that while subjective probability is not reducible to preference, we can measure subjective probability by observable preferences.

¹Buchak (2017) calls this 'interpretative decision theory': we use decision-theoretic principles to interpret an agent's mental states in a process of 'radical interpretation' (Davidson 1973; Lewis 1974).

Call this view *realism*. I will mostly remain neutral between constructivism and realism. Like Ramsey’s approach, my representation theorem can be interpreted as defining or measuring subjective probability in terms of preference. However, I will later suggest that the theorem naturally fits with an intermediate position between constructivism and realism: comparative probability is psychologically real but numerical probability functions are merely ‘representational devices’ for talking about comparative probability.

Here is the plan. I start by introducing some terminology (2) and explain Ramsey’s method for measuring subjective probability (3). I discuss some problems for Ramsey’s method (4). I introduce better foundations for subjective probability (5) and explain how they overcome the problems for Ramsey’s method (6). I finish by sketching two ideas suggested by the representation theorem (7): the view that comparative probability is more fundamental than numerical probability and a subjectivist version of the classical interpretation of probability.

2 Set-up

We have a set Ω of *states* which describe the world apart from our agent’s preferences and a σ -algebra \mathcal{F} of subsets of Ω which are called *events*.² For any $X \in \mathcal{F}$, we denote the relative complement of X in Ω by X^C . We have a set \mathcal{O} of *outcomes* which contain everything our agent cares about.

Following Savage (1972), *acts* are functions from states to outcomes. An act is *finite-valued* iff it only takes finitely many different outcomes. Our *act space* \mathcal{A} is the set of all finite-valued acts. This means that we can write each act $f \in \mathcal{A}$ as $\{o_1, E_1; \dots; o_n, E_n\}$, where events E_1, \dots, E_n are pairwise disjoint sets and their union is Ω and for each E_i with $1 \leq i \leq n$, o_i is the unique outcome $o_i \in \mathcal{O}$ such that $f(\omega) = o_i$ for all $\omega \in E_i$.

I write ‘ \succsim ’ for our agent’s preference ordering over acts, a binary relation on \mathcal{A} . The intended interpretation of $f \succsim g$ is that our agent weakly prefers f to g . Strict preference (\succ) and indifference (\sim) are defined in the usual

²A σ -algebra on Ω is a set of subsets of Ω which contains Ω and is closed under complementation and countable unions.

way.³ For each outcome $o \in \mathcal{O}$, the *constant act yielding o* , written \underline{o} , is the act which assigns o to all $\omega \in \Omega$. I define, for any $o, o' \in \mathcal{O}$, $o \succsim o'$ iff $\underline{o} \succsim \underline{o'}$. I use the term ‘option’ to talk about both acts and outcomes.

3 Ramsey’s Method

Ramsey (1926) proposes axioms on preferences which imply that our agent is representable as expected utility maximizer.⁴ This means that there is some probability function and utility function such that our agent always prefers acts with higher expected utility.⁵ Ramsey then gives us a way to construct or infer our agent’s utility function without already knowing our agent’s probability function.

Once we have the utility function, Ramsey pins down the subjective probability of any event $E \in \mathcal{F}$ as follows. First, we find three outcomes $b, m, w \in \mathcal{O}$ (best, medium and worst) such that our agent strictly prefers b over w and is indifferent between getting m for certain and a bet which yields b if E happens and w otherwise:

$$b \succ w, \tag{1}$$

$$m \sim \{b, E; w, E^C\}. \tag{2}$$

Then we use the assumption that our agent maximizes expected utility to infer that $p(E) = \frac{u(m)-u(w)}{u(b)-u(w)}$.⁶ So for Ramsey, subjective probabilities are betting odds. Since event E was arbitrary, we can use Ramsey’s method to uniquely pin down the subjective probability of all events. Depending

³So $f \succ g \iff (f \succsim g) \wedge \neg(g \succsim f)$ and $f \sim g \iff (f \succsim g) \wedge (g \succsim f)$.

⁴Jeffrey (1990, ch. 3) and Bradley (2004) reconstruct Ramsey’s reasoning. Ramsey originally used a different framework which does not distinguish states and outcomes, while I am reconstructing Ramsey’s reasoning in terms of the Savage framework.

⁵Relative to probability function $p : \mathcal{F} \rightarrow [0, 1]$ and utility function $u : \mathcal{O} \rightarrow \mathbb{R}$, the *expected utility* of act $f = \{o_1, E_1; \dots; o_n, E_n\}$ is $\mathbb{E}_{u,p}(f) = \sum_{i=1}^n p(E_i)u(o_i)$.

⁶Proof: since our agent maximizes expected utility, (1) and (2) entail that $u(m) = u(w) + p(E)(u(b) - u(w))$. Therefore, $u(m) - u(w) = p(E)(u(b) - u(w))$, so $p(E) = \frac{u(m)-u(w)}{u(b)-u(w)}$. The utility function is unique up to positive affine transformation so $p(E)$ is unique.

on whether we accept constructivism or realism, we can think of this as a definition of subjective probability in terms of preferences or as a way to measure subjective probability by preferences.

Ramsey had a lasting influence on decision theory.⁷ Savage (1972) also lays down axioms on the preference relation and proves a representation theorem which shows that any agent obeying these axioms can be represented as expected utility maximizer with a unique probability function. Many of the problems for Ramsey's approach I will discuss below generalize to Savage's representation theorem. However, as we will see later, the work of Savage holds key insights for an alternative approach to measure subjective probability.⁸

4 Problems for Ramsey

I now turn to explain why Ramsey's method is not an adequate foundation for subjective probability.

4.1 Strong Rationality Assumptions

Ramsey assumes that the agent under consideration is an expected utility maximizer. However, there are good reasons to think that real-life agents are not expected utility maximizers. Suppose you only care about money and choose between the following two lotteries:⁹

1. One million dollars for certain.
2. 89 % chance of winning one million dollars, 10 % chance of winning five million dollars, 1 % chance of winning nothing.

⁷Fishburn (1981) provides a great survey of decision theory after Ramsey. Misak (2020) places Ramsey's work in its broader intellectual context.

⁸Jeffrey (1990) develops a different framework for decision theory in which states, acts and outcomes are all propositions. However, Jeffrey's axioms do not pin down a unique probability function.

⁹A *lottery* is a probability distribution over outcomes and can be realized by multiple acts.

You also choose between the following two lotteries:

3. 89 % chance of winning nothing, 11 % chance of winning one million dollars.
4. 90 % chance of winning nothing, 10 % chance of winning five million dollars.

If you strictly prefer (1) over (2) and (4) over (3), your preferences are incompatible with expected utility maximization (Allais 1953).¹⁰ However, real-life agents sometimes exhibit this pattern of preferences (Oliver 2003). My point here is not that this pattern of preferences is rationally permissible, as argued by Buchak (2013). Rather, my point is that real-life agents apparently have such ‘Allais-preferences’. Therefore, we cannot use Ramsey’s method to define or measure their subjective probabilities. However, it still seems like such agents have subjective probabilities—after all, they are presumably *using* their subjective probabilities to reason that (1) is better than (2) and (4) is better than (3). So Ramsey’s method is not a good foundation for subjective probability. To make this vivid, imagine you find out that Ann has the preferences described above. Should you conclude that Ann cannot have any subjective probabilities or that there is no way for us to find out what these probabilities are? I think not.

People also sometimes choose *stochastically dominated* options. Option A stochastically dominates option B if for every outcome $o \in \mathcal{O}$, the probability that A yields an outcome weakly preferred to o is greater than or equal to the probability that B yields an outcome weakly preferred to o . It is generally agreed that you should:

Respect Stochastic Dominance. If f stochastically dominates g , then $f \succsim g$.

¹⁰If you strictly prefer (4) over (3), $.1u(\$5 \text{ Million}) > .11u(\$1 \text{ Million})$. So, adding the same term on both sides, $.1u(\$5 \text{ Million}) + .89u(\$1 \text{ Million}) > .11u(\$1 \text{ Million}) + .89u(\$1 \text{ Million})$, which means that $.1u(\$5 \text{ Million}) + .89u(\$1 \text{ Million}) > u(\$1 \text{ Million})$. But if you strictly prefer (1) over (2), $u(\$1 \text{ Million}) > .89u(\$1 \text{ Million}) + .1u(\$5 \text{ Million})$.

This principle follows from many normative decision theories, such as expected utility theory, risk-weighted expected utility and others.¹¹ However, empirical studies show that people sometimes violate this principle. Consider the following two lotteries:

5. 5% chance of \$12, 5% chance of \$14, 90% chance of \$96.

6. 10% chance of \$12, 5% chance of \$90, 85% chance of \$96.

