The Qualitative Grounds of Quantitative Confidence

Edward Elliott

University of Notre Dame

Abstract

This paper explores the qualitative basis for deriving meaningful psychological information from probabilistic representations of confidence within decision theory. Conditions on two finite binary relations—a preference relation $\succsim$ and a comparative confidence relation $\succsim_c$—are shown to imply the existence of a probability function and a utility function that jointly represent $\succsim$ under the expected utility rule; furthermore, the probability function is unique and also represents $\succsim_c$. It is shown that, under certain conditions, these probabilities will contain strictly more *meaningful* information than is contained in the confidence relation (and natural extensions thereof). In other words, there's more to what's meaningful in decision-theoretic representations of confidence than can generally be explained by reference to confidence relations alone.

1. Introduction

There are two main ways of thinking about the relationship between preferences and probabilities in subjective expected utility theory (SEUT). On the one hand, *constructivists* take preferences to be primary—conceptually, ontologically, or both—while probabilities are a 'mathematical construction' introduced to represent structural features of a preference ordering, without necessarily corresponding to any independent psychological quantity. On the other hand, *realists* interpret probabilities as a measure of *confidence*, where this is assumed to be neither conceptually nor ontologically reducible to preferences, but instead a major factor in what causally explains and normatively constrains the shape of an agent's preference ordering.[1]

This paper presupposes a realist perspective. But if probabilities measure an independent psychological quantity, then questions arise as to what structural features of that quantity are meaningfully captured by the numerical representation. Start with

[1] There are two main conceptions of preference—a *behavioral* conception according to which preferences are choice dispositions, and a *mentalistic* conception according to which preference is a comparative desire-like attitude. While I prefer the latter, nothing much in the present discussion hangs on this distinction and so it won't be discussed further. For discussion on the different conceptions of preference, see (Hausman 2012) and (Dietrich & List 2016).

the what-question: *what is meaningful in standard probabilistic representations of confidence?* Suppose, for example, that an agent's degrees of confidence are represented by a probability function p that assigns events E_1 and E_2 probabilities 0.9 and 0.3 respectively. It is uncontroversial that p carries meaningful ordinal information—that is, p represents an agent with more confidence in E_1 than in E_2. But do the probabilities represent anything more? In particular, is the fact that 0.9 is three times 0.3 meaningful, or just an artifact of a conventional choice of scale? Suppose it is. Then there's the why-question: *how is that meaning explained?* The mere fact that $0.9 = 3 \times 0.3$ doesn't by itself entail that the $3\times$ ratio is meaningful with respect to the measurement of confidence, any more than it entails that 90° F ($\approx 32^\circ$ C) is three times hotter than 30° F ($\approx -1^\circ$ C). Rather, if p carries meaningful ratio information, then that meaning must ultimately derive from features of the agent's psychology. Thus the why-question is really about the kinds of psychological structures implicated by SEUT—if ratios in probabilistic representations of confidence are indeed meaningful, then what is it about *confidence* itself that makes them so?

This paper presents a decision-theoretic representation theorem with the goal of shedding light on these questions. On the basis of this theorem, I will argue that in the context of SEUT under a realist construal,

Preference-Grounded Ratios. There are circumstances in which (i) two measures of confidence are equivalent in meaning if and only if they agree on all ratio information, and moreover (ii) the explanation for this fact relates at least in part to the interaction of confidence and preference.

The first part, in other words, is that degrees of confidence that can be measured on a ratio scale, at least under certain conditions. Though it's not often discussed, this seems to be widely accepted among realists.[2] More controversial is the second part, especially as it stands in contrast to well-known explanatory strategies that appeal instead to the structure of the agent's *comparative confidence* ordering—i.e., a relation capturing purely ordinal confidence comparisons, such as "… has more confidence in E than in F". I will discuss how these alternative explanations work in due course; what's important is that they require no direct reference to preferences, and therefore cast doubt on Preference-Grounded Ratios.[3]

My argument requires a representation theorem with certain distinctive properties, setting it apart from more typical representation theorems. For one thing, the typical theorem establishes conditions under which an agent's preferences maximize

[2] But not universally! Cf. Joyce (2015: 416–9), who argues that probabilities in SEUT represent confidence on an interval scale—so ratios of differences are meaningful, but not ratios *simpliciter*.

[3] The idea that we can explain confidence ratios in terms of comparative confidence relations has an extensive history; see (de Finetti 1931), (Koopman 1940), (Savage 1954), (Villegas 1964), (Luce 1968), (Domotor 1969), (Krantz et al. 1971: ch. 5), (Suppes and Zanotti 1982), (Fine 1973: ch. 2), (Suppes 1994); and for more recent discussions, see (DiBella 2018), (Stefánsson 2018), (Konek 2019), and (Elliott 2024). Such relations have been given a variety of overlapping names across several literatures, including but not limited to *comparative probabilities*, *qualitative probabilities*, *qualitative confidence*, *comparative beliefs*, and *belief orderings*.

expected utility relative to a unique probability function p; however, such theorems do not guarantee that p measures the agent's confidence—let alone that it does so on a ratio scale. As such, the goal is to construct a representation theorem that implies, under a realist interpretation, that p represents the agent's confidence in the minimal sense of reflecting their comparative confidence, and furthermore that it does so on a ratio scale. Call these the *representation challenge* and the *ratio challenge* respectively; as we'll see, they present only minor difficulties. Harder is the *differentiation challenge*: in order to show that the proper answer to the why-question relates to the interaction of confidence and preference, we need to isolate special conditions under which a preference-based explanation *can* apply whereas the aforementioned comparative confidence explanations *cannot*. The difficulty here is that existing theorems such as Savage's (1954) and those like it derive probabilities from an induced comparative confidence ordering and secure their uniqueness relative to that ordering, and consequently presume preference conditions that preclude precisely the kind of differentiation required to support PREFERENCE-GROUNDED RATIOS. As such, we want a theorem that establishes meaningful confidence ratios *specifically* in circumstances where comparative confidence explanations are, in principle, unavailable.

The paper is structured as follows. Section 2 lays out the formal setting and some conventions to be employed throughout the discussion, and Section 3 then presents the representation theorem in full. Then Section 4, Section 5 and Section 6 discuss the representation, ratio and differentiation challenges in turn. Finally, Section 7 draws out some of the broader philosophical implications of the result.

2. Preliminaries

We adopt Savage's (1954) analytic framework, though restricted to the finite case. To start, there are two primitive sets ($\mathcal{S}$ and $\mathcal{C}$) and two derived sets ($2^{\mathcal{S}}$ and $\mathcal{C}^{\mathcal{S}}$):

- $\mathcal{S}$ is a non-empty finite set of *states*, with typical element s
- $\mathcal{C}$ is a non-empty finite set of *consequences*, with typical elements $\alpha, \beta, \gamma, \dots$
- $2^{\mathcal{S}}$ is the set of *events* (i.e., sets of states), with typical elements E, F
- $\mathcal{C}^{\mathcal{S}}$ is the set of *acts* (i.e., functions from states to consequences), with typical elements f, g, h

For any event E, the complement of E with respect to $\mathcal{S}$ is denoted $\overline{E}$; and s will be used ambiguously for both the state and the singleton event containing that state. Next, we use fEg for the mixture of acts f and g on event E; that is, fEg denotes the act h such that $h(s) = f(s)$ for all states $s \in E$ and $h(s) = g(s)$ otherwise. Furthermore, we let $f(E) = \alpha$ whenever $f(s) = \alpha$ for all $s \in E$; and we identify each consequence with its corresponding constant act, so $f = \alpha$ whenever $f(\mathcal{S}) = \alpha$. Given the above, note in particular that $\alpha E \beta$ designates the *binary act* f such that $f(E) = \alpha$ and $f(\overline{E}) = \beta$, and that $\alpha E \alpha$ is just another way of writing α. Finally, observe that every act f will induce at least one (and typically more than one) partition $\mathcal{E}_f$ of $\mathcal{S}$ such that f is constant with respect to each E in $\mathcal{E}_f$; we will call these *consequence partitions*.

In addition, there are two primitive relations:

- $\succsim$ is a binary *preference relation* on $\mathcal{C}^{\mathcal{S}}$
- $\succsim_c$ is a binary *confidence relation* on $2^{\mathcal{S}}$

$\sim$ and $\succ$ are defined using $\succsim$ in the usual way; likewise for $\sim_c$ and $\succ_c$. In saying that $\succsim$ and $\succsim_c$ are *primitives*, I mean that the formal framework imposes no definitional or structural connection between them beyond those explicitly introduced by the axioms of the theorem; any further conceptual or metaphysical linkage between their interpretations lies outside the scope of the theorem. Keep in mind also that $\succsim_c$ is a relation between individual events, with $E \succsim_c F$ intended to say that the agent has at least as much confidence in E as in F. In Section 6 we will also consider a richer *conditional* confidence relation that holds between ordered pairs of events, but for now we restrict attention to *unconditional* confidence relations.

Next, we refer to any real-valued function on $\mathcal{C}$ as a *utility function*, and any real-valued function on $2^{\mathcal{S}}$ as a *confidence function*. Given that, $\succsim$ *maximizes EU* relative to a confidence function c and a utility function u just in case

$$f \succsim g \Leftrightarrow \sum_{E \in \mathcal{E}_f} c(E) \cdot u\big(f(E)\big) \geq \sum_{E \in \mathcal{E}_g} c(E) \cdot u\big(g(E)\big) \tag{MEU}$$

for any choice of consequence partitions $\mathcal{E}_f$ and $\mathcal{E}_g$. Likewise, $\succsim$ maximizes EU relative to c whenever there exists some utility function u such that $\succsim$ maximizes EU relative to c and u.

Finally, note that the c in (MEU) need not be a *probability function*, defined strictly as any confidence function satisfying the usual Kolmogorov properties:

- $c(\mathcal{S}) = 1$ (*normalization*)
- $c(E) \geq 0$ (*non-negativity*)
- $E \cap F = \varnothing$ implies $c(E \cup F) = c(E) + c(F)$ (*additivity*)

For all that follows, p is reserved for confidence functions presumed to satisfy all three properties, whereas c will be used for confidence functions that may satisfy some, all, or none of these properties.

3. Representation theorem

Before we get to the theorem, a couple clarificatory points are warranted. First, my topic is the meaningfulness of confidence ratios within the context of SEUT, and the discussion won't much depend on whether that theory is better interpreted descriptively (e.g., describing the typical relationship between ordinary humans' preferences and their degrees of confidence and utility) or normatively (e.g., characterizing what that relationship ideally ought to be). I take no stance here on such matters, nor on the truth of SEUT under either interpretation. For ease of expression, though, I will adopt language fitting a descriptive interpretation throughout.

For the same reason, the reader is forewarned against interpreting the axioms in this section either as descriptive generalizations or as normative ideals; they need be neither, given the purpose for which the theorem has been constructed. Decision-theoretic representation theorems are often—though not always—intended to provide maximally general conditions under which the corresponding decision theory holds; and where that is the goal, then it makes good sense to evaluate the axioms for their descriptive accuracy or normative force depending on whether the theory is to be interpreted descriptively or normatively. However, in this paper we seek not a general axiomatization of SEUT, but to show how meaningful confidence ratios can arise within SEUT under conditions that enable us to differentiate between explanatory hypotheses that would otherwise be indistinguishable. As such, what's important is that the "axioms" are consistent with SEUT—that they characterize how an expected utility maximizer *could* in principle be, not how we humans *typically are* nor how we *rationally ought* to be.

With that out of the way, there are eight axioms in total. The first two are:

A1. $\succsim$ is transitive and connected (i.e., $f \succsim g$ or $g \succsim f$ for all f, g)
A2. $\alpha \succsim \beta$ implies $\alpha E \beta \succsim \alpha F \beta \Leftrightarrow E \succsim_c F$

A1 is obviously necessary for $\succsim$ to maximize EU relative to any confidence and utility functions. A2 says that preferences between binary acts with the same two consequences depend only on relative confidence with respect to the conditioning events. It serves as the main route by which confidence gets tied to preference, and is implied by any realist interpretation of SEUT according to which the confidence function is taken to be a measure of confidence.

The next three axioms give rise to an *equally-spaced utility function*, in which adjacent points on a finite utility scale are numerically spaced equally far apart. The basic idea is similar to the method employed by Davidson and Suppes (1956), whose representation theorem also establishes an equally-spaced utility function, though here reconstructed for the more general Savage-style setting. First we define the integer-valued function Δ on $\mathcal{C} \times \mathcal{C}$ recursively like so:

$$\Delta(\alpha, \beta) = \begin{cases} 0 & \text{if } \alpha \sim \beta \\ 1 & \text{if } \alpha \succ \beta \text{ and there's no } \gamma \in \mathcal{C} \text{ such that } \alpha \succ \gamma \succ \beta \\ \mathrm{n} & \text{if there is a } \gamma \text{ such that } \Delta(\alpha, \gamma) = \mathrm{n} - 1 \text{ and } \Delta(\gamma, \beta) = 1 \\ -\mathrm{n} & \text{if } \Delta(\beta, \alpha) = \mathrm{n} \end{cases}$$

Essentially, $\Delta(\alpha, \beta)$ denotes the number of successive 'steps' (from better to worse) to get from α to β in the preference ordering over consequences. Given that,

A3. There is an event E such that $E \sim_c \overline{E}$
A4. If $E \sim_c \overline{E}$, then $\alpha E \beta \sim \gamma E \delta$ implies $\Delta(\alpha, \gamma) = \Delta(\delta, \beta)$
A5. $\Delta(\alpha, \alpha') = \Delta(\beta, \beta')$ implies $\alpha E f \succsim \beta E g \Leftrightarrow \alpha' E f \succsim \beta' E g$

Henceforth, we use H (or *half-probability event*) for an event that satisfies A3. Thus A3 states that at least one half-probability event exists, the sole point of which is to ensure that A4 is non-trivially satisfied. In that case, A4 reflects the fact that if adjacent consequences are spaced equally far apart in utility, then the difference between α and γ is the same as the difference between δ and β only if there's the same number of 'steps' between them:

$$u(\alpha) - u(\gamma) = u(\delta) - u(\beta) \;\Rightarrow\; \Delta(\alpha, \gamma) = \Delta(\delta, \beta)$$

To see this, observe that $\succsim$ maximizes EU relative to a probability function p and utility function u only if for any half-probability event H, $p(H) = p(\overline{H}) = 0.5$ and therefore

$$\alpha H \beta \sim \gamma H \delta \;\Leftrightarrow\; \frac{u(\alpha) + u(\beta)}{2} = \frac{u(\gamma) + u(\delta)}{2} \;\Leftrightarrow\; u(\alpha) - u(\gamma) = u(\delta) - u(\beta)$$

In terms of the final representation, $\alpha H \beta \sim \gamma H \delta$ can therefore be thought of as saying that the midway point in utility between α and β is the same as the midway point in utility between γ and δ, from which it follows that the difference in utility between α and γ must be equal to that between δ and β. (See Figure 1.)

