A Behavioral Characterization of Plausible Priors

Marciano Siniscalchi*
Economics Department, Northwestern University,
302 Arthur Andersen Hall, 2001 Sheridan Rd., Evanston, IL 60208.
marciano@northwestern.edu

March 6, 2005

*An earlier draft of this paper was circulated under the title "Expected Utility with Many Unique Priors." I would like to thank Michèle Cohen, Eddie Dekel, Paolo Ghirardato, Faruk Gul, Alessandro Lizzeri, Fabio Maccheroni, Wolfgang Pesendorfer, and Pietro Veronesi, as well as several RUD 2002 participants, for helpful discussion. All errors are my own.

Abstract

Recent decision theories represent ambiguity via multiple priors, interpreted as alternative probabilistic models of the relevant uncertainty. This paper provides a robust behavioral foundation for this interpretation. A prior P is "plausible" if preferences over some subset of acts admit an expected utility representation with prior P, but not with any other prior $Q \neq P$. Under suitable axioms, plausible priors can be elicited from preferences, and fully characterize them; also, probabilistic sophistication implies that there exists only one plausible prior; finally, "plausible posteriors" can be derived via Bayesian updating. Several familiar decision models are consistent with the proposed axioms.

JEL classification: D81

Running Title: PLAUSIBLE PRIORS

Keywords: Ambiguity, Multiple priors, Maxmin Expected Utility

This paper has an **Online Appendix**: please visit
`http://faculty.econ.northwestern.edu/faculty/siniscalchi`.

1 Introduction

Multiple priors, or sets of probabilities over relevant states of nature, are a distinguishing feature of several decision models that depart from subjective expected utility maximization (SEU) in order to account for perceived ambiguity; examples include the maxmin expected-utility (MEU) model (Ellsberg [8]; Gilboa and Schmeidler [14]), and Hurwicz' α-maxmin expected utility criterion (α-MEU; see e.g. Luce and Raiffa [22], and Ghirardato, Maccheroni and Marinacci [11]). Also, the Choquet Expected Utility model (CEU; Schmeidler [32]) does not employ sets of priors, but admits a multiple-priors interpretation (see Section 2.4 below).

The literature suggests that sets of priors may reflect the decision-maker's subjective assessment of available information: if the latter is perceived as being sufficiently precise, the decision-maker's beliefs can be represented by a single probability distribution; but, if information is subjectively ambiguous, the decision-maker may wish to consider multiple possible probabilistic descriptions of the underlying uncertainty.[1]

This "intuitive" interpretation of sets of priors has played a central role in motivating both the early literature and more recent seminal contributions on choice under ambiguity (cf. footnote 1). It is also typically invoked in applications that adopt multiple-prior preferences. In fact, the economic intuition underlying modeling assumptions and formal results often builds directly upon this interpretation: see e.g. Mukerji [28, p. 1209], Hansen, Sargent and Tallarini [16, p. 878], or Billot, Chateauneuf, Gilboa and Tallon [4, p. 686]. Finally, the literature often also suggests that the decision criteria (functional forms) employed in specific multiple-priors models may be viewed as reflecting the decision-maker's attitudes towards ambiguity. For instance, one can associate the use of the "min" operator in the MEU model with "extreme" ambiguity aversion. The cited interpretation of priors is fully consistent with this view: sets of probabilities are assumed to describe the individual's perception of ambiguity, independently of how they are used in order to rank alternative prospects;[2] consequently, specific functional representations reflect behavioral traits other than "beliefs"—in particular, attitudes towards ambiguity.

However, existing axiomatic characterizations of *specific* decision models (e.g. [14], [32], etc.) do not provide full support for the "intuitive" interpretation of sets of probabilities and decision rules described above. Even in simple, canonical examples, *preference relations admit multiple different representations, involving different sets of priors and decision criteria.*

[1] See e.g. Ellsberg [8, pp. 657 and 661]), Gilboa and Schmeidler [14, p. 142] and Schmeidler [32, p. 584]; also cf. Luce and Raiffa [22, pp. 304-305], and Bewley [3].