It is not hard to see that (5) stochastically dominates (6).¹² Nonetheless, in a study reported by Birnbaum and Navarrete (1998), most subjects chose (6) over (5). This is presumably because they rely on quick but imperfect heuristics in their decision making. Again, my claim is not that these preferences are rational but only that real-life agents have such preferences. Therefore, we cannot use Ramsey's method to define or measure their subjective probabilities. But again, while it might be irrational to have preferences which violate stochastic dominance, such preferences do not seem to preclude agents from having subjective probabilities.

Finally, people sometimes regard options as *incomparable* in value. Consider, for example, the choice between a career as a doctor and a career as a rock star. Both career choices can lead to a fulfilling and valuable life. However, what makes them valuable is radically different. It is difficult to see how one could compare the two options. Someone could reasonably think that one is not better than the other but neither are they exactly equally good (Chang 2002).

Incomparability arises both at the level of outcomes and acts. It is natural to understand the career choice example in terms of incomparable outcomes. In contrast, a second kind of incomparability might arise because it is too difficult to compare acts even if all of their outcomes are comparable. Suppose, for example, that you like more money rather than less. When faced with

¹¹Buchak (2013) and Tarsney (2020) defend stochastic dominance. Bader (2018) points out the wide applicability of stochastic dominance reasoning even if outcomes are incomparable. Russell (forthcoming) discusses some problems arising in such a setting.

¹²Both lotteries are sure to pay at least \$12. The probability of winning at least \$14 is 95% in (5) and 90% in (6). The probability of winning at least \$90 is 90% in (5) and (6).

a choice between two complicated investment portfolios, you might nonetheless not have any preference between them because it is too difficult for you to reason about the decision problem. Again, my claim is not that incomparability is rational, but merely that real-life agents sometimes have such preferences.¹³

Expected utility theory has no room for incomparability. This is because your utility function assigns a real number to each outcome and so renders all outcomes comparable. Each act is ranked by its expected utility so all acts are comparable as well. Since real-life agents sometimes regard both outcomes and acts as incomparable, their preferences cannot be represented as expected utility maximization. However, it is not plausible that incomparability precludes agents from having subjective probabilities.

Here is the upshot. There are good reasons to think that real-life agents are not expected utility maximizers. Therefore, we cannot use Ramsey's method to define or measure their subjective probabilities. Some might take this as reason to embrace a kind of nihilism: such agents do not have subjective probabilities or there is no way to measure what they are. A better response is to provide foundations for subjective probability which apply even to agents which fail to maximize expected utility, do not respect stochastic dominance and consider some options incomparable. One might still think that we should model the beliefs of real-life agents by something other than precise probability functions. However, we can make room for irrational preferences without giving up decision-theoretic foundations for precise subjective probability.

4.2 No Useful Advice

The standard decision-theoretic advice is to maximize expected utility relative to your subjective probability function and utility function. For this

¹³Many have defended the stronger claim that incomparability can be rational. Joyce (1999, p. 102) writes that 'a decision maker can be perfectly rational even when her preferences do not satisfy the completeness axiom'. Aumann (1962, p. 446) writes that '[o]f all the axioms of utility theory, the completeness axiom is perhaps the most questionable'. Similar points are made by Hare (2010), Bales, Cohen, and Handfield (2014), Schoenfield (2014), Bader (2018), and Sen (2018).

advice to be useful, we first need to figure out what your probability function *is*.¹⁴ However, if we define or measure your probability function on the assumption that you maximize expected utility, the advice to maximize expected utility can never be useful. Therefore, decision theory cannot play the role of giving useful advice.¹⁵

This puzzle arises on both constructivism and realism. For constructivists, your probability function is defined in terms of preferences which satisfy certain axioms. If you violate the axioms, you simply do not *have* a probability function and the advice to maximize expected utility is meaningless. For realists, you might still have a probability function if you violate the axioms. However, standard representation theorems give us no way to infer what this probability function is, so we cannot use decision theory to give useful advice.¹⁶

One reaction to this problem is to say that the only advice decision theory provides is: ‘obey the axioms!’. On this view, decision theory is merely a theory of consistency (Dreier 1996; Okasha 2016). While I have no knock-down objection to this position, it is unattractive because it makes decision theory largely irrelevant to non-ideal agents like us who are pretty much guaranteed to violate some normative principle of decision making. It would be better if we could make sense of how decision theory can provide useful advice to partly irrational agents. As I will show, we can indeed make sense of this, which considerably weakens the plausibility of this response.

¹⁴The same point applies to the utility function but I focus on subjective probability.

¹⁵Resnik (1987, p. 99) writes, about representation theorems in decision theory: ‘the theorem can be applied only to those agents with a sufficiently rich preference structure; and if they have such a structure, they will not need utility theory—because they will already prefer what it would advise them to prefer’. Meacham and Weisberg (2011) and Easwaran (2014) deploy similar arguments. Beck and Jahn (2021) also discuss the question of how decision-theoretic models can provide useful advice.

¹⁶This puzzle also arises for non-standard decision theories such as risk-weighted expected utility theory (Buchak 2013) and weighted-linear utility theory (Bottomley and Williamson forthcoming). On these theories, we also need an independent grasp on your subjective probability function, your utility function and possibly other functions like your risk function in order for the theory to provide useful advice.

4.3 Dependence on Utility

Ramsey defines subjective probability as ratio of utilities. This requires a very rich space of outcomes. For Ramsey, outcomes must allow for continuous gradations of value.¹⁷ However, it seems like agents can have subjective probabilities while not making such fine-grained distinctions of value. We could even imagine agents who do not have a utility function at all but merely an ordinal ranking of outcomes. For example, we can imagine an agent which only distinguishes between two outcomes, GOOD and BAD. Ramsey must deny that such an agent could have subjective probabilities or that we can find out what they are. This seems implausible.

More broadly, Ramsey's approach gives utility a certain kind of priority over subjective probability. But you might think that subjective probability is conceptually independent of utility. It would be great to disentangle the assumptions needed to measure subjective probability from strong assumptions about the structure of value, for example that the value of all outcomes is comparable and that value can be measured by a real-valued utility function. Such assumptions about value have seemed implausible to many philosophers and it would be great to have foundations for subjective probability which do not rely on them.

5 Better Foundations

We can provide better foundations for subjective probability. I introduce and explain a representation theorem building on Savage (1972) and Krantz et al. (1971) which yields a unique probability function representing our agent's beliefs on weak rationality assumptions. The key idea is to start with *comparative probability judgments* and to construct a unique probability function which represents these comparative judgments.

¹⁷For example, Fishburn (1981, p. 152) writes that in Ramsey's approach, the set of outcomes 'must be infinite and give arbitrarily fine gradations in utility'.

5.1 Comparative Probability

What does it mean to think that one event is more probable than another? Savage (1972) proposes to define comparative probability judgments in terms of certain kinds of preferences. Suppose our agent strictly prefers outcome b over outcome w . Now the intuition is that our agent *prefers the better prize on the more probable event*. So if our agent prefers the act $\{b, X; w, X^C\}$ over $\{b, Y; w, Y^C\}$, this means that our agent believes that event X is at least as likely as event Y , written $X \succcurlyeq Y$. So we can use acts of the form $\{b, X; w, X^C\}$, where $b \succ w$, to define or infer our agent's comparative probability judgments. Let us call these *test acts*.

We define the following relation \succcurlyeq on \mathcal{F} :

Definition 1. $X \succcurlyeq Y$ iff $\{b, X; w, X^C\} \succeq \{b, Y; w, Y^C\}$ for some $b, w \in \mathcal{O}$ with $b \succ w$.

Strict comparative probability (\succ) and indifference (\approx) are defined in the standard way.¹⁸

We can understand Savage's proposal in two ways. For constructivists, comparative probability reduces to preferences. (This is Savage's own view.) For realists, comparative probability does not reduce to preferences, but we can use preferences to measure comparative probability judgments. The core proposal of this paper is compatible with both ways of understanding comparative probability. However, I think that realism about comparative probability is ultimately more plausible and can tell a better story about some of the axioms below. I will return to this issue later.

5.2 Axioms

The first axiom ensures that the comparative probability ordering does not depend on our particular choice of outcomes:

¹⁸ $X \succ Y \iff (X \succcurlyeq Y) \wedge \neg(Y \succcurlyeq X)$ and $X \approx Y \iff (X \succcurlyeq Y) \wedge (Y \succcurlyeq X)$. In a slight abuse of notation, I use the same symbol (\succ) for both strict preference and strict comparative probability. This way to link comparative probability to preferences is standard (Fishburn 1986; Icard 2016).

Outcome Independence. For all $X, Y \in \mathcal{F}$, if $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ for some $b, w \in \mathcal{O}$ such that $b \succ w$, then $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ for all $b, w \in \mathcal{O}$ such that $b \succ w$.