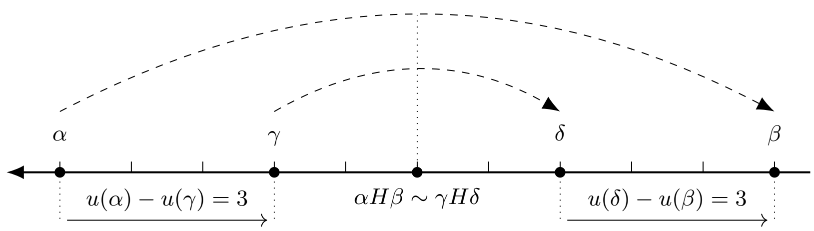

Figure 1

Combining A4 and A5, we can show that for all half-probability events H,

$$\alpha H \beta \sim \gamma H \delta \;\Leftrightarrow\; \Delta(\alpha, \gamma) = \Delta(\delta, \beta),$$

which establishes the equal-spacing structure (see Lemma 1(c)). In the presence of that structure, A5 then expresses the fact that relative expected utility depends not on the absolute utility of each consequence at a state, but instead on the differences between those utilities for each state. Compare the following decision matrices, where A5 tells us that $f \succsim g \Leftrightarrow f' \succsim g'$ precisely because the pattern of utility differences $(+2, -1, 0)$ is identical between them:

	s_1	s_2	s_3
f	1	2	4
g	3	1	4

	s_1	s_2	s_3
f'	2	1	3
g'	4	0	3

While A5 is far from necessary for consistency with SEUT in general, it *is* necessary given an equally-spaced utility distribution. In fact, in that context A5 is a natural strengthening of Savage's postulate P2, which (assuming finite $\mathcal{S}$ and $\mathcal{C}$) can be expressed:

P2. $\alpha E f \succsim \alpha E g \Leftrightarrow \beta E f \succsim \beta E g$

P2 comprises one half of Savage's Sure Thing Principle, and follows immediately from A5 since setting $\alpha' = \alpha$ and $\beta' = \beta$ gets $\Delta(\alpha, \beta) = \Delta(\alpha, \beta)$. The other half of Savage's principle is given by his P3, which is here relabeled for consistency as A6. Say that an event E is *null* just in case $\alpha E f \sim \beta E f$ for all consequences α, β; and otherwise it is *non-null*. Then we have a simple event-wise dominance condition:

A6. If E is non-null, then $\alpha \succsim \beta \Leftrightarrow \alpha E f \succsim \beta E f$

If $\succsim$ is non-trivial (i.e., $f \succ g$ for at least one pair of acts f, g), then A1–A6 already suffice for the existence of a utility function u such that

(i) $\alpha \succsim \beta \Leftrightarrow u(\alpha) \geq u(\beta)$
(ii) $\alpha H \beta \succsim \gamma H \delta \Leftrightarrow u(\alpha) - u(\gamma) \geq u(\delta) - u(\beta)$

Moreover, this utility function u will be unique up to a positive affine transformation, in the sense that u' satisfies (i) and (ii) just in case $u' = \mathrm{a} \cdot u + \mathrm{b}$ for any constants $\mathrm{a} > 0$ and b. (See Lemma 2.) As Ramsey (1931) pointed out, we can use the fact that difference ratios are always preserved under positive affine transformations to extract a unique normalized confidence function out of u under the presumption that $\succsim$ maximizes expected utility—at least provided that certain further conditions obtain.

There's more than one way to do this, and Ramsey's original strategy requires a few important assumptions. First assumption: every binary act $\alpha E \beta$ has a *certainty equivalent*—i.e., a consequence γ such that $\gamma \sim \alpha E \beta$. This is needed to pin the value every binary act to a point on the utility scale. Second assumption: the confidence function satisfies *normalized complement-additivity*—i.e., $c(E) + c(\overline{E}) = 1$ for all E. Given that, suppose $\alpha \succ \beta$ and $\gamma \sim \alpha E \beta$; then $\succsim$ maximizes EU relative to a normalized complement-additive confidence function c and a utility function u only if

$$c(E) = \frac{u(\gamma) - u(\beta)}{u(\alpha) - u(\beta)}$$

Thus Ramsey proposes to define $c(E)$ in terms of the ratios of differences in the utilities of the consequences of binary acts conditional on E and their certainty equivalents under the presumption of normalized complement-additivity. To ensure that the definition is consistent, this requires one final assumption: the ratio of $u(\gamma) - u(\beta)$ to $u(\alpha) - u(\beta)$ must be independent of the choice of consequences α, β, γ such that $\alpha \succ \beta$ and $\gamma \sim \alpha E \beta$ for any E.

For the present theorem, we pursue an alternative strategy that is broadly similar to Ramsey's approach, but requires neither certainty equivalents nor any prior assumption of normalized complement-additivity. Instead, we can use 'exchange rates' between utilities to derive confidence ratios between events. For a simple example of what I mean, consider the following matrix:

	E	$\overline{E}$
f	1	0
g	0	3

If $f \sim g$, then intuitively one unit of utility at E can be 'exchanged' for 3 units of utility at $\overline{E}$; under (MEU), this implies $c(E) = 3 \cdot c(\overline{E})$. More generally, if E and F are non-null events and there are consequences α, β, γ such that $\alpha, \beta \succ \gamma$ and $\alpha E \gamma \sim \beta F \gamma$, then $\succsim$ maximizes EU relative to any confidence function c and utility function u only if

$$\frac{c(E)}{c(F)} = \frac{u(\beta) - u(\gamma)}{u(\alpha) - u(\gamma)}$$

Fortunately, we do not have to directly determine the exchange rates that hold between every pair of non-null events—given that $\mathcal{S}$ is finite, it's enough if we pin down the exchange rates between non-null states, as that fixes the remainder of the confidence function up to multiplication by a positive constant. For this final step we need two more axioms. First, we need to ensure coherence in exchange rates. To that end, observe that the ratio between any two non-null E and E' can be established indirectly via their respective relationships with a common benchmark F:

$$\left.\begin{aligned} c(F) &= {}^{\mathrm{d}}\!/\!_{\mathrm{a}} \cdot c(E) \\ c(F) &= {}^{\mathrm{d}}\!/\!_{\mathrm{b}} \cdot c(E') \end{aligned}\right\} \Rightarrow \mathrm{b} \cdot c(E) = \mathrm{a} \cdot c(E')$$

Thus we have the following axiom, which holds necessarily for any preferences consistent with SEUT:

A7. If there are consequences $\alpha, \beta, \gamma, \delta$ and events E, E', F such that $\alpha, \beta, \gamma \succ \delta$, $\gamma E \delta \sim \alpha F \delta$, and $\gamma E' \delta \sim \beta F \delta$, then $\beta E \delta \sim \alpha E' \delta$

Finally, we need to ensure that there are sufficiently many consequences of differing value to determine an exchange rate between any two non-null states. Among other things, this is guaranteed by the final axiom:

A8. There is a non-null state s such that for all non-null states s', there are $\alpha, \beta, \gamma \in \mathcal{C}$ such that $\Delta(\beta, \gamma) \geq \Delta(\alpha, \gamma) = 1$ and $\alpha s' \gamma \sim \beta s \gamma$

With those final axioms in place, we can state the main result—essentially, if an expected utility maximizer's preferences and comparative confidences jointly satisfy A1–A8, then their confidence can be represented by exactly one probability function:

Theorem 1. *If $\mathcal{S}$ and $\mathcal{C}$ are non-empty and finite, then A1–A8 are satisfied only if (i) there is a unique probability function p such that $\succsim$ maximizes EU relative to p; and, furthermore, (ii) $E \succsim_c F$ just in case $p(E) \geq p(F)$.*

See the Appendix for the proof. I emphasize that the conclusions (i) and (ii) of Theorem 1 hold *at least* whenever A1–A8 hold. Those axioms are jointly sufficient for the result, but several of them—namely A3, A4, A5 and A8—are far from necessary. It is entirely possible, and not especially difficult, to construct cases where both conclusions are true even if some (or all) of A3, A4, A5 and A8 are violated.

This is worth noting because the non-necessary axioms are directly responsible for several intuitively undesirable restrictions on the shape of the expected utility representation. That is, A3–A5 underwrite the equal-spacing structure in the utility distribution, whereas A8 imposes a factorizability constraint on the representation of confidence: all confidence weights must be a non-negative integer multiple of the weight assigned to some minimal non-null state. These have the unfortunate consequence of ruling out representations that are clearly compatible with SEUT. However, they also ensure that differences in expected utilities are always spaced equally far apart, which greatly simplifies the existence proof. In that vein, they should be seen as relatively easy-to-state simplifying assumptions that suffice for the purposes for which Theorem 1 has been developed—most importantly, they help with the differentiation challenge. Since my argument only requires sufficient conditions, less restrictive axioms would only add complexity without altering any of the main philosophical conclusions drawn from the result.

4. The representation challenge

Having presented the theorem, we can now examine its distinctive features and how these lend support to Preference-Grounded Ratios. The three main challenges, recall, were to show that the probability function we derive from an agent's preferences

(i) represents the agent's confidence (the representation challenge),
(ii) that it does so on a ratio scale (the ratio challenge), and
(iii) that it does so in a manner which supports a preference-based explanation rather than a comparative confidence explanation (the differentiation challenge).

The present section focuses on the representation challenge, then Section 5 and Section 6 will address the ratio challenge and the differentiation challenge in turn. In all three cases, it will prove useful to compare Theorem 1 with a more typical representation theorem—namely, Savage's.

Without going into the specifics, Savage proved that if a preference relation $\succsim$ satisfies certain conditions, then it maximizes EU relative to exactly one probability function p. It is an oft-noted fact, though, that nothing about Savage's theorem implies that this function is a representation of confidence. (See, e.g., Christensen 2001: 358–364; Eriksson and Hájek 2007: 196–202; Hájek 2008: 803–4; Meacham and Weisberg 2011: 644–5; Titelbaum 2022: 298–9.) At the *very* least, if p is to count as a representation of confidence then it should track the agent's confidence relation $\succsim_c$ in the sense that

$$p(E) \geq p(F) \Leftrightarrow E \succsim_c F$$

For all Savage's theorem tells us, though, p and $\succsim_c$ may have no connection whatsoever. Moreover, the realist clearly cannot say that *what it is* for p to represent an agent's degrees of confidence *just is* for the agent to have preferences that maximize EU relative to p—at least, not without giving up realism.

Another option would be to argue that if an agent's preferences maximize EU relative to p, then the best explanation is that p is an accurate representation of the agent's degrees of confidence. That would address the representation challenge in a more realist-friendly way, but the premise strikes me as dubious at best if it's to be taken as anything more than rule of thumb. In any case, for present purposes there's a more straightforward solution: first, we include a comparative confidence relation $\succsim_c$ as an independent primitive alongside the preference relation $\succsim$; then we specify conditions under which $\succsim_c$ *correlates* with certain patterns in $\succsim$, such that information about the one can be gleaned from the other. We need not commit to saying that such conditions typically hold for ordinary agents, nor that they're normatively required; for my argument to work it is enough that the correlation is implied by SEUT under a realist construal.

In practice, the requisite conditions are easy to find, since they fall out naturally from the constraints imposed by expected utility maximization. In particular, suppose that α is better than β, and then compare the binary acts $\alpha E\beta$ versus $\alpha F\beta$. Now $\succsim$ maximizes EU relative to p only if the ordering over events induced by p corresponds to a certain pattern of preferences over such acts:

$$p(E) \geq p(F) \;\Leftrightarrow\; \alpha E\beta \succsim \alpha F\beta$$

Consequently, the requisite correlation exists whenever, for $\alpha \succsim \beta$,

$$\alpha E\beta \succsim \alpha F\beta \;\Leftrightarrow\; E \succsim_c F$$

This is exactly what A2 says. Readers familiar with Savage's work will recognize that it is just his proposed definition of a 'personal probability' ordering in terms of preferences—the difference being that because $\succsim$ and $\succsim_c$ are both treated as primitives in the framework of the theorem, we are free to interpret A2 as a condition under which two independent orderings $\succsim$ and $\succsim_c$ are correlated, rather than a stipulative reduction of one in terms of the other.

There are other representation theorems which posit a primitive confidence relation alongside the usual preference relation; Joyce's (1999) theorem is a representative example. In that respect, Theorem 1 certainly isn't unique. However, there is an important difference that's worth highlighting. In Joyce's theorem and others like it, there are several axioms that *directly* constrain the structure of $\succsim_c$—that is, independently of how $\succsim_c$ relates to $\succsim$—so as to ensure it can be represented by a probability function. Consider, for instance, what Joyce calls the *quasi-additivity* axiom, which in the present framework is given by:

$$(E \cap F = E' \cap F = \varnothing) \;\Rightarrow\; (E \succsim_c E' \Leftrightarrow E \cup F \succsim_c E' \cup F)$$

For realists, of course, there's nothing objectionable about such axioms—there may be (descriptive or rational) constraints on $\succsim_c$ that are wholly independent of how $\succsim_c$ relates to $\succsim$. However, my aim is to show that there are circumstances within expected utility theory wherein meaningful confidence ratios depend essentially on the relationship between preference and confidence, and *that* aim is best served by minimizing direct constraints on $\succsim_c$. I do not want to be accused of smuggling any ratio structure into the representation via the comparative confidence relation. Consequently, axioms A1–A8 impose almost no direct constraints on $\succsim_c$ by design. (The singular exception is A3, which only requires that at least one event is ranked equal to its complement.) Instead, A2 establishes a simple correlation between $\succsim_c$ and $\succsim$, then almost everything else about the structure of the confidence relation derives ultimately from conditions that directly constrain preferences.