[2]In particular, Ellsberg [8, especially pp. 661-665] and (to a lesser extent) Gilboa and Schmeidler [14, p. 142] suggest this interpretation.

For instance, a preference that satisfies the Gilboa-Schmeidler [14] axioms admits a MEU representation; however, as will be demonstrated below, the *same* preference typically also admits an α-MEU representation, and the sets of priors appearing in the two representations are different. Thus, additional considerations must be invoked in order to determine which of these sets, if any, comprises all possible probabilistic descriptions of the uncertainty, and hence which decision criterion reflects the decision-maker's attitudes towards ambiguity.

This paper proposes a definition of "plausible prior" that identifies probabilities directly from preferences, without relying on a specific functional representation. The proposed definition is thus fully "behavioral"; furthermore, it is independent of the individual's attitudes towards ambiguity.

The main result of this paper shows that, under suitable axioms on preferences, a unique collection of plausible priors can be identified. Moreover, under the same axioms:

- Plausible priors fully characterize preferences: there exists a unique decision criterion that determines the ordering of any two acts as a function of their expected utilities computed with respect to each plausible prior. Thus, the proposed definition yields a more robust foundation for the intuitive interpretation of multiple-prior models. Preferences that satisfy the proposed axioms will be deemed "plausible-priors preferences".

- The class of plausible-priors preferences is closed under Bayesian updating. Consider an "unconditional" plausible-priors preference, and a "conditional" preference that is related to the latter by means of suitable consistency conditions. Then conditional preferences also satisfy the axioms I propose; furthermore, the corresponding "plausible posteriors" are derived from plausible priors by Bayesian updating.

- Finally, under appropriate regularity conditions, a plausible-prior preference is probabilistically sophisticated in the sense of Machina and Schmeidler [24] *if and only if* it is consistent with SEU—and hence admits a single plausible prior. Thus, a multiplicity of plausible priors necessarily reflects the decision-maker's perception of ambiguity.

The axioms I consider are compatible with a variety of known decision models, reflecting a broad range of attitudes towards ambiguity: cf. Section 2.4 for examples.

A Motivating Example

In order to make the preceding discussion more concrete, consider Daniel Ellsberg's celebrated three-color-urn experiment ([8]). An urn contains 30 red balls, and 60 green and blue balls, in unspecified proportions; subjects are asked to rank bets on the realizations of a draw from the urn. Denote by r, g and b the possible realizations of the draw, in obvious notation. The following, typical pattern of preferences suggests that subjects dislike ambiguity about the relative likelihood of g vs. b: (\$10 if r, \$0 otherwise) is strictly preferred to (\$10 if g, \$0 otherwise), and (\$10 if g or b, \$0 otherwise) is strictly preferred to (\$10 if r or b, \$0 otherwise). These rankings violate SEU, but are consistent with the MEU decision model. According to the latter, for all "acts" f, g mapping realizations to prizes, f is weakly preferred to g if and only if

$$\min_{p\in\mathcal{P}} \int u(f(s))\,p(ds) \geq \min_{p\in\mathcal{P}} \int u(g(s))\,p(ds), \tag{1}$$

where u is a utility index and $\mathcal{P}$ a set of priors. The preferences described above for the three-color urn example are consistent with the MEU decision model in Eq. (1) if $u(\$10) > u(\$0)$ and, for instance, $\mathcal{P}$ is the set of all probability distributions p on $\{r, g, b\}$ such that $p(r) = \frac{1}{3}$ and $\frac{1}{6} \leq p(g) \leq \frac{1}{2}$ (other choices of priors are possible).