You would violate this axiom if you prefer to bet one dollar on event X rather than event Y but you also prefer to bet two dollars on Y rather than X . In this case, we cannot elicit stable comparative probability judgments from your preferences.

The next axiom demands that our agent is not indifferent among all outcomes:

Non-Degeneracy. There are outcomes $b, w \in \mathcal{O}$ with $b \succ w$.

What is the status of this axiom? Does the existence of subjective probability really require that you are not indifferent between all outcomes? Eriksson and Hájek (2007) point out that we can imagine a Zen monk who is indifferent between all outcomes but nonetheless has subjective probabilities. Thus, there are problems with **Non-Degeneracy** understood along constructivist lines. However, if we are realists, we can accept **Non-Degeneracy** as a condition under which we can measure comparative probability. The Zen monk might have subjective probabilities, but if they are really indifferent among everything, there is simply no way for us to find out what these subjective probabilities are. So we can think of this axiom as a *structure axiom* which ensures that preferences are rich enough to measure subjective probability.¹⁹

Here is the third axiom:

Restricted Ordering. The relation \succsim restricted to test acts with the same outcomes is complete and transitive. This means for any $b, w \in \mathcal{O}$ with $b \succ w$, for all $X, Y \in \mathcal{F}$ we have either $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ or $\{b, Y; w, Y^C\} \succsim \{b, X; w, X^C\}$.²⁰

¹⁹Joyce (1999, p. 82) distinguishes structure axioms and rationality axioms.

²⁰Two test acts with different outcomes needn't be comparable if the outcomes themselves are incomparable. Thanks to an anonymous referee for bringing this issue to my attention.

And if $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ and $\{b, Y; w, Y^C\} \succsim \{b, Z; w, Z^C\}$, then $\{b, X; w, X^C\} \succsim \{b, Z; w, Z^C\}$.

I do not constrain the preference relation in general to be complete and transitive, which leaves room for incomparability.

Why accept **Restricted Ordering**? Given our definition of comparative probability, **Restricted Ordering** requires that the comparative probability judgments of our agent are complete and transitive. There are reasons to be skeptical of both.²¹ For proponents of imprecise credences, rejecting completeness is particularly natural. Perhaps you have some opinion about how likely it is that there is life on mars and some opinion about how likely it is to rain tomorrow but no opinion about which is more likely. This seems particularly plausible if we consider agents which are not perfectly rational.

In response, remember that I want to explain how it is possible to ascribe *precise* credences to agents with some irrational preferences. For this reason, I will not consider agents whose comparative probability judgments fail to be complete and transitive. Such agents fall outside of the scope of my project.

The next two axioms are where the main action is. Let us begin with:

Certain Prize. For any $b, w \in \mathcal{O}$, if $b \succ w$, then for any $X \in \mathcal{F}$, $\underline{b} \succsim \{b, X; w, X^C\}$ and $\{b, X; w, X^C\} \succsim \underline{w}$.

This principle states a plausible minimal rationality condition. It says that if you strictly prefer b to w , then you must weakly prefer getting b for certain to an act which yields b if X happens and w otherwise. Further, you must weakly prefer this act to getting w for certain.

While **Certain Prize** is quite weak, it is possible to imagine agents which violate this axiom. For example, agents might prefer a risky option over a sure thing because they enjoy the thrill of gambling.²² Relatedly, **Certain Prize** might be violated by agents who prefer randomization (Icard 2021). One response to this concern is to make more fine-grained distinctions

²¹Fishburn (1986, p. 339) discusses examples in which comparative probability judgments violate completeness and transitivity. Ding, Holliday, and Icard (2021) study logics for comparative probability without completeness.

²²Thanks to an anonymous referee for raising this objection.

among outcomes (Dreier 1996). For example, a prize obtained for sure would be a different outcome from the same prize obtained by a risky gamble. However, this move threatens to trivialize decision-theoretic norms. So it is best to concede that while the axiom is weaker than rationality axioms in standard representation theorems, it still makes substantive demands which some agents might violate.

Do agents which violate **Certain Prize** not have subjective probabilities? This is not very plausible. After all, it is precisely their subjective probabilities which lead them to prefer the risky option. It is more plausible to think that if agents love the thrill of gambling, it might be difficult to determine their subjective probabilities from their preferences. As I show below, **Certain Prize** is a necessary condition for the agent's comparative probability judgments to be representable by a probability function, so measuring the subjective probability of agents which violate **Certain Prize** would require a fundamentally different approach to measuring subjective probability.

Here is another key axiom:

Alternative Prize. For any $X, Y, Z \in \mathcal{F}$ and $b, w \in \mathcal{O}$, if $b \succ w$ and Z is such that $X \cap Z = Y \cap Z = \emptyset$, then $\{b, X; w, X^C\} \succ \{b, Y; w, Y^C\}$ iff $\{b, X \cup Z; w, (X \cup Z)^C\} \succ \{b, Y \cup Z; w, (Y \cup Z)^C\}$.

Alternative Prize says the following. Suppose you strictly prefer b over w and you prefer $\{b, X; w, X^C\}$ over $\{b, Y; w, Y^C\}$. Now we modify both acts as follows: You also get b if some event Z disjoint from both X and Y happens. Now you should prefer $\{b, X \cup Z; w, (X \cup Z)^C\}$ to $\{b, Y \cup Z; w, (Y \cup Z)^C\}$. This reasoning also works backwards. **Alternative Prize** has a clear interpretation in terms of probability. If you prefer $\{b, X; w, X^C\}$ to $\{b, Y; w, Y^C\}$, you think that X is at least as likely as Y . Therefore, $X \cup Z$ must be at least as likely as $Y \cup Z$ given that Z is disjoint from both X and Y . So you should prefer $\{b, X \cup Z; w, (X \cup Z)^C\}$ to $\{b, Y \cup Z; w, (Y \cup Z)^C\}$ since you want the better prize on the more probable event.

You might violate **Alternative Prize** if you have credences which are not additive and represented by an alternative formalism like Dempster-Shafer

functions or ranking theory.²³ But my goal is to provide foundations for ascribing subjective *probability* to partly irrational agents. So agents modeled by such formalisms fall outside the scope of my project.²⁴ It would be desirable to have more general foundations for measuring belief which apply to agents with non-probabilistic credences, but I will not consider such agents here.

My axioms on the preference relation are necessary and sufficient for the comparative probability ordering to be a *qualitative probability* (de Finetti 1931):

Definition 2. A binary relation \succsim on \mathcal{F} is a qualitative probability iff for all $X, Y, Z \in \mathcal{F}$:

1. \succsim is complete and transitive (*Ordering*),
2. $\Omega \succsim X \succsim \emptyset$ (*Boundedness*),
3. $\Omega \succ \emptyset$ (*Non-Triviality*),
4. if $X \cap Z = Y \cap Z = \emptyset$, then $X \succ Y \iff X \cup Z \succ Y \cup Z$ (*Qualitative Additivity*).

Theorem 1. The preference relation \succsim satisfies **Outcome Independence**, **Non-Degeneracy**, **Restricted Ordering**, **Certain Prize** and **Alternative Prize** if and only if the comparative probability ordering \succsim is a qualitative probability.

A proof is provided in the appendix. I follow Savage’s definition of comparative probability. Savage also assumes **Outcome Independence** and **Non-Degeneracy**. The key difference is that Savage uses much stronger axioms to derive the result that the comparative probability ordering is a qualitative probability. Instead of **Restricted Ordering**, Savage assumes

²³Ellsberg (1961) gives an example of preferences which violate **Alternative Prize**.

²⁴Titelbaum (2022, Ch. 14.3) gives a brief introduction to Dempster-Shafer functions and Spohn (2012) discusses ranking theory. As both authors note, it is unclear how these alternatives to probability interact with decision making, which is a reason to set them aside for our purposes. Thanks to an anonymous referee for the suggestion to consider these frameworks.

that preferences are complete and transitive, which rules out incomparable options. This strong assumption is unnecessary to establish that the comparative probability ordering is complete and transitive. It suffices to assume that a small fragment of the preference relation is complete and transitive.

Further, Savage appeals to the ‘Sure-Thing Principle’ in order to establish that the comparative probability ordering satisfies Boundedness, Non-Triviality and Qualitative Additivity. The Sure-Thing-Principle is a strong axiom which rules out the Allais-preferences discussed earlier and plays a crucial role in establishing the existence of an expected utility representation. The key observation is that we can replace the Sure-Thing-Principle by the much weaker rationality axioms **Certain Prize** and **Alternative Prize** and still show that the comparative probability ordering is a qualitative probability.²⁵

Krantz et al. (1971, pp. 208-11) prove a similar result.²⁶ But instead of **Restricted Ordering**, they assume that preferences are complete and transitive, which rules out incomparable options. Furthermore, I have shown that my axioms are not only sufficient but necessary for the comparative probability ordering to be a qualitative probability. So my result is a strengthening of Krantz et al. (1971), maximally paring down the axioms on the preference relation required to show that the comparative probability ordering is a qualitative probability.