5. The ratio challenge

Next is the ratio challenge. Again, it is helpful to begin by relating the issue to a familiar complaint about typical representation theorems such as Savage's—namely, that they presuppose a specifically probabilistic representation of confidence. As Mike Titelbaum put it in a recent discussion,

> The phrase "there exists a unique probabilistic credence distribution" contains a key ambiguity. One might be tempted to read it as saying that given an agent's full preference ranking of acts, there will be exactly one credence distribution that matches those preferences via expected utility maximization, and *moreover* that credence distribution will be probabilistic. But that's not how [Savage's] theorem works. The proof begins by assuming that we're looking for a probabilistic credence distribution, and then showing that out of all the probabilistic distributions, there is exactly one that will match the agent's preferences… What if it turns out that any agent who can be represented as if she is maximizing EU with respect to a *probabilistic* distribution can also be represented as maximizing EU with respect to a *non-probabilistic* distribution? (Titelbaum 2022: 300; see also Hájek 2008: 803; Meacham and Weisberg 2011: 657)

While it's not quite an *ambiguity*—being easy-to-misunderstand is not the same as being ambiguous—it *is* true that typical statements of Savage's theorem and the like leave open whether $\succsim$ might also maximize EU relative to some other *non-probabilistic* function, and if so then how those other functions might relate to p. That is, they do not specify which numerical properties of p are genuinely meaningful, as opposed to merely reflecting the decision to restrict to probabilistic representations.

Fortunately, it is not difficult to separate those properties of the probabilistic representation which correspond to genuine psychological structure from those which do not. Say that a property P is *meaningful* just in case a *non-trivial* $\succsim$ maximizes EU relative to some confidence function that satisfies P only if it *doesn't* maximize EU relative to any confidence function that *doesn't* satisfy P.[4] The meaningful properties of a confidence function c are, in other words, those properties that are forced by the structure of $\succsim$ under the assumption that $\succsim$ and c are related via (MEU). Once that relationship is fixed, the preferences leave no freedom of choice with respect to meaningful properties.[5]

Given that, let Π be the set of all non-trivial preference orderings $\succsim$ such that $\succsim$ maximizes EU relative to some confidence function c. P is meaningful just in case there's a non-empty $\Pi^\star \subseteq \Pi$ such that for all $\succsim \in \Pi$,

(i) if $\succsim \in \Pi^\star$, then $\succsim$ maximizes EU relative to c only if c satisfies P
(ii) if $\succsim \notin \Pi^\star$, then $\succsim$ maximizes EU relative to c only if c satisfies $\neg$P

This lets us distinguish two ways in which P might be meaningful. On the one hand, P is *trivially* meaningful if $\Pi^\star = \Pi$—that is, if P is strictly required for any kind of expected utility maximization at all, in the sense that there are no $\succsim$ and c related via (MEU) such that c doesn't satisfy P. On the other hand, P is *substantively* meaningful if $\Pi^\star \subset \Pi$, in which case $\neg$P is also substantively meaningful. Every complementary pair of substantively meaningful properties P and $\neg$P will partition Π and thereby corresponds to a genuine distinction within the class of possible expected utility maximizing preferences.

It is straightforward to prove that additivity (in the confidence function) and non-constancy (in both the confidence function and the utility function) are trivially meaningful, each falling directly out of the decision rule encoded by (MEU) in all but the degenerate case where $\succsim$ is trivial:

Proof. Assume $\succsim$ maximizes EU relative to some c and u. Given that $\succsim$ is non-trivial, u must be non-constant. As such, consider any two consequences α and β where $u(\alpha) \neq u(\beta)$, and any pair of disjoint events E and F. Since $\alpha E{\cup}F\alpha = \alpha E\alpha F\alpha$ and $\beta E{\cup}F\beta = \beta E\beta F\beta$, the partition-independence constraint on (MEU) implies

$$\begin{aligned} c(E \cup F) \cdot u(\alpha) + c(\overline{E \cup F}) \cdot u(\alpha) &= c(E) \cdot u(\alpha) + c(F) \cdot u(\alpha) + c(\overline{E \cup F}) \cdot u(\alpha) \\ c(E \cup F) \cdot u(\beta) + c(\overline{E \cup F}) \cdot u(\beta) &= c(E) \cdot u(\beta) + c(F) \cdot u(\beta) + c(\overline{E \cup F}) \cdot u(\beta) \end{aligned}$$

[4] If $\succsim$ is trivial, then it trivially maximizes EU relative to every confidence function. This tells us nothing interesting about what is or is not meaningful in the probabilistic representation of confidence, only that we ought to ignore the degenerate case.

[5] This characterization of meaningfulness is intended to fit with standard measurement-theoretic usage; i.e., invariance under automorphisms *given* a fixed choice of representational structure (Pfanzagl 1968; Luce 1978; Narens 1985; Luce et al. 1990). Do keep in mind that meaningfulness is defined *relative to* a fixed assumption about how confidence and preference are related as encoded by the rule (MEU). This is worth flagging especially in relation to a well-known and much-discussed point from Zynda (2000); namely, that a probability function p may non-arbitrarily satisfy a certain property P (e.g., additivity) under the assumption that $\succsim$ and p are related via expected utility maximization, but that same property P need not be satisfied by some other confidence function c where $\succsim$ and c are related via some rule *other than* (MEU). For the reasons pointed out recently in (Elliott 2024: 18–20, 65–69), the existence of these alternative representations is irrelevant to what is and is not meaningful within SEUT as standardly represented. I acknowledge them here only so that I may explicitly set them aside.

Since $u(\alpha) \neq u(\beta)$, together these imply $c(E \cup F) = c(E) + c(F)$. So c is additive. Additivity implies $c(\varnothing) = 0$, so c can be constant only if $c(E) = 0$ for all E; but that would imply $\succsim$ is trivial, so c is non-constant. □

By contrast, neither normalization nor non-negativity are trivially meaningful, since we can readily concoct preferences consistent with (MEU) relative to confidence functions that satisfy neither property. For instance, suppose that $E, \overline{E}$ are non-empty, and let

$$\begin{array}{lll} c(\varnothing) = 0 & c(E) = 1 & u(\alpha) = 1 \\ c(\mathcal{S}) = 0 & c(\overline{E}) = -1 & u(\beta) = 0 \end{array}$$

Now consider the following preferences, which maximize EU relative to c and u:

$$\alpha E \beta \succ \alpha E \alpha \sim \beta E \beta \succ \beta E \alpha$$

You'll note, though, that those preferences violate axiom A6: E is non-null and $\alpha \sim \beta$, yet $\alpha E \beta \succ \beta E \beta$ and $\beta \overline{E} \alpha \succ \alpha \overline{E} \alpha$. Say that c is *zero-bounded* just in case it's either non-negative or non-positive; otherwise it is *mixed-sign*. As it turns out, these kinds of violations arise exactly whenever c is mixed-sign. That is, zero-boundedness and mixed-signedness are meaningful, corresponding precisely to those preferences in Π which do and do not satisfy A6 respectively:

Proof. Assume that $\succsim$ maximizes EU relative to (additive, non-constant) c and (non-constant) u, and take any E with $c(E) \neq 0$. Let $v(f)$ be the expected utility of f. Then for any α, β, f,

$$v(\alpha E f) - v(\beta E f) = c(E)\big(u(\alpha) - u(\beta)\big)$$
$$v(\alpha) - v(\beta) = c(\mathcal{S})\big(u(\alpha) - u(\beta)\big)$$

If c is zero-bounded, $c(E)$ and $c(\mathcal{S})$ have the same sign; so

$$v(\alpha) - v(\beta) \geq 0 \Leftrightarrow v(\alpha E f) - v(\beta E f) \geq 0,$$

which implies A6 under (MEU). On the other hand, if c is mixed sign, then there are three possibilities: $c(\mathcal{S}) = 0$, $c(\mathcal{S}) > 0$, and $c(\mathcal{S}) < 0$. If $c(\mathcal{S}) = 0$, pick any α, β such that $u(\alpha) > u(\beta)$, and pick an event E such that $c(E) < 0$ and $c(\overline{E}) > 0$. Observe that $v(\alpha E \alpha) = c(\mathcal{S}) \cdot u(\alpha) = 0$ and likewise $v(\beta E \beta) = c(\mathcal{S}) \cdot u(\beta) = 0$, so (MEU) implies $\alpha \sim \beta$. However, (MEU) also implies $v(\beta E \alpha) > v(\alpha E \alpha)$ and so $\beta E \alpha \succ \alpha E \alpha$, which violates A6. That violations of A6 also arise whenever $c(\mathcal{S}) > 0$ or $c(\mathcal{S}) < 0$ can be readily established using similar patterns of reasoning, so c is zero-bounded just in case $\succsim$ satisfies A6. □

We can now specify the exact role that normalization and non-negativity play within SEUT. In the presence of additivity, they pin down one particular way of instantiating non-constancy and zero-boundedness; however, once those properties are fixed, normalization and non-negativity impose no further restrictions on the kinds of preferences consistent with SEUT. This is trivially easy to see in the case of normalization. Suppose $\succsim$ maximizes EU relative to c and u, and let c' be any positive scalar transformation of c (i.e., $c' = \mathrm{a} \cdot c$ for $\mathrm{a} > 0$). Then $\succsim$ also maximizes EU relative to c' and u, since multiplication by a positive scalar preserves the ordering of expected utilities. It is similarly easy to see in the case of non-negativity: let $-c = -1 \cdot c$ and likewise for $-u$, then

$$\sum_{E \in \mathcal{E}_f} c(E) \cdot u(f(E)) = \sum_{E \in \mathcal{E}_f} -c(E) \cdot -u(f(E))$$

In general, then, c relates to $\succsim$ via (MEU) just in case the sign-inverse of c does too (i.e., provided the utility function is adjusted accordingly). Putting the two points together: $\succsim$ maximizes EU relative to any c just in case it also does so relative to any (positive or negative) scalar transformation of c. Non-negativity and normalization therefore amount to an arbitrary choice of sign and unit respectively—they ensure the confidence function is non-constant and zero-bounded, but beyond that they are theoretically idle.

There are two lessons here. Lesson one: if $\succsim$ maximizes EU relative to any c, then c is additive, non-constant, and unique *at most* up to a scalar transformation. The 'at most' is emphasized for a reason—for all I've said so far, it is still possible that $\succsim$ might maximize EU relative to c and *also* relative to c', even where c' is *not* a scalar transformation of c. Absent further constraints on the representation, it is not possible to fix a uniqueness condition *stronger than* multiplication by a non-zero constant. But now the second lesson: to show that (a) $\succsim$ maximizes EU relative to exactly one probability function is essentially the same thing as showing that (b) $\succsim$ maximizes EU relative to an additive, non-constant and zero-bounded confidence function c that's unique up to a scalar transformation. The proof of their equivalence is straightforward:

Proof. That (b) implies (a) is immediate. For the other direction, assume that $\succsim$ maximizes EU relative to a unique probability function p. From the points above, $\succsim$ therefore maximizes EU relative to every c that's a scalar transformation of p, and *doesn't* maximize EU relative to any c that's non-additive, constant or mixed-sign. As such, let $c^\star$ be any additive, non-constant and zero-bounded confidence function that's not a scalar transformation of p, and suppose $\succsim$ maximizes EU relative to $c^\star$. Then there will exist some $p^\star \neq p$ such that $p^\star$ is a scalar transformation of $c^\star$ and $\succsim$ maximizes EU relative to $p^\star$, contradicting the initial assumption. □

Now, finally, we can address the ratio challenge. If A1–A8 are satisfied, Theorem 1 tells us that $\succsim$ maximizes EU relative to a unique probability function p, and therefore also relative to a confidence function that's unique up to a scalar transformation. Scalar transformations are precisely those which preserve ratio information, in the sense that c is a scalar transformation of c' only if

$$\frac{c(E)}{c(F)} = \frac{c'(E)}{c'(F)}$$

whenever $c(F) \neq 0$, otherwise $c(F) = 0$ implies $c'(F) = 0$. That is to say, the structure of the preference relation forces a particular *pattern of ratios among events* in any confidence function to which $\succsim$ relates via (MEU)—so that pattern characterizes precisely a class of confidence functions that are indistinguishable from the perspective of expected utility maximization.

6. The differentiation challenge

Having addressed the representation and ratio challenges, what remains is the differentiation challenge. To that end, this section argues that there are cases in which a preference ordering maximizes EU relative to exactly one probability function, and that function contains meaningful ratio information that cannot in principle be explained solely by the structure of the comparative confidence relation (even when that structure is strengthened in natural ways). The strategy is therefore to construct a case in which two probability functions correspond to the same confidence ordering, yet are distinguished by preferences in a way that Theorem 1 implies is significant.

I should start by saying more about comparative confidence explanations of confidence ratios. The simplest version rests on what might be called the *equal-parts principle*:

> For any event E, if (i) the agent has more confidence in E than in $\varnothing$, (ii) E can be partitioned into n sub-events $F_1, \ldots, F_n$, and (iii) for any two of those sub-events F_i, F_j, the agent has the same amount of confidence in F_i as they do in F_j, then the agent has $1/n$ times as much confidence in each of the sub-events F_i as they do in E.

An intuitive way to see how this works is to imagine that $\succsim_c$ represents the relative *size* of events—so $E \sim_c F$ means E is the same size as F, while $E \succ_c F$ means E is bigger than F. We assume that every event is at least as big as $\varnothing$, which has zero size; and that sizes combine additively in the usual way, so the size of any event is just the sum of the sizes of its disjoint parts. Now suppose that $\mathcal{S}$ contains just three states s_1, s_2, s_3, and the total size ordering is given by

$$\mathcal{S} \succ_c \{s_1, s_2\} \sim_c \{s_1, s_3\} \succ_c \{s_2, s_3\} \sim_c s_1 \succ_c s_2 \sim_c s_3 \succ_c \varnothing$$

As pictured in Figure 2(a), the ordering implies that $\mathcal{S}$ can be divided into two equal-sized parts, s_1 and $\{s_2, s_3\}$; and then $\{s_2, s_3\}$ can in turn be divided into two equal-sized parts, s_2 and s_3. Given that, s_1 and $\{s_2, s_3\}$ must each be half the size of $\mathcal{S}$, while s_2 and s_3 must be half the size of $\{s_2, s_3\}$ and therefore one quarter the size of $\mathcal{S}$. Thus by application of the equal-parts principle, it is *sometimes* possible to determine the scale of each state in relation to the size of the total state space merely by considering the structure of $\succsim_c$ alone—that is, without any direct reference to preferences.

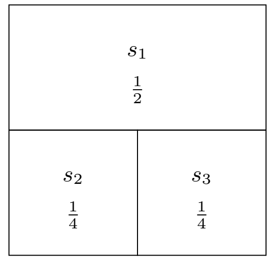

Figure 2

On the other hand, it is often impossible to determine precise ratio relations between events merely given the structure of $\succsim_c$. Say that a probability function p is *ordinally determined* just in case, if p represents a confidence ordering $\succsim_c$, then no other probability function also represents $\succsim_c$. Clearly, $\succsim_c$ suffices to explain precise ratio relations between all events only if it can be represented by an ordinally determined probability function. So to find problem cases for the comparative confidence explanation of confidence ratios—that is, cases that differentiate between comparative confidence versus preference-based explanations—we must look towards probability functions that are ordinally *under*determined.

I will describe such a case momentarily, but first it is worth noting that the axioms of Savage's theorem are inconsistent with exactly this kind of differentiation. That is, for any preferences that satisfy Savage's axioms, those preferences will maximize EU only relative to an ordinally determined probability function. This is ultimately a result of Savage's derivation strategy, which proceeds in three main steps: first, he derives a confidence ordering from the preference relation; then he employs a generalization of the equal-parts principle to show that this confidence ordering has a unique probabilistic representation; and finally he derives the utility function that fits those preferences relative to those probabilities. The crucial part is the second step—it precludes any probability functions that are ordinally underdetermined. But those are precisely the kinds of probability functions that are needed for the present argument. So Savage's theorem, and for that matter any other theorems that employ the typical *preferences* $\rightarrow$ *probabilities* $\rightarrow$ *utilities* derivation method, are inadequate for addressing the differentiation challenge.

By contrast, for Theorem 1 the usual order of derivation is reversed: first we derive a utility function from preferences over binary acts, then we derive the unique probability function that fits the full preference ordering over all acts in light of those utilities. This alternative *preferences* $\rightarrow$ *utilities* $\rightarrow$ *probabilities* method bypasses any appeal to equal-parts structure in the confidence ordering, appealing instead to an equal-spacing structure in the preference ordering over consequences. Moreover, it permits cases where the preferences maximize EU relative to a unique probability function even if that function does not *uniquely* represent its corresponding confidence ordering.

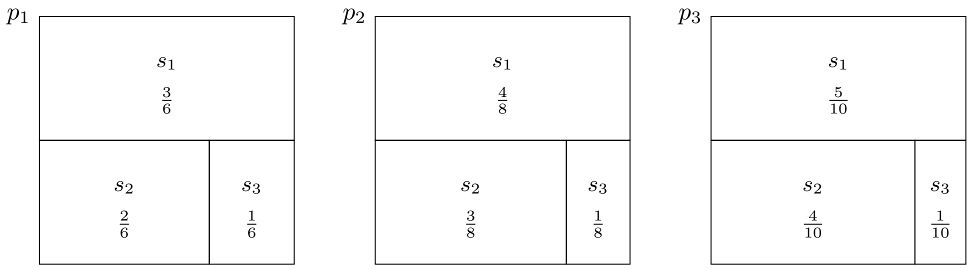

Figure 3

Consider the three probability functions p_1, p_2 and p_3 represented in Figure 3. They each carry rather different ratio information, but correspond to the same confidence ordering:

$$\mathcal{S} \succ_c \{s_1, s_2\} \succ_c \{s_1, s_3\} \succ_c \{s_2, s_3\} \sim_c s_1 \succ_c s_2 \succ_c s_3 \succ_c \varnothing$$

Set aside p_3 for now; it'll be important later. We want to show that there exists a pair of preference orderings $\succsim_1$ and $\succsim_2$ such that, for $\succsim_c$ as just specified,

(i) $\succsim_1$ and $\succsim_c$ satisfy A1–A8, and $\succsim_1$ maximizes EU relative to p_1 (and not p_2)
(ii) $\succsim_2$ and $\succsim_c$ satisfy A1–A8, and $\succsim_2$ maximizes EU relative to p_2 (and not p_1)

The simplest possible example requires a minimum of five consequences. Let $\mathcal{C} = \{0, 1, 2, 3, 4\}$ and $u(i) = i$ for $i = 0, \ldots, 4$; then Table 1 represents the ordering $\succsim_1$ generated by p_1 and u under the expected utility rule, while Table 2 represents the ordering $\succsim_2$ generated by p_2 and u. Both $\succsim_1$ and $\succsim_2$ satisfy all the axioms of Theorem 1. Since the acts in Table 1 and Table 2 are arranged by expected utility, it is immediate that the necessary axioms A1, A2, A6 and A7 are satisfied. Furthermore, since the utilities are equally spaced, it is also immediate that A4 and A5 are also satisfied. A3 is satisfied by $s_1 \sim_c \{s_2, s_3\}$. Finally, A8 simply requires the existence of a non-null state—in both cases it's s_3—such that for some $x, y > 0$, $f \sim g$ and $f' \sim g'$:

	s_1	s_2	s_3
f	1	0	0
g	0	0	x

	s_1	s_2	s_3
f'	0	1	0
g'	0	0	y

For p_1, $x = 3$ and $y = 2$; whereas for p_2, $x = 4$ and $y = 3$.

So Theorem 1 implies that p_1 is the unique probability function to which $\succsim_1$ maximizes EU, and p_2 is the unique probability function to which $\succsim_2$ maximizes EU—even though p_1 and p_2 correspond to the same confidence ordering. Since we have already concluded that, under these circumstances, the ratio relations in p_1 and p_2 are meaningful, the result is that there's more to what makes those ratios meaningful in this case than can be explained solely in terms of $\succsim_c$.

s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu
0	0	0	0.0	0	2	4	1.33	2	3	0	2.0	3	2	3	2.67
0	0	1	0.17	1	2	1	1.33	3	0	3	2.0	3	3	1	2.67
0	0	2	0.33	0	3	3	1.5	3	1	1	2.0	4	0	4	2.67
0	1	0	0.33	0	4	1	1.5	4	0	0	2.0	4	2	0	2.67
0	0	3	0.5	1	1	4	1.5	1	4	2	2.17	2	4	3	2.83
0	1	1	0.5	1	2	2	1.5	3	1	2	2.17	3	3	2	2.83
1	0	0	0.5	1	3	0	1.5	1	3	4	2.17	3	4	0	2.83
0	1	2	0.67	2	0	3	1.5	2	2	3	2.17	4	1	3	2.83
0	0	4	0.67	2	1	1	1.5	2	3	1	2.17	3	2	4	2.83
0	2	0	0.67	3	0	0	1.5	3	0	4	2.17	4	2	1	2.83
1	0	1	0.67	0	4	2	1.67	3	2	0	2.17	2	4	4	3.0
0	1	3	0.83	2	1	2	1.67	4	0	1	2.17	3	3	3	3.0
1	0	2	0.83	0	3	4	1.67	1	4	3	2.33	3	4	1	3.0
1	1	0	0.83	1	2	3	1.67	2	3	2	2.33	4	1	4	3.0
0	2	1	0.83	1	3	1	1.67	2	4	0	2.33	4	2	2	3.0
0	1	4	1.0	2	0	4	1.67	3	1	3	2.33	4	3	0	3.0
0	2	2	1.0	2	2	0	1.67	4	0	2	2.33	3	4	2	3.17
0	3	0	1.0	3	0	1	1.67	4	1	0	2.33	3	3	4	3.17
1	0	3	1.0	0	4	3	1.83	2	2	4	2.33	4	2	3	3.17
1	1	1	1.0	1	3	2	1.83	3	2	1	2.33	4	3	1	3.17
2	0	0	1.0	1	4	0	1.83	1	4	4	2.5	3	4	3	3.33
1	1	2	1.17	2	1	3	1.83	2	3	3	2.5	4	3	2	3.33
0	2	3	1.17	3	0	2	1.83	2	4	1	2.5	4	4	0	3.33
0	3	1	1.17	3	1	0	1.83	3	1	4	2.5	4	2	4	3.33
1	0	4	1.17	1	2	4	1.83	3	2	2	2.5	3	4	4	3.5
1	2	0	1.17	2	2	1	1.83	3	3	0	2.5	4	3	3	3.5
2	0	1	1.17	0	4	4	2.0	4	0	3	2.5	4	4	1	3.5
0	3	2	1.33	1	3	3	2.0	4	1	1	2.5	4	4	2	3.67
0	4	0	1.33	1	4	1	2.0	2	4	2	2.67	4	3	4	3.67
1	1	3	1.33	2	1	4	2.0	4	1	2	2.67	4	4	3	3.83
2	0	2	1.33	2	2	2	2.0	2	3	4	2.67	4	4	4	4.0
2	1	0	1.33												

Table 1: $p(s_1) = 3/6$, $p(s_2) = 2/6$, $p(s_3) = 1/6$

s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu	s_1	s_2	s_3	eu
0	0	0	0.0	1	1	4	1.38	3	0	4	2.0	4	1	2	2.63
0	0	1	0.13	2	0	3	1.38	3	1	1	2.0	3	3	1	2.75
0	0	2	0.25	2	1	0	1.38	4	0	0	2.0	2	4	2	2.75
0	0	3	0.38	0	3	3	1.5	1	3	4	2.13	3	2	4	2.75
0	1	0	0.38	0	4	0	1.5	1	4	1	2.13	4	2	0	2.75
0	0	4	0.5	1	2	2	1.5	2	3	0	2.13	4	1	3	2.75
0	1	1	0.5	2	0	4	1.5	4	0	1	2.13	3	3	2	2.88
1	0	0	0.5	2	1	1	1.5	2	2	3	2.13	4	2	1	2.88
1	0	1	0.63	3	0	0	1.5	3	1	2	2.13	2	4	3	2.88
0	1	2	0.63	0	3	4	1.63	2	3	1	2.25	4	1	4	2.88
0	2	0	0.75	0	4	1	1.63	1	4	2	2.25	2	4	4	3.0
1	0	2	0.75	1	3	0	1.63	2	2	4	2.25	3	3	3	3.0
0	1	3	0.75	3	0	1	1.63	3	2	0	2.25	3	4	0	3.0
0	2	1	0.88	1	2	3	1.63	4	0	2	2.25	4	2	2	3.0
0	1	4	0.88	2	1	2	1.63	3	1	3	2.25	3	3	4	3.13
1	0	3	0.88	1	3	1	1.75	2	3	2	2.38	3	4	1	3.13
1	1	0	0.88	0	4	2	1.75	3	2	1	2.38	4	3	0	3.13
0	2	2	1.0	1	2	4	1.75	1	4	3	2.38	4	2	3	3.13
1	0	4	1.0	2	2	0	1.75	3	1	4	2.38	4	3	1	3.25
1	1	1	1.0	3	0	2	1.75	4	0	3	2.38	3	4	2	3.25
2	0	0	1.0	2	1	3	1.75	4	1	0	2.38	4	2	4	3.25
0	3	0	1.13	1	3	2	1.88	1	4	4	2.5	4	3	2	3.38
2	0	1	1.13	2	2	1	1.88	2	3	3	2.5	3	4	3	3.38
0	2	3	1.13	0	4	3	1.88	2	4	0	2.5	3	4	4	3.5
1	1	2	1.13	2	1	4	1.88	3	2	2	2.5	4	3	3	3.5
0	3	1	1.25	3	0	3	1.88	4	0	4	2.5	4	4	0	3.5
0	2	4	1.25	3	1	0	1.88	4	1	1	2.5	4	3	4	3.63
1	2	0	1.25	0	4	4	2.0	2	3	4	2.63	4	4	1	3.63
2	0	2	1.25	1	3	3	2.0	2	4	1	2.63	4	4	2	3.75
1	1	3	1.25	1	4	0	2.0	3	3	0	2.63	4	4	3	3.88
0	3	2	1.38	2	2	2	2.0	3	2	3	2.63	4	4	4	4.0
1	2	1	1.38												

Table 2: $p(s_1) = 4/8$, $p(s_2) = 3/8$, $p(s_3) = 1/8$

You might be wondering, though, whether this depends on a relatively impoverished understanding of comparative confidence. Most notably, rather than focusing on an *unconditional* confidence relation over individual events, we might instead follow Koopman (1940) in taking *conditional* comparative confidences to be basic. Let $\succsim_k$ be a binary relation over $2^{\mathcal{S}} \times (2^{\mathcal{S}} \setminus \{\varnothing\})$, where $(E, E') \succsim_k (F, F')$ means that the agent has at least as much confidence in E *given* E' as in F *given* F'. We assume that *unconditional* confidence comparisons are really just comparisons conditional on the necessary event, so

$$E \succsim_c F \Leftrightarrow (E, \mathcal{S}) \succsim_k (F, \mathcal{S})$$

So $\succsim_k$ includes $\succsim_c$ as a proper part; however, since $\succsim_k$ also includes comparisons not captured by $\succsim_c$, the richer conditional confidence ordering permits more fine-grained differentiation among probability functions.