One could also axiomatize comparative probability directly and argue that the qualitative probability axioms are reasonable constraints on belief without trying to justify them by more fundamental axioms about prefer-

²⁵Machina and Schmeidler (1992) also weaken Savage’s axiom to give a ‘more robust definition of subjective probability’. However, their axioms are stronger than the ones given here, as they entail that preferences always respect stochastic dominance—a property they refer to as ‘probabilistic sophistication’—and that preferences are complete. My representation theorem shows how to define subjective probability *without* probabilistic sophistication (and without completeness). Elliott (2017) provides a representation theorem for ‘frequently irrational’ agents and uses a restricted class of two-outcome acts to construct a unique credence and utility function. However, this credence function is not necessarily a probability function, so this approach does not provide foundations for ascribing subjective *probability* to partly irrational agents.

²⁶**Outcome Independence** is equivalent to the first axiom by Krantz et al. (1971), **Certain Prize** is equivalent to their second axiom and **Alternative Prize** is equivalent to their third axiom. They also mention **Non-Degeneracy**.

ences (Joyce 1999, p. 91). However, my project is to show how we can infer credences from preferences without already assuming that we have access to comparative probability judgments. Therefore, I start with axioms on the preference relation.

So far, we have seen how preference reveals qualitative probability. How do we get from qualitative probability to quantitative probability? Probability function p represents the qualitative probability \succsim if for all $X, Y \in \mathcal{F}$,

$$p(X) \geq p(Y) \iff X \succsim Y.$$

The axioms introduced so far are necessary but not sufficient for the existence of a probability function representing our qualitative probability (Kraft, Pratt, and Seidenberg 1959). To get around this problem, I add an axiom which ensures that the space of events is sufficiently rich to pin down a (unique) probability function. Here is Savage's proposal:

Event Richness. For all $X, Z \in \mathcal{F}$ and outcomes $b, w \in \mathcal{O}$ with $b \succ w$, if $\{b, X; w, X^C\} \succ \{b, Z; w, Z^C\}$, there is a finite partition $\mathcal{Y} = \{Y_1, \dots, Y_n\}$ of Ω such that for all $Y_i \in \mathcal{Y}$, $\{b, X; w, X^C\} \succ \{b, (Z \cup Y_i); w, (Z \cup Y_i)^C\}$.

This axiom says that we can cut up events very finely. If you strictly prefer the good prize on X rather than Z , there is a finite partition of our state space such that you still prefer the good prize on X rather than Z or one of the elements of our partition. It is instructive to state **Event Richness** in terms of comparative probability. In these terms, it says that if $X \succ Z$, then there exists a finite partition $\mathcal{Y} = \{Y_1, \dots, Y_n\}$ of Ω such that for all $Y_i \in \mathcal{Y}$, $X \succ Z \cup Y_i$.

Why accept **Event Richness**? Savage (1972, p. 38) gives the following argument. Suppose you judge X to be more probable than Z . Savage points out that we could plausibly choose a coin and throw it sufficiently often such that you would still judge X to be more probable than Z or any particular sequence of heads and tails. As Savage notes, this doesn't require that you consider the coin to be fair. The possible outcomes of the coin flip form the required partition.

Let us end by briefly reflecting on the plausibility of **Event Richness**. Does rationality require that you cut up events very finely? Despite Savage’s argument, this does not seem very plausible. Like **Non-Degeneracy**, we should think about **Event Richness** not as rationality axiom but rather as structure axiom which ensures that preferences are rich enough to fix subjective probability. This means that realism can tell a more plausible story about this axiom than constructivism. According to the realist story, it is not the existence of subjective probability which requires such a rich event space. Rather, the rich event space is necessary to infer (precise) probability from preference.

Event Richness implies that Ω is infinite.²⁷ This might strike you as problematic because it seems possible to have subjective probabilities with a finite state space. One option is to look for another structure axiom which is compatible with finite state spaces but still allows us to derive a unique probability function. As Luce (1967) and Fishburn (1986) point out, there are such axioms, but they are rather complicated and do not have intuitive plausibility of **Event Richness**. Since we need some structure axiom anyways, it seems best to stick with **Event Richness** because of its intuitive plausibility. However, finding a good replacement for **Event Richness** which is compatible with finite state spaces is a way in which the representation theorem could be improved.²⁸

5.3 Representation Theorem

The axioms allow us to prove:

²⁷Proof sketch: Assume Ω is finite. Then consider the least probable event X such that $X \succ \emptyset$. **Event Richness** demands that there exists a finite partition \mathcal{Y} of Ω such that for all $Y_i \in \mathcal{Y}$, $\{b, X; w, X^C\} \succ \{b, Y_i; w, Y_i^C\}$, so $X \succ Y_i \succ \emptyset$, which contradicts our assumption.

²⁸Another option would be axioms ensuring that the comparative ordering can be represented by some probability function which needn’t be unique (Scott 1964). However, this is not compatible with providing decision-theoretic foundations for *precise* subjective probability and so I will set it aside. One could also argue that comparative probability orderings on finite spaces should be *extendable* to orderings on infinite state spaces which satisfy **Event Richness**.

Theorem 2. *If the preference relation \succsim satisfies **Outcome Independence**, **Non-Degeneracy**, **Restricted Ordering**, **Certain Prize**, **Alternative Prize** and **Event Richness**, there is a unique finitely additive probability function $p : \mathcal{F} \rightarrow [0, 1]$ representing the comparative probability ordering \succcurlyeq , so for all $X, Y \in \mathcal{F}$,*

$$p(X) \geq p(Y) \iff X \succcurlyeq Y.$$

Once we have shown that the comparative probability ordering is a qualitative probability, the rest of the proof is due to Savage. Here is a quick proof sketch inspired by Kreps (1988, pp. 120-125):

Proof. The axioms entail that for any $n \in \mathbb{N}$, there is a partition \mathcal{Y} of Ω into n equiprobable events: events such that $Y_i \approx Y_j$ for each $Y_i, Y_j \in \mathcal{Y}$.²⁹ We write $C(k, n)$ for a union of k cells of this partition. We define, for any $X \in \mathcal{F}$:

$$k(X, n) = \max_k (X \succcurlyeq C(k, n)).$$

So given a n -fold equiprobable partition, $k(X, n)$ is the unique maximal positive integer such that X is at least as probable as the union of k cells of our partition. We define

$$p(X) = \lim_{n \rightarrow \infty} \frac{k(X, n)}{n}.$$

One can show that p is a finitely additive probability function which represents \succcurlyeq and that it is unique. \square

In this proof, we divide Ω into more and more fine-grained equiprobable partitions. For every such partition, we ‘approximate’ $p(X)$ by the largest number of cells collectively less probable (according to our comparative probability ordering) than X divided by the number of all cells. Step by step, we get a closer approximation, until we recover the ‘true’ probability of X in the limit. As a simple analogy, think of approximating the area of a two-dimensional figure by drawing more and more fine-grained grids and

²⁹Fishburn (1970, pp. 195-8) provides a detailed reconstruction of this step of the proof. Gaifman and Liu (2018) discuss how it relies on the assumption that the events form a σ -algebra.

counting the number of squares covered by the figure divided by the number of all squares. As the grid gets more and more fine-grained, we approximate the area of our figure more and more closely and we recover the true area in the limit.

We can think of the theorem in two ways. If we are inclined towards constructivism, we can think of it as a definition of subjective probability in terms of preferences. In this case, the fact that Ann's subjective probability of rain tomorrow is .3 is *constituted* by the fact that

$$\lim_{n \rightarrow \infty} \frac{k(\text{rain}, n)}{n} = .3,$$

and from this perspective, my axioms are conditions under which subjective probability exists. If we are inclined towards realism, we think that there is a probability function encoding Ann's beliefs not defined in terms of her preferences. From a realist point of view, we can interpret the proof as giving an algorithm to *measure* Ann's subjective probability by constructing better and better approximations. From this perspective, my axioms are conditions under which subjective probability can be measured by this algorithm.

5.4 Countable Additivity

As it stands, the representation theorem delivers a *finitely additive* probability function which represents our agent's beliefs. This probability function might fail to be *countably additive*.³⁰ Some decision theorists, for example de Finetti and Savage, have argued that rationality only requires finite additivity and violations of countable additivity are fine. However, there are also reasons to want countable additivity. Most importantly, there are convergence theorems in Bayesian statistics which show that under certain conditions, agents with different priors converge to similar opinions after

³⁰The probability function $p : \mathcal{F} \rightarrow [0, 1]$ is countably additive if for any countable sequence X_1, X_2, \dots of pairwise disjoint events in \mathcal{F} , $p(\bigcup_{n=1}^{\infty} X_i) = \sum_{n=1}^{\infty} p(X_i)$.

learning enough shared evidence.³¹ Many of these convergence theorems require countable additivity (Elga 2016). So if convergence is a central part of your conception of subjective probability, finitely additive probability is not enough. This is not the place to settle whether arguments for countable additivity are conclusive. The key point is that if countable additivity is desirable, we can add another plausible axiom on preferences to ensure that subjective probabilities are countably additive, building on work by Villegas (1964). Details are in the appendix.