The following tables represent conditional probabilities $p(E \mid F)$ only for the cases where $E \subset F$ and $E \neq \varnothing$. The trivial cases where $E = \varnothing$ (so $p(E|F) = 0$) and $E \cap F = F$ (so $p(E|F) = 1$) have been excluded; and all other comparisons (i.e., where $E \not\subseteq F$ and $E \cap F \neq \varnothing$) are fully determined by those comparisons represented here. What they tell us is that even though p_1 and p_2 represent the same *unconditional* confidence ordering, they correspond to different *conditional* confidence orderings:

E	F	$p_1(E)$	$p_1(F)$	$p_1(E\|F)$
$\{s_1, s_2\}$	$\mathcal{S}$	$5/6$	$6/6$	$5/6 \approx 0.83$
s_1	$\{s_1, s_3\}$	$3/6$	$4/6$	$3/4 = 0.75$
s_2	$\{s_2, s_3\}$	$2/6$	$3/6$	$2/3 \approx 0.67$
$\{s_1, s_3\}$	$\mathcal{S}$	$4/6$	$6/6$	$4/6 \approx 0.67$
s_1	$\{s_1, s_2\}$	$3/6$	$5/6$	$3/5 = 0.60$
s_1	$\mathcal{S}$	$3/6$	$6/6$	$3/6 = 0.50$
$\{s_2, s_3\}$	$\mathcal{S}$	$3/6$	$6/6$	$3/6 = 0.50$
s_2	$\{s_1, s_2\}$	$2/6$	$5/6$	$2/5 = 0.40$
s_2	$\mathcal{S}$	$2/6$	$6/6$	$2/6 \approx 0.33$
s_3	$\{s_2, s_3\}$	$1/6$	$3/6$	$1/3 \approx 0.33$
s_3	$\{s_1, s_3\}$	$1/6$	$4/6$	$1/4 = 0.25$
s_3	$\mathcal{S}$	$1/6$	$6/6$	$1/6 \approx 0.17$

E	F	$p_2(E)$	$p_2(F)$	$p_2(E\|F)$
$\{s_1, s_2\}$	$\mathcal{S}$	$7/8$	$8/8$	$7/8 \approx 0.88$
s_1	$\{s_1, s_3\}$	$4/8$	$5/8$	$4/5 = 0.80$
s_2	$\{s_2, s_3\}$	$3/8$	$4/8$	$3/4 = 0.75$
$\{s_1, s_3\}$	$\mathcal{S}$	$5/8$	$8/8$	$5/8 \approx 0.63$
s_1	$\{s_1, s_2\}$	$4/8$	$7/8$	$4/7 \approx 0.57$
s_1	$\mathcal{S}$	$4/8$	$8/8$	$4/8 = 0.50$
$\{s_2, s_3\}$	$\mathcal{S}$	$4/8$	$8/8$	$4/8 = 0.50$
s_2	$\{s_1, s_2\}$	$3/8$	$7/8$	$3/7 \approx 0.43$
s_2	$\mathcal{S}$	$3/8$	$8/8$	$3/8 \approx 0.38$
s_3	$\{s_2, s_3\}$	$1/8$	$4/8$	$1/4 = 0.25$
s_3	$\{s_1, s_3\}$	$1/8$	$5/8$	$1/5 = 0.20$
s_3	$\mathcal{S}$	$1/8$	$8/8$	$1/8 \approx 0.13$

In other words, if $\succsim_k^1$ is the conditional confidence ordering corresponding to p_1, and $\succsim_k^2$ is the conditional confidence ordering corresponding to p_2, then there are (at least) two locations where $\succsim_k^1$ and $\succsim_k^2$ must diverge:

1. $(s_2, \{s_1, s_3\}) \sim_k^1 (\{s_1, s_3\}, \mathcal{S})$, whereas $(s_2, \{s_1, s_3\}) \succ_k^2 (\{s_1, s_3\}, \mathcal{S})$
2. $(s_2, \mathcal{S}) \sim_k^1 (s_3, \{s_2, s_3\})$, whereas $(s_2, \mathcal{S}) \succ_k^2 (s_3, \{s_2, s_3\})$

However, while this *does* help for the particular example chosen above, there are plenty of *other* cases where it does not help. Rather than comparing p_1 and p_2, suppose we compare p_2 and p_3 instead. Just as there are preferences $\succsim_1$ and $\succsim_2$ that maximize EU relative to p_1 and p_2 respectively, it is straightforward to show that there is are preferences $\succsim_3$ that maximizes EU relative to p_3 (and not p_1 or p_2).[6] Unlike p_1 and p_2, though, p_2 and p_3 correspond to the exactly the same conditional confidence ordering:

E	F	$p_2(E)$	$p_2(F)$	$p_2(E\|F)$
$\{s_1, s_2\}$	$\mathcal{S}$	$7/8$	$8/8$	$7/8 \approx 0.88$
s_1	$\{s_1, s_3\}$	$4/8$	$5/8$	$4/5 = 0.80$
s_2	$\{s_2, s_3\}$	$3/8$	$4/8$	$3/4 = 0.75$
$\{s_1, s_3\}$	$\mathcal{S}$	$5/8$	$8/8$	$5/8 \approx 0.63$
s_1	$\{s_1, s_2\}$	$4/8$	$7/8$	$4/7 \approx 0.57$
s_1	$\mathcal{S}$	$4/8$	$8/8$	$4/8 = 0.50$
$\{s_2, s_3\}$	$\mathcal{S}$	$4/8$	$8/8$	$4/8 = 0.50$
s_2	$\{s_1, s_2\}$	$3/8$	$7/8$	$3/7 \approx 0.43$
s_2	$\mathcal{S}$	$3/8$	$8/8$	$3/8 \approx 0.38$
s_3	$\{s_2, s_3\}$	$1/8$	$4/8$	$1/4 = 0.25$
s_3	$\{s_1, s_3\}$	$1/8$	$5/8$	$1/5 = 0.20$
s_3	$\mathcal{S}$	$1/8$	$8/8$	$1/8 \approx 0.13$

E	F	$p_3(E)$	$p_3(F)$	$p_3(E\|F)$
$\{s_1, s_2\}$	$\mathcal{S}$	$9/10$	$10/10$	$9/10 = 0.90$
s_1	$\{s_1, s_3\}$	$5/10$	$6/10$	$5/6 \approx 0.83$
s_2	$\{s_2, s_3\}$	$4/10$	$5/10$	$4/5 = 0.80$
$\{s_1, s_3\}$	$\mathcal{S}$	$6/10$	$10/10$	$3/5 = 0.60$
s_1	$\{s_1, s_2\}$	$5/10$	$9/10$	$5/9 \approx 0.56$
s_1	$\mathcal{S}$	$5/10$	$10/10$	$5/10 = 0.50$
$\{s_2, s_3\}$	$\mathcal{S}$	$5/10$	$10/10$	$5/10 = 0.50$
s_2	$\{s_1, s_2\}$	$4/10$	$9/10$	$4/9 \approx 0.44$
s_2	$\mathcal{S}$	$4/10$	$10/10$	$2/5 = 0.40$
s_3	$\{s_2, s_3\}$	$1/10$	$5/10$	$1/5 = 0.20$
s_3	$\{s_1, s_3\}$	$1/10$	$6/10$	$1/6 \approx 0.17$
s_3	$\mathcal{S}$	$1/10$	$10/10$	$1/10 = 0.10$

So enriching the confidence structure *à la* Koopman will not suffice to cover all the cases that are apt for a preference-based explanation. But we could have seen that this must be true even without considering particular cases. Keeping $\mathcal{S} = \{s_1, s_2, s_3\}$, there are 56 ordered pairs in $2^{\mathcal{S}} \times (2^{\mathcal{S}} \setminus \{\varnothing\})$ and so at most 2^{3136} possible conditional confidence relations—only a tiny subset of which will determine a unique probabilistic representation. On the other hand, p_1, p_2, p_3 represent just the start of an infinite sequence of probability functions all of which can be given a preference-based explanation. Take any probability function p such that for $\mathrm{n} \geq 3$,

$$p(s_1) = \frac{1}{2}, \quad p(s_2) = \frac{\mathrm{n}-1}{2\mathrm{n}}, \quad p(s_3) = \frac{1}{2\mathrm{n}}$$

Given a large enough $\mathcal{C}$, we can easily characterize a preference ordering $\succsim$ such that p is the unique probability function relative to which $\succsim$ maximizes expected utility. So there are more probability functions that can be uniquely recovered from some preference ordering than can in principle be uniquely recovered from conditional comparative confidence relations.

The argument generalizes. Besides Koopman-style conditional confidence relations, others have suggested enriching the simple comparative confidence framework by supplementing $\succsim_c$ with a separate *independence* relation between events (e.g., Domotor 1970; Kaplan and Fine 1977), or a qualitative *uncertainty* relation over sets of events (e.g., Suppes 2014). In both cases, the additional structure helps draw distinctions between probability functions that cannot be drawn just in terms of a simple comparative confidence relation. But for *any* finite set of relations over events (or sets of events), all of finite arity, there can be at most finitely many probability functions that are uniquely determined by the relations in that set—whatever they may be. So there will always be cases in which a preference ordering maximizes EU relative

[6] The minimal example of such a preference ordering requires at least six consequences, with $\succsim_3$ therefore being defined over a space of $6^3 = 216$ acts—so I have opted not to reproduce that ordering explicitly here.

to exactly one probability function, where that function contains meaningful ratio information that cannot be explained solely by the structure of the comparative confidence relation—even when that structure is enriched or supplemented in any of the ways that have thus far been suggested in the literature.

Finally, one might worry that there may also be cases in which a purely confidence-based strategy can determine precise ratios while a preference-based explanation cannot. This possibility can be ruled out in the case of the simple comparative confidence strategy. It is a consequence of the correlation between $\succsim$ and $\succsim_c$ expressed by A2 that, if any probability function p is ordinally determined (so p is the kind of function for which the simple comparative confidence strategy is apt), then $\succsim$ maximizes EU relative to p if and only if p is the unique probability function relative to which $\succsim$ maximizes EU (so p is also the kind of function for which a preference-based strategy is apt).[7] So if the simple comparative confidence explanation is available then so is a preference-based explanation, but not vice versa.

The situation is less clear once the confidence structure is enriched, for instance, by adopting conditional comparative confidence relations in the style of Koopman. It may be that some such structures determine ratios in cases where preference-based explanations do not apply—I do not have a proof that such cases are impossible, though I'd be surprised if they weren't. In any case, it doesn't matter. My thesis is not that preference-based explanations account for *every* possible instance of meaningful ratio information. It is consistent with PREFERENCE-GROUNDED RATIOS that different explanatory strategies may apply in different circumstances, or that some cases require a combination of qualitative structures in order to recover precise ratios. It is also consistent with the possibility that some patterns of confidence admit no ratio-scale representation at all. The point established here is simply that there exist circumstances in which meaningful confidence ratios arise in virtue of the systematic interaction between confidence and preference—circumstances that cannot be captured by confidence relations alone.

7. Conclusion

Representation theorems have often been used to motivate constructivist views on which probabilities are merely mathematical constructions derived from preferences, and all this talk of 'deriving' a confidence function from preferences certainly seems reminiscent of constructivism. This association, I believe, accounts for much of the suspicion towards preference-based explanations of confidence ratios among realists—especially when these are set in contrast to explanations that require no more than an appropriately structured confidence ordering. But I would prefer to dispel the association between decision-theoretic representation theorems and constructivism. The latter can be useful tools for realists, too, and the claim that there are circumstances

[7] Such preferences need not satisfy all the axioms of Theorem 1. A preference-based explanation is available at least whenever $\succsim$ maximizes EU relative to exactly one probability function (and there's an appropriate correlation between $\succsim$ and $\succsim_c$). A1–A8 therefore express jointly sufficient conditions for when such an explanation to be possible, but not all of them are individually necessary.

where what's meaningful in a confidence function can only be properly explained by reference to preference should in no way be taken to entail a constructivist theory of confidence.

As a remedy, it can be useful to think of confidence as being more like density than it is like size. Size ratios reflect a *direct* connection between the structure of the relative size ordering and concatenations of its relata—if a is n times bigger than b, then you can concatenate n copies of b to form an object the same size as a. That is the intuitive thought at the heart of the equal-parts strategy discussed in Section 6. And as a consequence, it is possible to explain how size ratios come to be meaningful without essential reference to any other quantities. One doesn't need to know anything about the weights of a and b, for example, or their temperature, surface reflectances, and so on, in order to comprehend what it means for a to be n times bigger than b. But with density the situation is more complicated—take n objects of density ρ and their concatenation won't have density $\mathrm{n} \cdot \rho$, unless some mass is added or some volume is lost in the process. Instead, where mass (m) and volume (V) are represented in the usual way, the meaningfulness of density ratios can be explained by the fact that

$$\frac{\rho_i}{\rho_j} = \frac{m_i}{m_j} \cdot \frac{V_j}{V_i},$$

where the terms on the right-hand-side are all scale-independent. More generally, if you want to know why density ratios make sense then you're going to need to consider how density relates to mass and volume.

The lesson here is not that mass and volume are *primitive* quantities while density is *derived*, as if this distinction somehow reflects a necessary conceptual or ontological hierarchy between them. We could just as well treat density as a primitive while treating mass or volume, or both, as the derived quantities; the distinction just corresponds to the order in which we choose to derive certain quantities from others in a dimensionally coherent system of units. Rather, what matters is the structure of the relations between the three quantities—namely, the fact that fixing any two of them determines the third; and the fact that the connection between them can be represented in product/quotient form so as to imply the scale-independence of density ratios. Along similar lines, confidence ratios can inherit meaning from the systematic way that confidence and preference relate in a decision-theoretic context. For this explanation to work, we do not have to imagine that confidence is somehow conceptually or ontologically subordinate to preference—realism about confidence simply doesn't mean we have to ignore the obvious interaction with preferences when it comes to explaining meaningful confidence ratios!

Theorem 1 shows that there are cases in which preferences uniquely determine a unique probability function even though the associated comparative confidence structure underdetermines it. In such cases, the probability ratios carry meaningful information that cannot be explained purely by the structure of the agent's confidence ordering (and natural strengthenings thereof). Instead, the explanation must appeal to the way confidence and preference jointly constrain the structure of expected-utility representation. This is precisely the point of Preference-Grounded Ratios. Under suitable conditions, degrees of confidence admit a ratio-scale representation, and the explanation for this fact lies not solely in the internal structure of confidence itself but in its systematic interaction with preference. Decision-theoretic representation theorems therefore do more than recover probabilities from preferences: they help reveal the qualitative psychological structure that makes quantitative confidence meaningful in the first place.

Although the argument of this paper has been formulated solely within the context of SEUT, essentially similar explanations can arise across a wide variety of mainstream decision theories—namely, those under which an act's value is obtained by integrating a utility (or value) function against multiplicative weights that are derived in whole or in part from the decision-maker's numerically-represented degrees of confidence, possibly following some distortion or ranking. The benefit of appealing to expected utility theory is that the relevant interactions are straightforward: double the confidence weight at a state and you double the contribution of the utilities of the consequences at that state to the value of the act. But in prospect theory, for example, or in risk-weighted utility, rank-dependent expected utility, Choquet expected utility theory, and so on, there is still a systematic—albeit generally less direct—interaction between confidence and preference to which we can appeal in the explanation of meaningful confidence ratios.