6 Problems Solved

I explain how my representation theorem does better than Ramsey’s method.

6.1 Weak Rationality Assumptions

My axioms do not entail that our agent is an expected utility maximizer. They do not even entail the weaker claim that our agent always respects stochastic dominance. A quick way to see this is that my axioms only constrain preferences over a very restricted set of acts—two-outcome acts where one outcome is strictly preferred—while expected utility maximization and stochastic dominance constrain preferences over all acts. My axioms do not even require preferences over arbitrary acts to be transitive. So the axioms are compatible with Allais-preferences and violations of stochastic dominance.

Further, the axioms allow agents to consider many options incomparable. To be sure, **Non-Degeneracy** requires the existence of at least two comparable outcomes. However, the axioms allow agents to consider *all other* outcomes incomparable. Thus, we can make room for the kind of outcome incomparability discussed above (career as a doctor vs. career as a rock star). Furthermore, we allow agents to consider acts with more than two outcomes

³¹Subjective Bayesians draw on such convergence theorems to argue that, despite different priors, rational agents will agree in the long run (Earman 1992, ch. 6). Convergence arguments also play an important role in some versions of objective Bayesianism (Neth 2023).

incomparable even if all outcomes are comparable, like in the complicated portfolio choice discussed earlier. So, speaking a bit loosely, we allow agents to consider almost all options incomparable.

I still make substantive rationality assumptions. In particular, as discussed above, we can imagine agents which violate **Certain Prize** and **Alternative Prize**. It would be desirable to have even more general foundations for subjective probability. But there is a trade-off between substantive rationality axioms which allow us to measure subjective probability but exclude some agents and weak rationality axioms which include these agents but might make measuring subjective probabilities impossible. In particular, Theorem 1 shows that my rationality axioms are necessary conditions for the agent's comparative probability judgments to be representable by a probability function. The comparative probability judgments of agents which violate these axioms cannot be represented by any probability function. So if we want to further weaken these axioms, a fundamentally different approach to measuring subjective probability is needed.

6.2 Useful Advice

As explained above, my axioms allow agents to have some irrational preferences. Thus, we can give useful advice. We can define or measure the subjective probabilities of partly irrational agents from their preferences over simple acts and use these probabilities to give useful advice for how to choose among more complicated acts.

You might complain that my axioms are too weak because they only constrain preferences over test acts. But this is a feature, not a bug. We can measure or define subjective probability from preferences over test acts and then apply your favorite decision-theoretic norm to give advice for choices among more complicated acts. I remain agnostic on what exactly this advice looks like. Beyond the basic requirement to respect stochastic dominance, different decision theorists will give different advice: some of them will advise you to maximize expected utility, others will advise you to maximize risk-weighted expected utility and so on. Since I allow incomparable options,

there is also the question of how to decide when options are incomparable. But any decision theorist needs to know at least your credences to give useful advice.³² My representation theorem shows how we can measure or define your credences without already presupposing that your preferences are fully rational and so enables the decision theorist to give useful advice.

Here is a simple toy example for how we can give useful advice. There is an urn with some red marbles, some yellow marbles and some black marbles. A marble will be drawn from this urn.³³ We observe that Ann strictly prefers winning one dollar if the marble is red over winning one dollar if the marble is black. So Ann prefers $\{\$1, R; \$0, R^C\}$ over $\{\$1, B; \$0, B^C\}$, where R is the event that the marble is red and B the event that the marble is black. We know that Ann likes more money rather than less, so Ann thinks R is more likely than B . Using **Alternative Prize**, we can infer that Ann must judge $R \cup Y$ to be more likely than $B \cup Y$, where Y is the event that the marble is yellow.

Now suppose Ann faces another choice. The first option pays one dollar if the marble is red, two dollars if the marble is yellow and nothing otherwise: $\{\$1, R; \$2, Y; 0$, $B\}$. The second option pays one dollar if the marble is black, two dollars if the marble is yellow and nothing otherwise: $\{\$1, B; \$2, Y; \$0, R\}$. We can advise Ann that, to avoid (strict) stochastic dominance, she should prefer the first option. This is genuinely useful advice because it is consistent with my axioms that Ann has no preference among these options or even prefers the stochastically dominated option.$

6.3 No Dependence on Utility

My axioms make minimal demands on the richness of the outcome space. I only require that there are at least two outcomes our agent is not indifferent between. Thus, we can define or measure subjective probabilities of

³²For decision-theoretic advice to be useful, we also need some way to measure your utility function (Narens and Skyrms 2020) and possibly other functions like your risk function (Neth 2019b).

³³To satisfy **Event Richness**, let us assume that this is the first draw in an infinite sequence of draws from the urn.

a very ‘simple-minded’ agent who only distinguishes between the outcome GOOD and the outcome BAD and has an ordinal ranking of these two outcomes. We can disentangle measuring subjective probability from the strong assumptions about value implicit in standard representation theorems.

There is a subtle difference in how my structure axioms compare to Ramsey and Savage. While I do not assume a rich space of outcomes, I do assume a rich space of events, as required by **Event Richness**. Ramsey’s original method does not need such a rich space of events.³⁴ So in terms of structural richness, my method does better than Ramsey’s in one way but worse in another way. This means that our axioms are incomparable in terms of their logical strength. However, from a philosophical point of view, I think that Ramsey and my approach assume a similar amount of structural richness. Ramsey assumes that our agent makes very fine-grained distinctions with respect to the value of outcomes while I assume that our agent makes very fine-grained distinctions with respect to the comparative probability of events. In contrast, my rationality axioms are much weaker than Ramsey’s rationality assumptions. Compared with Savage, I make the same structural assumptions about event richness but much weaker rationality assumptions, so we have a strictly more general decision-theoretic foundation for subjective probability than Savage’s.

The upshot: I have shown how to define or measure subjective probability with much weaker rationality axioms than standard representation theorems. If we are interested in ascribing precise subjective probability to partly irrational agents, this is definite progress. One might also take this result as illustration of how strong **Event Richness** really is. This axiom is doing the heavy lifting in my construction of subjective probability. On the one hand, this might incline some of us to be skeptical of this structure axiom.³⁵ On the other hand, nobody has figured out how to derive subjective probability without rich preferences and it is probably impossible to do so. So it is fair

³⁴When reconstructing Ramsey’s reasoning, Fishburn (1981, p. 151) writes that Ramsey assumes ‘a finite state set’. As noted above, no finite Ω can satisfy **Event Richness**.

³⁵Joyce (1999, p. 98) expresses skepticism about the structure axioms in Savage’s representation theorem, although one of Joyce’s main targets of completeness which I don’t assume.

to say that we have found *better foundations for subjective probability*.

7 Interpretation

I sketch two ways in which my representation theorem sheds light on the interpretation of subjective probability. First, it naturally fits with a view on which comparative probability is more fundamental than numerical probability. Second, it suggests a subjectivist version of the classical interpretation of probability.

7.1 Comparativism

Our starting point were comparative probability judgments defined in terms of preferences. I laid down axioms to ensure that this comparative probability ordering is a qualitative probability and an additional structure axiom to ensure that there is a unique probability function which represents this ordering. While we ultimately end up with a unique probability function which represents our agent's beliefs, this approach naturally suggests a picture on which comparative probability is *more fundamental* than numerical probability.³⁶ This is in sharp contrast to Ramsey's approach. For Ramsey, subjective probabilities are ratios of utilities and so they are fundamentally quantitative.

The idea that comparative probability is more fundamental than numerical probability has considerable intuitive appeal. It is more natural to think about which of two events is more likely than to assign numerical probabilities. Furthermore, as I will turn to explain now, taking comparative

³⁶Comparativism is discussed by Koopman (1940), Fine (1973), Zynda (2000), Hawthorne (2017), Stefánsson (2017), Konek (2019), and Elliott (2022). Of course, the idea of starting with comparative probability is well-known in decision theory (Fishburn 1986). However, it is valuable to make explicit that we can be realists about comparative probability and constructivists about numerical probability, while many decision theorists like Savage are constructivist all the way down. As Holliday and Icard (2013) point out, comparative probability can also shed light on probability operators in natural language and so might help us with puzzles about which inferences with these operators are valid (Yalcin 2010; Neth 2019a).

probability as fundamental allows us to tell a plausible story about the axioms.

My representation theorem naturally fits with a combination of realism and constructivism: realism about comparative probability and constructivism about numerical probability. According to this picture, the comparative probability ordering is psychologically real and not reducible to preferences—rather, preferences serve to measure comparative probability. This is the realist aspect. The advantage of this bit of realism is that we can tell a plausible story about some axioms, in particular **Non-Degeneracy**, which requires our agent not to be indifferent among all outcomes. It is not very plausible to think that the *existence* of comparative probability requires this axiom, but much more plausible to think that *measuring* comparative probability requires this axiom.

However, in contrast to the comparative probability ordering, the probability function constructed in the representation theorem is not psychologically real but only a ‘representational device’ to talk about the underlying comparative probability ordering. This is the constructivist aspect. The advantage of this bit of constructivism is that we can tell a plausible story about **Event Richness**. It is implausible to think that subjective probability requires the rich event space postulated by this axiom. If we think of comparative probability as fundamental, we can say that agents might have comparative subjective probabilities even if they do not satisfy this axiom. The axiom describes a condition under which we can *represent* comparative subjective probability by a unique probability function, not a condition under which subjective probability *exists*.

7.2 Vindicating the Classical Picture

According to the *classical interpretation of probability* associated with Laplace, we can determine the probability of some event as follows.³⁷ First, we find a suitable set of ‘equally possible’ cases. Then, we count the number of cases

³⁷Gillies (2000, ch. 2) gives an overview of the classical interpretation and Diaconis and Skyrms (2018, ch. 1) briefly recount the history of reasoning about probability in terms of ‘equally possible’ cases.

in which the event occurs and divide this number by the number of all cases. For example, if we want to find out the probability of snake eyes (two 1's) when rolling two fair dice, there are 36 'equally possible' cases and exactly one of these cases is snake eyes, so the probability of snake eyes is $\frac{1}{36}$.

There are many well-known objections to the classical interpretation of probability. First, you might complain that the definition given above is circular. Laplace defines probability in terms of 'equally possible' cases, but it is hard to see what 'equally possible' could mean other than 'equally probable'. Second, the classical interpretation entails that all probabilities are rational, since they are the ratio of two positive integers. But there is nothing incoherent about irrational-valued probabilities.³⁸ Third, what guarantees that we can always find 'equally possible' cases? They are easy to find in games of chance but much harder to find in real-life situations, where we might try to find the probability that a nuclear power plant will have a catastrophic accident in the next 100 years (Halpern 2003, p. 18).

My representation theorem can be construed as *subjectivist version of the classical interpretation of probability*. Recall how we construct the subjective probability function. To find Ann's subjective probability for rain tomorrow, we find n mutually exclusive and collectively exhaustive events which Ann considers to be equally probable. We write down k , the greatest number of events Ann considers to be collectively less likely than rain tomorrow. This is a bit like counting the number of 'cases' in which it rains tomorrow. We approximate Ann's subjective probability of rain tomorrow by k divided by the total number of cases. We define Ann's subjective probability of rain tomorrow as the limit of this procedure as n goes to infinity.

The shift towards the subjective helps with some of the well-known worries for the classical interpretation. First, what does it mean to say that the cases are 'equally possible'? In our picture, it means that they are judged to be equally probable by our subject and we can know that they are so judged by looking at our subject's preferences. Since we have not defined comparative probability in terms of numerical probability but rather directly in terms of preferences, we can sidestep the circularity worry. Second, since we define

³⁸Hájek (1996) uses this as an argument against (finite) frequentism.

subjective probability as limit of a sequence of rational numbers, we can have irrational-valued subjective probabilities.

Some worries for our subjectivist Laplacean picture still remain. What guarantees that, for any n , we can find a partition of n mutually exclusive and collectively exhaustive events which our subject considers to be equally likely? In our construction, this follows from **Event Richness**, which ensures that the event space of our subject is sufficiently fine-grained. But is it a requirement for the existence of subjective probability to have such a fine-grained event space? Arguably not. If we accept comparativism, we can reply that the more fundamental comparative subjective probabilities still exist without **Event Richness**, but they may not admit of representation by a unique probability function. From this point of view, Laplace’s ‘equiprobable cases’ highlight a condition under which comparative probability judgments can be represented by a unique probability function.

8 Conclusion

Ramsey wants to reduce subjective probability to preference but makes very demanding rationality assumptions—the agent under consideration has perfectly coherent preferences. I have shown how to provide better foundations for subjective probability: axioms which ensure that there is a unique probability function representing our agent’s beliefs while leaving room for mistakes.

Let me close by observing that if we are convinced that my axioms are rationally required, we can also read my representation theorem as an answer to the question: *Why be probabilistically coherent?* Because rationality requires you to obey the axioms and if you obey the axioms, there is a unique probability function representing your comparative probability judgments. In contrast to other decision-theoretic arguments for probabilistic coherence, such as dutch book arguments or standard representation theorems, this argument does not presuppose or entail expected utility maximization. So we have a new argument for probabilistic coherence from weak assumptions

about practical rationality in the face of uncertainty.³⁹

Acknowledgements

I'm very grateful to Kenny Easwaran, Wesley Holliday, Marcel Jahn, Mikayla Kelley, John MacFarlane, Edward Schwartz and two anonymous referees for helpful comments on previous drafts. Special thanks to Lara Buchak for several rounds of comments and her enthusiasm and support for this project. Further thanks to audiences at the Formal Epistemology Workshop 2022 at UC Irvine and Berkeley's Formal Epistemology Reading Course (FERC) for helpful discussion, especially Mathias Böhm, Yifeng Ding, and Hanti Lin. While working on this paper, I was supported by a Global Priorities Fellowship by the Forethought Foundation and a Josephine De Karman Fellowship.

References

- Allais, M. (1953). "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine". In: *Econometrica* 21.4, pp. 503–546. DOI: 10.2307/1907921.
- Aumann, Robert J. (1962). "Utility Theory Without the Completeness Axiom". In: *Econometrica* 30.3, pp. 445–462. DOI: <https://doi.org/10.2307/1909888>.
- Bader, Ralf M. (2018). "Stochastic Dominance and Opaque Sweetening". In: *Australasian Journal of Philosophy* 96.3, pp. 498–507. DOI: 10.1080/00048402.2017.1362566.
- Bales, Adam, Daniel Cohen, and Toby Handfield (2014). "Decision Theory for Agents with Incomplete Preferences". In: *Australasian Journal of Philosophy* 92.3, pp. 453–70. DOI: 10.1080/00048402.2013.843576.

³⁹In contrast to accuracy arguments for probabilistic coherence, we also avoid commitments about epistemic value. Furthermore, accuracy arguments run into difficulties when there are infinitely many possibilities (Kelley and Neth 2023).

- Beck, Lukas and Marcel Jahn (2021). “Normative Models and Their Success”. In: *Philosophy of the Social Sciences* 51.2, pp. 123–150. DOI: 10.1177/0048393120970908.
- Birnbaum, Michael H. and Juan B. Navarrete (1998). “Testing Descriptive Utility Theories: Violations of Stochastic Dominance and Cumulative Independence”. In: *Journal of Risk and Uncertainty* 17.1, pp. 49–79. DOI: 10.1023/A:1007739200913.
- Bottomley, Christopher and Timothy Luke Williamson (forthcoming). “Rational Risk-Aversion: Good Things Come to Those Who Weight”. In: *Philosophy and Phenomenological Research*. DOI: 10.1111/phpr.13006.
- Bradley, Richard (2004). “Ramsey’s Representation Theorem”. In: *Dialectica* 58.4, pp. 483–497. DOI: 10.1111/j.1746-8361.2004.tb00320.x.
- Buchak, Lara (2013). *Risk and Rationality*. Oxford University Press.
- (2017). “Decision Theory”. In: *The Oxford Handbook of Probability and Philosophy*. Ed. by Christopher Hitchcock and Alan Hájek. Oxford University Press, pp. 789–814.
- Chang, Ruth (2002). “The Possibility of Parity”. In: *Ethics* 112.4, pp. 659–688. DOI: 10.1086/339673.
- Davidson, Donald (1973). “Radical Interpretation”. In: *Dialectica* 27.1, pp. 314–328. DOI: 10.1111/j.1746-8361.1973.tb00623.x.
- de Finetti, Bruno (1931). “Sul significato soggettivo della probabilità”. In: *Fundamenta Mathematicae* 17.1, pp. 298–329. DOI: 10.4064/fm-17-1-298-329.
- Diaconis, Persi and Brian Skyrms (2018). *Ten Great Ideas About Chance*. Princeton University Press.
- Ding, Yifeng, Wesley H. Holliday, and Thomas Icard (2021). “Logics of Imprecise Comparative Probability”. In: *International Journal of Approximate Reasoning* 132, pp. 154–180. DOI: <https://doi.org/10.1016/j.ijar.2021.02.004>.
- Dreier, James (1996). “Rational Preference: Decision Theory as a Theory of Practical Rationality”. In: *Theory and Decision* 40.3, pp. 249–276. DOI: 10.1007/bf00134210.