Appendix

The proof of Theorem 1 proceeds as follows. First we use A1–A5 to show that binary acts on half-probability events behave as expected given the intended representation (Lemma 1). Along with A6, this lets us construct an equally-spaced utility function that represents $\succsim$ over $\mathcal{C}$ and is unique up to a positive affine transformation (Lemma 2). We then add A7 and A8 to show that $\succsim$ maximizes EU relative to at least one probability function (Lemma 3). Given the facts about expected utility maximization generally established in Section 5, Theorem 1 then follows readily.

Lemma 1. *If A1–A5 are satisfied, then for all consequences $\alpha, \beta, \gamma, \delta$ and all half-probability events H, H':*

(a) $\alpha H \beta \sim \beta H' \alpha$

(b) *If $\succsim$ is non-trivial then H is non-null*

(c) $\Delta(\alpha, \gamma) = \Delta(\delta, \beta) \Leftrightarrow \alpha H \beta \sim \gamma H \delta$

(d) *If $\alpha H \beta \succ \gamma H \delta$ then $\Delta(\alpha, \gamma) > \Delta(\delta, \beta)$*

Proof. (a) Consider the following decision matrix:

	$H \cap H'$	$H \cap \overline{H'}$	$\overline{H} \cap H'$	$\overline{H} \cap \overline{H'}$
$\alpha H \beta$	α	α	β	β
$\alpha H' \beta$	α	β	α	β
$\beta H \alpha$	β	β	α	α
$\beta H' \alpha$	β	α	β	α

Note that, since the consequences of $\alpha H \beta$ and $\alpha H' \beta$ are identical under the event $H \cap H'$, A5 therefore implies $\alpha H \beta \succsim \alpha H' \beta \Leftrightarrow \beta H' \alpha \succsim \beta H \alpha$. By completeness, either $\alpha H \beta \succsim \alpha H' \beta$ or $\alpha H' \beta \succsim \alpha H \beta$. Suppose it's the former, and assume (without loss of generality) that $\alpha \succsim \beta$. We know that $\alpha H \beta \sim \beta H \alpha$ for all half-probability events H, since $H \sim_c \overline{H}$ entails $\alpha H \beta \sim \alpha \overline{H} \beta = \beta H \alpha$ under A2. As such, A5 implies $\alpha H \beta \succsim \alpha H' \beta \sim \beta H' \alpha \succsim \beta H \alpha$, hence $\alpha H \beta \sim \beta H' \alpha$ follows by A1. If it's the latter then (for the same reasons) we get $\alpha H \beta \sim \beta H' \alpha$. So either way, $\alpha H \beta \sim \beta H' \alpha$.

(b) Suppose that $\succsim$ is non-trivial, with $\alpha \succ \beta$. By A1, if $\alpha H \beta \sim \alpha$ then $\alpha H \beta \nsim \beta$, whereas if $\alpha H \beta \sim \beta$ then $\beta H \alpha \nsim \alpha$. So at least one of $\alpha H \beta \nsim \alpha$ or $\alpha H \beta \nsim \beta$ holds, and either way H is non-null.

(c) For the left-to-right direction, suppose that $\Delta(\alpha, \gamma) = \Delta(\delta, \beta)$. Then from A5, $\alpha H \beta \sim \delta H \gamma \Leftrightarrow \gamma H \beta \sim \beta H \gamma$. The right-hand-side is guaranteed by Lemma 1(a), so $\Delta(\alpha, \gamma) = \Delta(\delta, \beta)$ implies $\alpha H \beta \sim \delta H \gamma \sim \gamma H \delta$. The right-to-left direction is immediate from A4 plus Lemma 1(a).

(d) Observe that, from A4 and Lemma 1(c), for any consequences α, β, there are (not necessarily distinct) consequences γ, δ such that $\alpha H \beta \sim \gamma H \delta$ and there's no further ϵ where $\gamma \succ \epsilon \succ \delta$. There are two possibilities: either $\Delta(\alpha, \beta)$ is even, in which case there is a $\gamma = \delta$ such that $\Delta(\alpha, \gamma) = \Delta(\gamma, \beta)$ and therefore $\alpha H\beta \sim \gamma H\gamma$; or $\Delta(\alpha, \beta)$ is odd, in which case there will be adjacent γ, δ such that $\Delta(\alpha, \gamma) = \Delta(\delta, \beta)$, and therefore $\alpha H\beta \sim \gamma H\delta$. In either case, there will be the same number of consequences between α and $\alpha H\beta$ as there are between $\alpha H\beta$ and β. Given this, assume $\alpha H\beta \succ \gamma H\delta$. There are three ways this might arise:

(i) $\alpha \succ \gamma$ and $\beta \succ \delta$
(ii) $\alpha \succ \gamma$, $\delta \succsim \beta$, and $\Delta(\alpha, \gamma) > \Delta(\delta, \beta)$
(iii) $\gamma \succsim \alpha$, $\beta \succ \delta$ and $\Delta(\gamma, \alpha) < \Delta(\beta, \delta)$

In all three circumstances it follows that $\Delta(\alpha, \gamma) > \Delta(\delta, \beta)$. □

Lemma 2. *If A1–A6 are satisfied and $\succsim$ is non-trivial, then there is a utility function u such that for all consequences $\alpha, \beta, \gamma, \delta$ and all half-probability events H,*

U1. $\alpha \succsim \beta \Leftrightarrow u(\alpha) \geq u(\beta)$
U2. $\alpha H\beta \succsim \gamma H\delta \Leftrightarrow u(\alpha) - u(\gamma) \geq u(\delta) - u(\beta)$
U3. $\Delta(\alpha, \beta) = \Delta(\gamma, \delta) \Leftrightarrow u(\alpha) - u(\beta) = u(\gamma) - u(\delta)$

Furthermore,

U4. *If any other utility function u' satisfies U1–U3, then there are constants* $\mathrm{a} > 0$ *and* b *such that* $u' = \mathrm{a} \cdot u + \mathrm{b}$

Proof. Assume that A1–A6 hold, and that $\succsim$ is non-trivial. Let $\alpha\beta \succsim_h \gamma\delta$ mean that $\alpha H\beta \succsim \gamma H\delta$ for some half-probability event H. Given Lemma 1(a), if $\alpha H\beta \succsim \gamma H\delta$ is true for some half-probability event H then it's true for all, so it doesn't matter which H we choose. We start by showing that $\langle \mathcal{C} \times \mathcal{C}, \succsim_h \rangle$ is an additive conjoint structure. The following definition is modified from (Krantz et al. 1971: 256), since finiteness obviates the usual need for an Archimedean condition:

Definition 1. Suppose that $\mathbf{X}$ and $\mathbf{Y}$ are non-empty finite sets and that $\succsim$ is a binary relation on $\mathbf{X} \times \mathbf{Y}$. Then $\langle \mathbf{X} \times \mathbf{Y}, \succsim \rangle$ is an *additive conjoint structure* just in case $\succsim$ satisfies the following:

B1. $\succsim$ is transitive and connected
B2. For all $x, x', x'' \in \mathbf{X}$ and all $y, y', y'' \in \mathbf{Y}$, $xy'' \sim x''y'$ and $x''y \sim x'y''$ implies $xy \sim x'y'$
B3. (a) For all $x, x' \in \mathbf{X}$, $xy \succsim x'y$ for any $y \in \mathbf{Y}$ implies $xy' \succsim x'y'$ for all $y' \in \mathbf{Y}$
(b) For all $y, y' \in \mathbf{Y}$, $xy \succsim xy'$ for any $x \in \mathbf{X}$ implies $x'y \succsim x'y'$ for all $x' \in \mathbf{X}$
B4. (a) There are $x, x' \in \mathbf{X}$ and $y \in \mathbf{Y}$ such that $xy \not\sim x'y$
(b) There are $y, y' \in \mathbf{Y}$ and $x \in \mathbf{X}$ such that $xy \not\sim xy'$
B5. (a) If there are $x, x', x'' \in \mathbf{X}$ and $y, y' \in \mathbf{Y}$ for which $x'y' \succsim xy \succsim x''y'$, then there is an $x''' \in \mathbf{X}$ such that $x'''y' \sim xy$
(b) If there are $x, x' \in \mathbf{X}$ and $y, y', y'' \in \mathbf{Y}$ for which $x'y' \succsim xy \succsim x'y''$, then there is a $y''' \in \mathbf{Y}$ such that $x'y''' \sim xy$

That $\langle \mathcal{C}\times\mathcal{C}, \succsim_h \rangle$ satisfies B1 follows immediately from A1. For B2, assume $\alpha H\gamma' \sim \gamma H\beta'$ and $\gamma H\alpha' \sim \beta H\gamma'$. By A4 + Lemma 1(a) these imply $\Delta(\alpha, \beta') = \Delta(\beta, \alpha')$; then from Lemma 1(c) we get $\alpha H\alpha' \sim \beta H\beta'$. B3(a) follows immediately from A5. B4(a) follows immediately from A6 + Lemma 1(b), given that $\succsim$ is non-trivial. For B5(a), assume $\gamma' H\delta \succ \alpha H\beta \succ \gamma'' H\delta$. (If either $\succ$ is replaced by $\sim$ then the conclusion is immediate.) By Lemma 1(d) this implies $\Delta(\gamma', \alpha) > \Delta(\beta, \delta) > \Delta(\gamma'', \alpha)$. There are two cases: (i) $\Delta(\beta, \delta) \geq 0$, and (ii) $\Delta(\beta, \delta) < 0$. For case (i), $\Delta(\gamma', \alpha) = \Delta(\beta, \delta) + \mathrm{n}$ for some natural number n. So let ϵ be any consequence such that $\Delta(\gamma', \epsilon) = \mathrm{n}$; then $\Delta(\epsilon, \alpha) = \Delta(\beta, \delta)$, which implies $\epsilon H\delta \sim \alpha H\beta$ under Lemma 1(c). The reasoning is more or less the same for case (ii). Finally, Lemma 1(a) implies that $\alpha\beta \sim_h \beta\alpha$ for all α, β, so the (b) parts of B3–B5 are each entailed by the respective (a) parts.

Since $\langle \mathcal{C}\times\mathcal{C}, \succsim_{\mathrm{h}} \rangle$ is an additive conjoint structure, we can use the following:

Theorem 2 (Krantz et al. 1971: 257)**.** *Suppose that $\langle \mathbf{X}\times\mathbf{Y}, \succsim \rangle$ is an additive conjoint structure. Then there exists a pair of functions $u_x : \mathbf{X} \to \mathbb{R}$ and $u_y : \mathbf{Y} \to \mathbb{R}$ such that, for all $x, x' \in \mathbf{X}$ and $y, y' \in \mathbf{Y}$,*

$$xy \succsim x'y' \;\Leftrightarrow\; u_x(x) + u_y(y) \geq u_x(x') + u_y(y')$$

Furthermore, if u'_x and u'_y are two other functions with the same property, then there exist constants $\mathrm{a} > 0$, b_1 and b_2 such that $u'_x = \mathrm{a}u_x + \mathrm{b}_1$ and $u'_y = \mathrm{a}u_y + \mathrm{b}_2$.

Let $u_x : \mathcal{C} \to \mathbb{R}$ and $u_y : \mathcal{C} \to \mathbb{R}$ be any pair of functions as specified in Theorem 2. Since $\alpha H\beta \sim \beta H\alpha$ for all consequences α, β, therefore

$$u_x(\alpha) + u_y(\beta) = u_x(\beta) + u_y(\alpha),$$

which is equivalent to

$$u_x(\alpha) - u_y(\alpha) = u_x(\beta) - u_y(\beta)$$

Thus $u_x = u_y + \mathrm{b}$ for some constant b. In light of this, Lemma 1(a) plus Theorem 2's uniqueness condition (i.e., with $\mathrm{a} = 1$, $\mathrm{b}_1 = 0$, $\mathrm{b}_2 = \mathrm{b}$) imply that there is a utility function u such that for all half-probability events H,

$$\alpha H\beta \succsim \gamma H\delta \;\Leftrightarrow\; u(\alpha) + u(\beta) \geq u(\gamma) + u(\delta) \;\Leftrightarrow\; u(\alpha) - u(\gamma) \geq u(\delta) - u(\beta)$$

So u satisfies U2. For U1, simply note:

$$\alpha \succsim \beta \;\Leftrightarrow\; \alpha H\alpha \succsim \beta H\beta \;\Leftrightarrow\; u(\alpha) + u(\alpha) \geq u(\beta) + u(\beta) \;\Leftrightarrow\; u(\alpha) \geq u(\beta)$$

The equal spacing property U3 follows from U2 plus Lemma 1(c). Finally, U4 follows straightforwardly from the uniqueness condition in Theorem 2. □

Lemma 3. *If A1–A8 hold, then there is a probability function p such that $\succsim$ maximizes EU relative to p.*

Proof. Assume A1–A8. Henceforth, let ω^0 designate a consequence that's minimal in $\succsim$ over $\mathcal{C}$ (i.e., $\alpha \succsim \omega$ for all $\alpha \in \mathcal{C}$); and let ω^i designate a consequence such that $\Delta(\omega^i, \omega^0) = i$. From A8 plus A5, there will exist a non-null state (call it $s_{\min}$) such that for every non-null state s there is a consequence (call it β_s) where $\beta_s \succsim \omega^1$ and $\omega^1\{s\}\omega^0 \sim \beta_s\{s_{\min}\}\omega^0$. Additionally, if there's any consequence $\gamma \neq \beta_s$ such that $\omega^1\{s\}\omega^0 \sim \gamma\{s_{\min}\}\omega^0$, then it follows from A7 that $\beta_s\{s\}\omega^0 \sim \gamma\{s\}\omega^0$, which by A6 implies $\gamma \sim \beta_s$ and so $\Delta(\beta_s, \omega^0) = \Delta(\gamma, \omega^0)$.

Given those facts, let p be the unique normalized function on $2^{\mathcal{S}}$ satisfying both

$$p(E) = \sum_{s \in E} p(s)$$

and

$$p(s) = \begin{cases} 0 & \text{if } s \text{ is null,} \\ \Delta(\beta_s, \omega^0) \cdot p(s_{\min}) & \text{if } s \text{ is non-null} \end{cases}$$

Observe that, since $\Delta(\beta_s, \omega^0)$ is always positive, so $p(s_{\min})$ must be positive and p is non-negative, hence a probability function. Furthermore, let u be the unique utility function satisfying U1–U3 such that $u(\omega^i) = i$. Now p and u jointly determine a preference ordering (call it $\succsim^\star$) under the expected utility rule. We need to show the following:

C1. If $f \sim^\star g$, then $f \sim g$
C2. If $f \succ^\star g$ and there's no h such that $f \succ^\star h \succ^\star g$, then $f \succ g$

Since C1 and C2 together imply that $\succsim = \succsim^\star$, it will follow that $\succsim$ maximizes EU relative to p.