- Earman, John (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. MIT Press.
- Easwaran, Kenny (2014). “Decision Theory Without Representation Theorems”. In: *Philosophers’ Imprint* 14. URL: <http://hdl.handle.net/2027/spo.3521354.0014.027>.
- Elga, Adam (2016). “Bayesian Humility”. In: *Philosophy of Science* 83.3, pp. 305–323. DOI: 10.1086/685740.
- Elliott, Edward (2017). “A Representation Theorem for Frequently Irrational Agents”. In: *Journal of Philosophical Logic* 46.5, pp. 467–506. DOI: 10.1007/s10992-016-9408-8.
- (2022). “Comparativism and the Measurement of Partial Belief”. In: *Erkenntnis* 87.6, pp. 2843–2870. DOI: 10.1007/s10670-020-00329-x.
- Ellsberg, Daniel (1961). “Risk, Ambiguity, and the Savage Axioms”. In: *The Quarterly Journal of Economics* 75.4, pp. 643–669. DOI: 10.2307/1884324.
- Eriksson, Lina and Alan Hájek (2007). “What Are Degrees of Belief”. In: *Studia Logica* 86.2, pp. 185–213. DOI: 10.1007/s11225-007-9059-4.
- Fine, Terrence L. (1973). *Theories of Probability: An Examination of Foundations*. Academic Press.
- Fishburn, Peter C. (1970). *Utility Theory for Decision Making*. Publications in Operations Research. John Wiley & Sons/Research Analysis Corporation.
- (1981). “Subjective Expected Utility: A Review of Normative Theories”. In: *Theory and Decision* 13.2, pp. 139–199. DOI: 10.1007/bf00134215.
- (1986). “The Axioms of Subjective Probability”. In: *Statistical Science* 1.3, pp. 335–345. DOI: 10.1214/ss/1177013611.
- Gaifman, Haim and Yang Liu (2018). “A Simpler and More Realistic Subjective Decision Theory”. In: *Synthese* 195.10, pp. 4205–4241. DOI: 10.1007/s11229-017-1594-6.
- Gillies, Donald (2000). *Philosophical Theories of Probability*. Routledge.
- Hájek, Alan (1996). “‘Mises Redux’—Redux: Fifteen Arguments Against Finite Frequentism”. In: *Erkenntnis* 45.2-3, pp. 209–27. DOI: 10.1007/bf00276791.

- Halpern, Joseph Y. (2003). *Reasoning About Uncertainty*. MIT Press.
- Hare, Caspar (2010). “Take the Sugar”. In: *Analysis* 70.2, pp. 237–247. DOI: 10.1093/analys/anp174.
- Hawthorne, James (2017). “A Logic of Comparative Support: Qualitative Conditional Probability Relations Representable by Popper Functions”. In: *The Oxford Handbook of Probability and Philosophy*. Ed. by Alan Hájek and Christopher Hitchcock, pp. 277–295.
- Holliday, Wesley H. and Thomas Icard (2013). “Measure Semantics and Qualitative Semantics for Epistemic Modals”. In: *Proceedings of SALT 23*, pp. 514–534. DOI: 10.3765/salt.v23i0.2670.
- Icard, Thomas (2016). “Pragmatic Considerations on Comparative Probability”. In: *Philosophy of Science* 83.3, pp. 348–370. DOI: 10.1086/685742.
- (2021). “Why Be Random?” In: *Mind* 130.517, pp. 111–139. DOI: 10.1093/mind/fzz065.
- Jeffrey, Richard C. (1990). *The Logic of Decision*. University of Chicago Press.
- Joyce, James M. (1999). *The Foundations of Causal Decision Theory*. Cambridge University Press.
- Kelley, Mikayla and Sven Neth (2023). “Accuracy and Infinity: A Dilemma for Subjective Bayesians”. In: *Synthese* 201.12, pp. 1–14. DOI: 10.1007/s11229-022-04019-9.
- Konek, Jason (2019). “Comparative Probabilities”. In: *The Open Handbook of Formal Epistemology*. Ed. by Richard Pettigrew and Jonathan Weisberg. PhilPapers Foundation, pp. 267–348.
- Koopman, Bernard O. (1940). “The Axioms and Algebra of Intuitive Probability”. In: *Annals of Mathematics* 41.2, pp. 269–292. DOI: 10.2307/1969003.
- Kraft, Charles H., John W. Pratt, and A. Seidenberg (1959). “Intuitive Probability on Finite Sets”. In: *Annals of Mathematical Statistics* 30.2, pp. 408–419. DOI: 10.1214/aoms/1177706260.
- Krantz, David, Duncan Luce, Patrick Suppes, and Amos Tversky (1971). *Foundations of Measurement, Vol. I: Additive and Polynomial Representations*. New York Academic Press.

- Kreps, David (1988). *Notes on the Theory of Choice*. Underground Classics in Economics. Westview Press.
- Lewis, David (1974). “Radical Interpretation”. In: *Synthese* 23, pp. 331–344. DOI: 10.1007/bf00484599.
- Luce, Duncan (1967). “Sufficient Conditions for the Existence of a Finitely Additive Probability Measure”. In: *The Annals of Mathematical Statistics* 38.3, pp. 780–786. DOI: 10.1214/aoms/1177698871.
- Machina, Mark J. and David Schmeidler (1992). “A More Robust Definition of Subjective Probability”. In: *Econometrica* 60.4, pp. 745–780. DOI: 10.2307/2951565.
- Meacham, Christopher J. G. and Jonathan Weisberg (2011). “Representation Theorems and the Foundations of Decision Theory”. In: *Australasian Journal of Philosophy* 89.4, pp. 641–663. DOI: 10.1080/00048402.2010.510529.
- Misak, Cheryl (2020). *Frank Ramsey: A Sheer Excess of Powers*. Oxford University Press.
- Narens, Louis and Brian Skyrms (2020). *The Pursuit of Happiness: Philosophical and Psychological Foundations of Utility*. Oxford University Press.
- Neth, Sven (2019a). “Chancy Modus Ponens”. In: *Analysis* 79.4, pp. 632–638. DOI: 10.1093/analys/anz022.
- (2019b). “Measuring Belief and Risk Attitude”. In: *Proceedings Seventeenth Conference on Theoretical Aspects of Rationality and Knowledge*. Toulouse, France, 17-19 July 2019. Ed. by Lawrence S. Moss. Vol. 297. Electronic Proceedings in Theoretical Computer Science, pp. 252–272. DOI: 10.4204/EPTCS.297.22.
- (2023). “A Dilemma for Solomonoff Prediction”. In: *Philosophy of Science* 90.2, pp. 288–306. DOI: 10.1017/psa.2022.72.
- Okasha, Samir (2016). “On the Interpretation of Decision Theory”. In: *Economics & Philosophy* 32.3, pp. 409–433. DOI: 10.1017/s0266267115000346.
- Oliver, Adam (2003). “A Quantitative and Qualitative Test of the Allais Paradox Using Health Outcomes”. In: *Journal of Economic Psychology* 24.1, pp. 35–48. DOI: 10.1016/s0167-4870(02)00153-8.

- Ramsey, Frank (1926). “Truth and Probability”. In: *The Foundations of Mathematics and other Logical Essays*. Ed. by R.B. Braithwaite. Originally published in 1931. 1999 electronic edition. Harcourt, pp. 156–198.
- Resnik, Michael (1987). *Choices: An Introduction to Decision Theory*. University of Minnesota Press.
- Russell, Jeffrey Sanford (forthcoming). “Fixing Stochastic Dominance”. In: *British Journal for the Philosophy of Science*.
- Savage, Leonard J. (1972). *The Foundations of Statistics*. Second Revised Edition. Wiley Publications in Statistics.
- Schoenfield, Miriam (2014). “Decision Making in the Face of Parity”. In: *Philosophical Perspectives* 28.1, pp. 263–277. DOI: 10.1111/phpe.12044.
- Scott, Dana (1964). “Measurement Structures and Linear Inequalities”. In: *Journal of Mathematical Psychology* 1.2, pp. 233–247. DOI: 10.1016/0022-2496(64)90002-1.
- Sen, Amartya (2018). “The Importance of Incompleteness”. In: *International Journal of Economic Theory* 14.1, pp. 9–20. DOI: 10.1111/ijet.12145.
- Spohn, Wolfgang (2012). *The Laws of Belief: Ranking Theory and its Philosophical Applications*. Oxford University Press.
- Stefánsson, H. Orri (2017). “What is ‘Real’ in Probabilism?” In: *Australasian Journal of Philosophy* 95.3, pp. 573–587. DOI: 10.1080/00048402.2016.1224906.
- Tarsney, Christian (2020). “Exceeding Expectations: Stochastic Dominance as a General Decision Theory”. In: *Global Priorities Institute Working Papers Series* 3-2020.
- Titelbaum, Michael G. (2022). *Fundamentals of Bayesian Epistemology 2: Arguments, Challenges, Alternatives*. Oxford University Press.
- Villegas, Carlos (1964). “On Qualitative Probability σ -Algebras”. In: *The Annals of Mathematical Statistics* 35.4, pp. 1787–1796. DOI: 10.1214/aoms/1177700400.
- Yalcin, Seth (2010). “Probability Operators”. In: *Philosophy Compass* 5.11, pp. 916–37. DOI: 10.1111/j.1747-9991.2010.00360.x.