For C1, the goal will be to show that if $f \sim^\star g$, then $f \sim g$ follows by a sequence of indifferences justified by A5. To that end, it will be helpful to start by reconstructing the act space as a set of integer-valued vectors. Say that acts f and g are *essentially identical* (written $f \simeq g$) just in case $f(s) \sim g(s)$ for all non-null states s. Let $\mathbf{A}$ be the partition of $\mathcal{C}^{\mathcal{S}}$ into $\simeq$-equivalence classes, so $\mathbf{f} \in \mathbf{A}$ just in case $\mathbf{f} = \{g \in \mathcal{C}^{\mathcal{S}} \mid g \simeq f\}$ for some act f. Note that the definition of $\simeq$ doesn't depend on the choice of $\succsim$ or $\succsim^\star$, since s is null just in case $p(s) = 0$ and $\alpha \sim \beta \Leftrightarrow u(\alpha) = u(\beta) \Leftrightarrow \alpha \sim^\star \beta$ for all consequences α, β. For the same reason, $f \simeq g$ implies $f \sim g$ and $f \sim^\star g$. Furthermore, let $\widetilde{\mathcal{C}}$ be the set of $\sim$-equivalence classes in $\mathcal{C}$, and let $\mathcal{N} \subset \mathcal{S}$ be the set of null states. Since A3 implies there are half-probability events and A8 implies that $\succsim$ is non-trivial, it follows that $\mathcal{S} - \mathcal{N} = \{s_1 \ldots, s_{\mathrm{n}}\}$ for some $\mathrm{n} \geq 2$ and $|\widetilde{\mathcal{C}}| = \mathrm{m} + 1$ for some $\mathrm{m} \geq 1$. The upshot is that $\mathbf{A}$ is isomorphic to (and will henceforth be identified with) the discrete n-dimensional box $\{0, 1, \ldots, \mathrm{m}\}^{\mathrm{n}}$. That is, every act f can be represented by an n-tuple (or vector) $\mathbf{f}$ taking integer values between 0 and m inclusive, where the i^{th} coordinate expresses the utility of $f(s_i)$ under u. Two acts correspond to the same vector $\mathbf{f}$ just in case they're essentially identical, in which case they're indifferent under both $\succsim$ and $\succsim^\star$.

Next we characterize the set of vector differences, $\mathbf{D} \subseteq \{-\mathrm{m}, \ldots, \mathrm{m}\}^{\mathrm{n}}$, such that $f \sim^{\star} g$ just in case $\mathbf{f} - \mathbf{g} \in \mathbf{D}$. First, where $\mathbb{Q}^{\mathrm{n}}$ is the vector space of functions from $\{1, \ldots, \mathrm{n}\}$ to $\mathbb{Q}$, define the linear map $eu : \mathbb{Q}^{\mathrm{n}} \to \mathbb{Q}$ such that for all $\mathbf{v} \in \mathbb{Q}^{\mathrm{n}}$,

$$eu(\mathbf{v}) = \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot p(s_i)$$

The integer kernel of eu is then the lattice

$$\mathbf{K} = \ker(eu) \cap \mathbb{Z}^{\mathrm{n}} = \left\{ \mathbf{v} : \{1, \ldots, \mathrm{n}\} \to \mathbb{Z} \,\middle|\, \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot p(s_i) = 0 \right\}$$

Henceforth, assume for notational convenience that $s_{\mathrm{n}} = s_{\min}$, and let b_i be the positive integer ($\leq$ m) such that $p(s_i) = b_i \cdot p(s_{\min})$. Then, for each non-null $s_i \neq s_{\min}$, define the *basic exchange vector* $\mathbf{b}^i_{\min}$ like so:

$$\mathbf{b}^i_{\min}(s) = \begin{cases} b_i & \text{if } s = s_{\min} \\ -1 & \text{if } s = s_i \\ 0 & \text{otherwise} \end{cases}$$

Every basic exchange vector lies in $\mathbf{K}$, since

$$eu(\mathbf{b}^i_{\min}) \;=\; b_i \cdot p(s_{\min}) + (-1)p(s_i) \;=\; \frac{p(s_i)}{p(s_{\min})} \cdot p(s_{\min}) - p(s_i) \;=\; 0$$

The same is true for every integer linear combination of basic exchange vectors, by the linearity of eu. Furthermore, every vector in $\mathbf{K}$ can be written as an integer linear combination of basic exchange vectors. Since $p(s_{\min}) > 0$,

$$\sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot p(s_i) = 0 \;\Leftrightarrow\; p(s_{\min}) \cdot \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot b_i = 0 \;\Leftrightarrow\; \sum_{i=1}^{\mathrm{n}} \mathbf{v}(i) \cdot b_i = 0$$

And since $b_{\mathrm{n}} = 1$, isolating $\mathbf{v}(\mathrm{n}) \cdot b_{\mathrm{n}}$ from the final equality gets us:

$$\mathbf{v}(\mathrm{n}) \;=\; -\sum_{i=1}^{\mathrm{n}-1} \mathbf{v}(i) \cdot b_i$$

Thus every kernel vector $\mathbf{v}$ takes the form:

$$\mathbf{v} = \sum_{i=1}^{\mathrm{n}-1} \big(-\mathbf{v}(i)\big) \cdot \mathbf{b}^i_{\min} = \left(\mathbf{v}(1), \ldots, \mathbf{v}(\mathrm{n}-1), -\sum_{i=1}^{\mathrm{n}-1} \mathbf{v}(i) \cdot b_i \right)$$

Now given $\mathbf{v}(i) \in \mathbb{Z}$ whenever $\mathbf{v} \in \mathbf{K}$, so $\mathbf{K}$ consists of the integer linear combinations of the $\mathbf{b}^i_{\min}$. Finally, $\mathbf{D}$ will be the intersection of $\mathbf{K}$ with those difference vectors that can in principle be realized from the elements of $\mathbf{A}$:

$$\mathbf{D} = \left\{ \sum_{i=1}^{\mathrm{n}-1} \mathrm{a}_i \cdot \mathbf{b}^i_{\min} \,\middle|\, \mathrm{a}_i \in \mathbb{Z} \right\} \cap \{-\mathrm{m}, \ldots, \mathrm{m}\}^{\mathrm{n}}$$

We will also make use of the following observations. The first is an equivalent way to express A5: two ordered pairs $(\mathbf{f},\mathbf{g})$ and $(\mathbf{f}',\mathbf{g}')$ in $\mathbf{A}\times\mathbf{A}$ realize the same difference vector only if preferences between the acts they represent are matched. That is, for all $\mathbf{f},\mathbf{g},\mathbf{f}',\mathbf{g}' \in \mathbf{A}$,

O1. $\mathbf{f}-\mathbf{g}=\mathbf{f}'-\mathbf{g}'$ implies $f \succsim g \Leftrightarrow f' \succsim g'$

Next, for any vector $\mathbf{v}$, let $\mathbf{v}^{(+)}$ and $\mathbf{v}^{(-)}$ denote the pointwise positive and negative components of $\mathbf{v}$ respectively; so

$$\mathbf{v}^{(+)}(i)=\max\{\mathbf{v}(i),0\}, \qquad \mathbf{v}^{(-)}(i)=\max\{-\mathbf{v}(i),0\}$$

If $\mathbf{f},\mathbf{g}\in\mathbf{A}$ and $\mathbf{f}-\mathbf{g}=\mathbf{h}$, then $-\mathrm{m}\le \mathbf{h}(i)\le \mathrm{m}$ and so $\mathbf{h}^{(+)}$ and $\mathbf{h}^{(-)}$ also belong to $\mathbf{A}$. Moreover, $\mathbf{f}-\mathbf{g}=\mathbf{h}^{(+)}-\mathbf{h}^{(-)}$. Since $h^{(+)}$ and $h^{(-)}$ have the lowest utility values for each state consistent with the vector $\mathbf{f}-\mathbf{g}$, $(\mathbf{h}^{(+)},\mathbf{h}^{(-)})$ is the *minimal realizer* of the difference vector realized by $(\mathbf{f},\mathbf{g})$. By O1, the preference ordering between acts f and g matches the ordering of those acts represented by the minimal realizer of $\mathbf{f}-\mathbf{g}$. That is, for all $\mathbf{f},\mathbf{g}\in\mathbf{A}$,

O2. $\mathbf{f}-\mathbf{g}=\mathbf{h}$ implies $f \succsim g \Leftrightarrow h^{(+)} \succsim h^{(-)}$

Recall also that for every non-null state s there's a consequence β_s where $\beta_s \succsim \omega^1$ and $\omega^1\{s\}\omega^0 \sim \beta_s\{s_{\min}\}\omega^0$. Since $u(\beta_{s_i})=b_i$, so

$$\beta_{s_i}\{s_{\min}\}\omega^0 \in \mathbf{b}^{i\,(+)}_{\min}, \qquad \omega^1\{s_i\}\omega^0 \in \mathbf{b}^{i\,(-)}_{\min}$$

From O2,

$$\mathbf{f}-\mathbf{g}=\mathbf{b}^i_{\min} \;\Rightarrow\; (f\sim g \;\Leftrightarrow\; \beta_s\{s_{\min}\}\omega^0 \sim \omega^1\{s\}\omega^0),$$

in which case the known indifference $\omega^1\{s\}\omega^0 \sim \beta_s\{s_{\min}\}\omega^0$ implies $f\sim g$. The same holds if $\mathbf{f}-\mathbf{g}=-\mathbf{b}^i_{\min}$. Rearranging terms gets us, for all $\mathbf{f}\in\mathbf{A}$,

O3. $\mathbf{f}\pm\mathbf{b}^i_{\min}=\mathbf{g}$ implies $f\sim g$ whenever $\mathbf{g}\in\mathbf{A}$

Finally, suppose that $f\succsim g$ and that there are f',g' such that $\mathbf{f}-\mathbf{g}=\mathbf{f}'-\mathbf{g}'$ and $\mathbf{f}'\pm\mathbf{b}^i_{\min}=\mathbf{h}$ for some $\mathbf{h}\in\mathbf{A}$. By O1, $f'\succsim g'$; and by O3, $f'\sim h$; hence $h\succsim g'$. Now observe that

$$\mathbf{h}-\mathbf{g}'=\begin{cases} (\mathbf{f}+\mathbf{b}^{i\,(+)}_{\min})-(\mathbf{g}+\mathbf{b}^{i\,(-)}_{\min}) & \text{if } \mathbf{f}'+\mathbf{b}^i_{\min}=\mathbf{h} \\ (\mathbf{f}+\mathbf{b}^{i\,(-)}_{\min})-(\mathbf{g}+\mathbf{b}^{i\,(+)}_{\min}) & \text{if } \mathbf{f}'-\mathbf{b}^i_{\min}=\mathbf{h} \end{cases}$$

Consequently, for all $\mathbf{f},\mathbf{g}\in\mathbf{A}$,

O4. $\mathbf{f}+\mathbf{b}^{i\,(+)}_{\min}=\mathbf{h}$ and $\mathbf{g}+\mathbf{b}^{i\,(-)}_{\min}=\mathbf{h}'$, or $\mathbf{f}+\mathbf{b}^{i\,(-)}_{\min}=\mathbf{h}$ and $\mathbf{g}+\mathbf{b}^{i\,(+)}_{\min}=\mathbf{h}'$, implies $f\succsim g \Leftrightarrow h\succsim h'$ whenever $\mathbf{h},\mathbf{h}'\in\mathbf{A}$

Putting all that together: if $f \sim^\star g$, then $\mathbf{f} - \mathbf{g} = \mathbf{h}$ for some $\mathbf{h} \in \mathbf{D}$, and so $f \sim g \Leftrightarrow h^{(+)} \sim h^{(-)}$. To establish C1 it therefore suffices to show that $h^{(+)} \sim h^{(-)}$. Keeping in mind the earlier stipulation that $s_n = s_{\min}$, recall that $\mathbf{b}^i_{\min}(i) = -1$ and $\mathbf{b}^i_{\min}(j) = 0$ whenever $i \neq j \in \{1, \ldots, \mathrm{n}-1\}$. As such, let $\mathbf{h}_{[n-1]}$ be the restriction of $\mathbf{h}$ to the first $n-1$ coordinates, so

$$\mathbf{h} = (x_1, \ldots, x_{n-1}, x_n) \;\Rightarrow\; \mathbf{h}_{[n-1]} = (x_1, \ldots, x_{n-1})$$

Then $\mathbf{h}_{[n-1]}(i)$ expresses the number of times ($\leq m$) the corresponding basic exchange vectors $\mathbf{b}^i_{\min}$ must be added (where $\mathbf{h}(i) > 0$) or subtracted (where $\mathbf{h}(i) < 0$) from $\mathbf{h}$ to obtain the zero vector (designated $\mathbf{0}$). For example, if $\mathbf{h} = (-2, 3, 4)$, then $\mathbf{h}_{[n-1]} = (-2, 3)$; and if $\mathbf{h} \in \mathbf{D}$, it follows that

$$\mathbf{h} - 2\mathbf{b}^1_{\min} + 3\mathbf{b}^2_{\min} = \mathbf{0}$$

In other words, we obtain $\mathbf{0}$ by taking $\mathbf{b}^1_{\min}$ from $\mathbf{h}$ twice and adding $\mathbf{b}^2_{\min}$ three times. We can then show that $h^{(+)} \sim h^{(-)} \Leftrightarrow f \sim f$ by chaining together indifferences justified by the foregoing observations. This amounts to a finite 'walk' through the space $\mathbf{A}\times\mathbf{A}$, from $(\mathbf{h}^{(+)}, \mathbf{h}^{(-)})$ to $(\mathbf{0}, \mathbf{0})$, with steps corresponding to the addition or subtraction of a basic exchange vector to/from $\mathbf{h}$.