Zynda, Lyle (2000). “Representation Theorems and Realism About Degrees of Belief”. In: *Philosophy of Science* 67.1, pp. 45–69. DOI: 10.1086/392761.

Appendix

Theorem 1. *The preference relation \succsim satisfies **Outcome Independence**, **Non-Degeneracy**, **Restricted Ordering**, **Certain Prize** and **Alternative Prize** if and only if the comparative probability ordering \succcurlyeq is a qualitative probability.*

Proof. I begin by showing the left-to-right direction. Assume \succsim satisfies the axioms. By **Outcome Independence** and Definition 1, \succcurlyeq is a binary relation on \mathcal{F} . By **Non-Degeneracy**, there are $b, w \in \mathcal{O}$ with $b \succ w$. By **Restricted Ordering**, for any $X, Y \in \mathcal{F}$, we have $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ or $\{b, Y; w, Y^C\} \succsim \{b, X; w, X^C\}$, so by Definition 1, $X \succcurlyeq Y$ or $Y \succcurlyeq X$. Therefore, \succcurlyeq is complete. Analogous reasoning shows that \succcurlyeq is transitive, so \succcurlyeq satisfies Ordering.

Consider any $X \in \mathcal{F}$. By **Certain Prize**, we have $\underline{b} \succsim \{b, X; w, X^C\}$ and $\{b, X; w, X^C\} \succsim \underline{w}$. Now $\underline{b} = \{b, \Omega; w, \emptyset\}$ and $\underline{w} = \{w, \Omega; b, \emptyset\}$. So $\{b, \Omega; w, \emptyset\} \succsim \{b, X; w, X^C\}$ and $\{b, X; w, X^C\} \succsim \{w, \Omega; b, \emptyset\}$, and by Definition 1, $\Omega \succcurlyeq X \succcurlyeq \emptyset$, so \succcurlyeq satisfies Boundedness. By analogous reasoning, \succcurlyeq satisfies Non-Triviality.

Now assume $X \cap Z = Y \cap Z = \emptyset$. We want to show that $X \succ Y \iff X \cup Z \succ Y \cup Z$. Assume $X \succ Y$. By Definition 1, $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ for some $b, w \in \mathcal{O}$ with $b \succ w$. By **Alternative Prize**, $\{b, X \cup Z; w, (X \cup Z)^C\} \succ \{b, Y \cup Z; w, (Y \cup Z)^C\}$, so $X \cup Z \succ Y \cup Z$ by Definition 1. Analogous reasoning shows the converse implication, so \succcurlyeq satisfies Qualitative Additivity.

I proceed to show the right-to-left direction. Assume \succcurlyeq is a qualitative probability. We want to show that \succsim satisfies the axioms. Assume $\{b, X; w, X^C\} \succsim \{b, Y; w, Y^C\}$ for some $b, w \in \mathcal{O}$ with $b \succ w$. By Definition 1, $X \succcurlyeq Y$. Now assume for *reductio* that for some $b', w' \in \mathcal{O}$ with $b' \succ w'$,

$\{b', X; w', X^C\} \not\sim \{b', Y; w', Y^C\}$. By Definition 1, $X \not\asymp Y$, which contradicts our assumption. Therefore, $\{b, X; w, X^C\} \sim \{b, Y; w, Y^C\}$ for all $b, w \in \mathcal{O}$ such that $b \succ w$, so **Outcome Independence** holds. We have $\Omega \succ \emptyset$ by Non-Triviality, so $\{b, \Omega; w, \emptyset\} \succ \{b, \emptyset; w, \Omega\}$ for some $b, w \in \mathcal{O}$ with $b \succ w$. Therefore, there are some $b, w \in \mathcal{O}$ with $b \succ w$, so **Non-Degeneracy** holds.

By Ordering, for all $X, Y \in \mathcal{F}$, $X \succcurlyeq Y$ or $Y \succcurlyeq X$. Consider some $b, w \in \mathcal{O}$ with $b \succ w$. We want to show that for all $X, Y \in \mathcal{F}$, either $\{b, X; w, X^C\} \sim \{b, Y; w, Y^C\}$ or $\{b, Y; w, Y^C\} \sim \{b, X; w, X^C\}$. Assume $X \succcurlyeq Y$. By Definition 1, $\{b', X; w', X^C\} \sim \{b', Y; w', Y^C\}$ for some $b', w' \in \mathcal{O}$ with $b' \succ w'$. So by **Outcome Independence**, $\{b, X; w, X^C\} \sim \{b, Y; w, Y^C\}$. An analogous argument applies if $Y \succcurlyeq X$. For transitivity, assume that for some $b, w \in \mathcal{O}$ with $b \succ w$, we have $\{b, X; w, X^C\} \sim \{b, Y; w, Y^C\}$ and $\{b, Y; w, Y^C\} \sim \{b, Z; w, Z^C\}$. By Definition 1, $X \succcurlyeq Y$ and $Y \succcurlyeq Z$, so by Ordering it follows that $X \succcurlyeq Z$. Again by Definition 1, $\{b', X; w', X^C\} \sim \{b', Y; w', Y^C\}$ for some $b', w' \in \mathcal{O}$ with $b' \succ w'$. By **Outcome Independence**, $\{b, X; w, X^C\} \sim \{b, Y; w, Y^C\}$, so **Restricted Ordering** holds. By Boundedness, for all $X \in \mathcal{F}$, $\Omega \succcurlyeq X$ and $X \succcurlyeq \emptyset$, so $\underline{b} \sim \{b, X; w, X^C\}$ and $\{b, X; w, X^C\} \sim \underline{w}$ for all $b, w \in \mathcal{O}$ with $b \succ w$ so **Certain Prize** holds.

Finally, let $X \cap Z = Y \cap Z = \emptyset$ and assume $\{b, X; w, X^C\} \succ \{b, Y; w, Y^C\}$ for some $b, w \in \mathcal{O}$ with $b \succ w$. By Definition 1, $X \succ Y$, so by Qualitative Additivity, $X \cup Z \succ Y \cup Z$. Again by Definition 1 and **Outcome Independence**, $\{b, X \cup Z; w, (X \cup Z)^C\} \succ \{b, Y \cup Z; w, (Y \cup Z)^C\}$. A similar argument shows the converse entailment, so **Alternative Prize** holds. \square

We can ensure countable additivity by adding this axiom:

Monotone Preference Continuity. For any $b, w \in \mathcal{O}$ with $b \succ w$ and any monotonically increasing sequence of events $X_1 \subseteq X_2 \subseteq \dots$ with $\bigcup_{n=1}^{\infty} X_n = X$, if for all n , $\{b, Y; w, Y^C\} \sim \{b, X_n; w, X_n^C\}$, then $\{b, Y; w, Y^C\} \sim \{b, X; w, X^C\}$.

Building on work by Villegas (1964), we can show:

Theorem 3. *If the preference relation \sim satisfies **Outcome Independence**, **Non-Degeneracy**, **Restricted Ordering**, **Certain Prize**, **Al-***

ternative Prize, Event Richness and Monotone Preference Continuity, there is a unique countably additive probability function representing \succsim .

Proof. Assume the preference relation \succsim satisfies **Outcome Independence**, **Non-Degeneracy**, **Restricted Ordering**, **Certain Prize**, **Alternative Prize** and **Event Richness**. Then \succsim is *atomless*, which means that for every $X \succ \emptyset$, there is some $Y \subseteq X$ such that $X \succ Y \succ \emptyset$. Now, given **Monotone Preference Continuity**, \succsim satisfies:

Monotone Probability Continuity. If X_1, X_2, \dots is a monotonically increasing sequence of events with $\bigcup_{n=1}^{\infty} X_n = X$, and for every n , $Y \succ X_n$, then $Y \succ X$.

Villegas (1964) shows that if \succsim is an atomless qualitative probability and **Monotone Probability Continuity** holds, there is a unique countably additive probability function representing \succsim . By Theorem 2, there is a unique probability function representing \succsim . By Villegas' result, **Monotone Preference Continuity** implies that this probability function must be countably additive. \square