Given the earlier decomposition of $\mathbf{h} \in \Delta$ into a sum of basic exchange vectors,

$$\mathbf{h} = \sum_{i=1}^{n-1} a_i \cdot \mathbf{b}^i_{\min}, \quad a_i \in \mathbb{Z},$$

set the *positive* and *negative exchange multisets* for $\mathbf{h}$ like so:[8]

$$\mathrm{Pos}(\mathbf{h}) \;=\; \{i^{\langle a_i \rangle} \mid a_i > 0\}, \qquad \mathrm{Neg}(\mathbf{h}) \;=\; \{i^{\langle |a_i| \rangle} \mid a_i < 0\}$$

For any pair $(\mathbf{f}, \mathbf{g})$, call $\mathbf{f}$ the *top vector* and $\mathbf{g}$ the *bottom vector*. We take a *$+i$-step* from $(\mathbf{f}, \mathbf{g})$ when we add ${\mathbf{b}^i_{\min}}^{(+)}$ to the top vector and add ${\mathbf{b}^i_{\min}}^{(-)}$ to the bottom vector, arriving at $(\mathbf{f} + {\mathbf{b}^i_{\min}}^{(+)}, \mathbf{g} + {\mathbf{b}^i_{\min}}^{(-)})$. Similarly, we take a *$-i$-step* from $(\mathbf{f}, \mathbf{g})$ when we add ${\mathbf{b}^i_{\min}}^{(-)}$ to the top and ${\mathbf{b}^i_{\min}}^{(+)}$ to the bottom vectors. A $+i$-step (or $-i$-step) is *available* just in case there is an i in $\mathrm{Pos}(\mathbf{h})$ (or $\mathrm{Neg}(\mathbf{h})$, respectively), and with each such step we eliminate an i from that multiset.[9] A $\pm i$-step *stays in bounds* whenever the coordinates of the resulting vectors stay within $\{0, \ldots, m\}$, and it's *valid* just in case it's available and stays in bounds. Finally, we take a *reduction step* (or r-step) by shifting from $(\mathbf{f}, \mathbf{g})$ to $((\mathbf{f} - \mathbf{g})^{(+)}, (\mathbf{f} - \mathbf{g})^{(-)})$—i.e., reducing to the minimal realizer of the corresponding difference vector.

[8] Here, $i^{\langle n \rangle}$ denotes a multiset (also known as a *bag*) comprising n copies of i. For example, $7^{\langle 3 \rangle} = \{7, 7, 7\}$ and $2^{\langle 3 \rangle} 3^{\langle 2 \rangle} = \{2, 2, 2, 3, 3\}$; and if $\mathbf{h}_{[n-1]} = (4, -3, 1)$, then $\mathrm{Pos}(\mathbf{h}) = \{1, 1, 1, 1, 3\}$ and $\mathrm{Neg}(\mathbf{h}) = \{2, 2, 2\}$.

[9] For example, if $\mathbf{h}_{[n-1]} = (-2, 3)$, then at the start of the walk a total of two -1-steps and three $+2$-steps will be available; and if a single $+2$-step is taken, then two -1-steps and two $+2$-steps will remain available.

By O2 and O4, if we start at any point $(\mathbf{f}, \mathbf{g})$ in $\mathbf{A}\times\mathbf{A}$ and then take any valid $\pm i$-step followed immediately by an r-step, we will land on a pair $(\mathbf{f}', \mathbf{g}')$ such that

$$\mathbf{f}' - \mathbf{g}' = (\mathbf{f} - \mathbf{g}) \pm \mathbf{b}^i_{\min} \in \mathbf{A},$$

and therefore $f \sim g \Leftrightarrow f' \sim g'$. Consequently, starting at $(\mathbf{h}^{(+)}, \mathbf{h}^{(-)})$ for any $\mathbf{h} \in \mathbf{D}$, we apply the following procedure:

S1. If there are any available $\pm i$-steps in $\mathrm{Pos}(\mathbf{h})$ or $\mathrm{Neg}(\mathbf{h})$, then go to S2; otherwise, terminate procedure.
S2. If there are no valid $+i$-steps in $\mathrm{Pos}(\mathbf{h})$, then move to S3; otherwise, take any valid $+i$-step followed by an r-step, then go to S1.
S3. If there are no valid $-i$-steps $\mathrm{Neg}(\mathbf{h})$, then move to S1; otherwise, take any valid $-i$-step followed by an r-step, then go to S1.

By the points above, if all available $\pm i$-steps are taken then the procedure will terminate at $(\mathbf{0}, \mathbf{0})$, with every $\pm i$- and r-step along the way landing on some $(\mathbf{f}, \mathbf{g}) \in \mathbf{A}\times\mathbf{A}$ such that $h^{(+)} \sim h^{(-)} \Leftrightarrow f \sim g$. So to complete the proof of C1 we just need to show that all available $\pm i$-steps will indeed be taken—that is,

(i) following any r-step, if there are no valid $+i$-steps but there are some available $-i$-steps, then at least one $-i$-step stays in bounds
(ii) following any r-step, if there are no valid $-i$-steps but there are some available $+i$-steps, then at least one $+i$-step stays in bounds

We focus on (i), since the case for (ii) is symmetric. Consider to start with the first $n-1$ coordinates. There are $\mathbf{h}^{(-)}(i)$ $-i$-steps available at the beginning of the walk, and $\mathbf{h}^{(+)}(i) = 0$ whenever $\mathbf{h}^{(-)}(i) > 0$. Furthermore, a $-i$-step adds 1 to the i^{th} coordinate of the top vector, otherwise leaving the first $n-1$ coordinates of both the top and bottom vectors untouched. So the net effect of a $-i$-step followed by an r-step on those coordinates is a reduction by 1 of the i^{th} coordinate of the bottom vector, leaving the i^{th} coordinate of the top vector at zero; and furthermore, taking $\mathbf{h}^{(-)}(i)$ $-i$-steps results in both top and bottom vectors taking value 0 at coordinate i. (The same is true for $+i$-steps, *mutatis mutandis.*) As such, an available $-i$-step never goes out of bounds with respect to the first $n-1$ coordinates. Restricting our attention to the n^{th} coordinate, then, observe that a $+i$-step always adds $1 \leq b_i \leq m$ to coordinate n at the top vector, and a negative j-step always adds $1 \leq b_j \leq m$ to coordinate n at the bottom vector. There are two possible circumstances where there are no valid $+i$-steps following an r-step:

1. There are no available $+i$-steps
2. There is an available $+i$-step but it goes out of bounds

If it's the latter, then the top vector must be positive at n, else a $+i$-step would not go out of bounds. If it's the former, and there are some negative j-steps available, then by the points above the top vector must be zero at the first $n-1$ coordinates and the bottom vector must be positive for at least one of those coordinates. Since every vector in the kernel of eu besides $\mathbf{0}$ must have non-zero positive and negative components, and since every step in the walk lands on a pair of vectors that realizes a kernel vector, it follows again that the top vector must be positive at n. So in both circumstances, the top vector is positive at n, in which case the bottom vector must be zero at n, and so taking any available $-i$-step will stay in bounds.

For C2, recall that for all non-null states s_i, $\mathbf{f}(i) = u(f(s_i))$ and $p(s_i) = b_i \cdot p(s_{\min})$ for some integer $1 \leq b_i \leq m$. Given this,

$$eu(\mathbf{f}) \;=\; p(s_{\min}) \cdot \sum_{i=1}^{\mathrm{n}} b_i \cdot \mathbf{f}(i) \;=\; \sum_{s \in \mathcal{S}} p(s) \cdot u(f(s))$$

Since eu tracks $\succsim^\star$, we know that

$$f \succ^\star g \;\Leftrightarrow\; eu(\mathbf{f}) - eu(\mathbf{g}) \geq p(s_{\min})$$

Let $\mathbf{e}_{\min}$ be the vector assigning $+1$ for $s_{\min}$ and 0 otherwise. Thus $\mathbf{f} - \mathbf{g} = \mathbf{e}_{\min}$ when $eu(\mathbf{f}) - eu(\mathbf{g}) = p(s_{\min})$, in which case $f \succ^\star g$ and (moreover) there can be no h such that $f \succ^\star h \succ^\star g$. Furthermore, let g be any act that's not maximal in $\succsim^\star$. Then there is an act g' such that $g \sim^\star g'$ and $\mathbf{g'} + \mathbf{e}_{\min} = \mathbf{f}$ for some act f; namely,

(i) If $u(g(s_{\min})) < \mathrm{m}$, then $\mathbf{g'} = \mathbf{g}$ and $\mathbf{g'} + \mathbf{e}_{\min} \in \mathbf{A}$.
(ii) If $u(g(s_{\min})) = \mathrm{m}$, then $\mathbf{g} + \mathbf{e}_{\min} \notin \mathbf{A}$. But since g is non-maximal, there must be some non-null state $s_i \neq s_{\min}$ such that $u(g(s_i)) < \mathrm{m}$. So $\mathbf{g'} = \mathbf{g} - \mathbf{b}^i_{\min} \in \mathbf{A}$ and $\mathbf{g'} + \mathbf{e}_{\min} \in \mathbf{A}$.

As such, suppose that $f \succ^\star g$ and there's no h such that $f \succ^\star h \succ^\star g$. Then there will be acts f' and g' such that $f \sim^\star f'$, $g \sim^\star g'$ and $\mathbf{f} - \mathbf{g'} = \mathbf{e}_{\min} - \mathbf{0}$. By O2 and A6, $\mathbf{f} - \mathbf{g'} = \mathbf{e}_{\min} - \mathbf{0}$ implies $f \succ g'$; by C1, $f \sim f'$ and $g \sim g'$; therefore $f \succ g$. □

From here the proof of Theorem 1 is straightforward. From Lemma 3 we know that $\succsim$ maximizes EU relative to some probability function p. Given the points discussed in Section 5, we just need to show that $\succsim$ maximizes EU relative to c just in case $c = \mathrm{a} \cdot p$ for some non-zero constant a. The right-to-left direction is obvious. For the left-to-right, we note that if $c \neq \mathrm{a} \cdot p$ for any non-zero a, then c is either (i) constant ($\mathrm{a} = 0$); or, if it's non-constant, then (ii) it's non-additive or (iii) inconsistent with the basic exchange rates between states. Since all three cases can be ruled out, p must be the unique probability function p such that $\succsim$ maximizes EU relative to p; and then $E \succsim_c F \Leftrightarrow p(E) \geq p(F)$ follows immediately from A2.

References

Christensen, D. (2001). Preference-based arguments for probabilism. *Philosophy of Science 68*(3), 356–376.

Davidson, D. and P. Suppes (1956). A finitistic axiomatization of subjective probability and utility. *Econometrica 24*(3), 264–275.

de Finetti, B. (1931). Sul Significato Soggettivo Della Probabilita. *Fundamenta Mathematicae 17*(1), 298–329.

DiBella, N. (2018). Qualitative Probability and Infinitesimal Probability.

Dietrich, F. and C. List (2016). Mentalism versus behaviourism in economics: a philosophy-of-science perspective. *Economics and Philosophy 32*(2), 249–281.

Domotor, Z. (1969). Probabilistic Relational Structures and Their Applications.

Domotor, Z. (1970). Qualitative information and entropy structures. In J. Hintikka and P. Suppes (Eds.), *Information and inference*, pp. 148–194. Reidel.

Elliott, E. (2024). *The Measurement of Subjective Probability.* Cambridge University Press.

Eriksson, Lina and Hájek, Alan (2007). What are degrees of belief? *Studia Logica 86*(2), 183–213.

Fine, T. L. (1973). *Theories of Probability: An Examination of Foundations.* Academic Press.

Hausman, D. (2012). *Preference, Value, Choice, and Welfare.* Cambridge University Press.

Hájek, A. (2008). Arguments for–or against–Probabilism? *British Journal for the Philosophy of Science 59*(4), 793–819.

Joyce, J. (1999). *The foundations of causal decision theory.* Cambridge University Press.

Joyce, J. (2015). The Value of Truth: A Reply to Howson. *Analysis 75*(3), 413–424.

Kaplan, M. and T. Fine (1977). Joint orders in comparative probability. *The Annals of Probability 5*, 161–179.

Konek, J. (2019). Comparative Probabilities. In *The Open Handbook of Formal Epistemology*, pp. 267–348. PhilPapers Foundation.

Koopman, B. O. (1940). The Bases of Probability. *Bulletin of the American Mathematical Society 46*(10), 763–774.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971). *Foundations of Measurement, Vol. I: Additive and polynomial representations.* Academic Press.

Luce, D. R. (1968). On the numerical representation of qualitative conditional probability. *Annals of Mathematical Statistics 39*, 481–491.

Luce, R. D. (1978). Dimensionally Invariant Numerical Laws Correspond to Meaningful Qualitative Relations. *Philosophy of Science 45*(1), 1–16.

Luce, R. D., D. H. Krantz, P. Suppes, and A. Tversky (1990). *Foundations of Measurement, Vol. III: Representation, axiomatization, and invariance.* Dover.

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

Narens, L. (1985). *Abstract Measurement Theory.* MIT Press.

Pfanzagl, J. (1968). *Theory of Measurement.* Wiley.

Ramsey, F. P. (1931). Truth and probability. In R. B. Braithwaite (Ed.), *The Foundations of Mathematics and Other Logical Essays*, pp. 156–198. Routledge.

Savage, L. J. (1954). *The Foundations of Statistics.* Dover.

Stefánsson, H. O. (2018). On the Ratio Challenge for Comparativism. *Australasian Journal of Philosophy 96*(2), 380–390.

Suppes, P. (1994). Qualitative theory of subjective probability. In G. Wright and P. Ayton (Eds.), *Subjective Probability*, pp. 17–38. Wiley.

Suppes, P. (2014). Using Padoa's principle to prove the non-definability, in terms of each other, of the three fundamental qualitative concepts of comparative probability, independence and comparative uncertainty, with some new axioms of qualitative independence and uncertainty included. *Journal of Mathematical Psychology 60*, 47–57. (Yes that's the real title).

Suppes, P. and M. Zanotti (1982). Necessary and Sufficient Qualitative Axioms for Conditional Probability. *Z. Wahrschelnllchkeltstheorle verw. Gebiete 60*, 163–169.

Titelbaum, M. (2022). *Fundamentals of Bayesian Epistemology 2: Arguments, Challenges, Alternatives.* Oxford University Press.

Villegas, C. (1964). On Qualitative Probability σ-Algebras. *The Annals of Mathematical Statistics 35*(4), 1787–1796.

Zynda, L. (2000). Representation Theorems and Realism About Degrees of Belief. *Philosophy of Science 67*(1), 45–69.