Now consider Hurwicz' α-maxmin expected utility (α-MEU) model, which prescribes that f be weakly preferred to g if and only if

$$\begin{aligned}\alpha \min_{q\in\mathcal{Q}} \int u(f(s))\,q(ds) + (1-\alpha)\max_{q\in\mathcal{Q}} \int u(f(s))\,q(ds) \geq \\ \alpha \min_{q\in\mathcal{Q}} \int u(g(s))\,q(ds) + (1-\alpha)\max_{q\in\mathcal{Q}} \int u(g(s))\,q(ds),\end{aligned} \tag{2}$$

where $\mathcal{Q}$ is a set of priors and $\alpha \in [0, 1]$. It is easy to see that, if $\alpha = \frac{3}{4}$ and $\mathcal{Q}$ comprises all probabilities q over $\{r, g, b\}$ such that $q(r) = \frac{1}{3}$, one obtains an alternative representation the MEU preferences characterized by the set of priors $\mathcal{P}$ specified above. In other words, the *same* preference ordering admits *two* representations: MEU with priors $\mathcal{P}$, or $\frac{3}{4}$-MEU with priors $\mathcal{Q}$.[3] Additional considerations are required to decide which of the two sets $\mathcal{P}$ and $\mathcal{Q}$, if any, consists of priors the decision-maker considers possible—hence, which decision criterion reflects this individual's attitudes towards ambiguity.

The proposed definition of "plausible priors".

A probability P is deemed a *plausible prior* if there exists a subset C of acts with the following properties: (i) when restricted to C, the decision-maker's preferences are consistent with SEU, i.e. conform to the Savage [31] or Anscombe-Aumann [2] axioms; and (ii) P is the

[3] A similar construction shows that these preferences admit α-MEU representations for any $\alpha \in [\frac{3}{4}, 1]$. Moreover, α-MEU-type representations featuring *arbitrarily small subsets* of $\mathcal{P}$ can also be constructed. Section 6.2 in the Online Appendix shows that similar constructions are feasible for all MEU preferences.

only probability that, jointly with a suitable utility function, provides a SEU representation of preferences restricted to the set C.

Consider the three-color urn example. Let C_1 be the set of acts f such that $f(b)$ is weakly preferred to $f(g)$, and let C_2 be the set of acts f for which $f(g)$ is weakly preferred to $f(b)$; as above, consider MEU preferences with priors $\mathcal{P}$. Then preferences restricted to C_1 are consistent with SEU, and uniquely identify the subjective probability P_1 characterized by $P_1(r) = \frac{1}{3}$ and $P_1(b) = \frac{1}{6} = \frac{2}{3} - P_1(g)$; this is because P_1 minimizes the expected utility of any act in C_1 over the set $\mathcal{P}$, so the MEU evaluation of any such act is precisely $\int u \circ f \, dP_1$. Similarly, preferences restricted to C_2 are consistent with SEU, with unique prior P_2, where $P_2(r) = \frac{1}{3}$ and $P_2(g) = \frac{1}{6} = \frac{2}{3} - P_2(b)$. Thus, P_1 and P_2 are plausible priors for this preference (and they are the only ones). Finally, recall that the same preferences also admit a $\frac{3}{4}$-MEU representation, with priors $\mathcal{Q}$; but since the two representations are numerically identical, it is still the case that the evaluation of an act f in C_1 is given by $\int u \circ f \, dP_1$, and similarly the evaluation of acts $f \in C_2$ is $\int u \circ f \, dP_2$. Thus, the same plausible priors are obtained, regardless of the overall representation of preferences one decides to work with.

To motivate the proposed definition, observe first that a decision-maker whose preferences admit some form of multiple-prior representation, such as MEU or α-MEU, can be described as (a) specifying a relevant collection of probabilities, and (b) evaluating any act according to its expected utility, computed with respect to a *suitably selected* probability drawn from this collection. Different acts may be evaluated using different probabilities; also, distinct multiple-prior models differ in the way evaluation probabilities are selected for each acts.

In the three-color urn example, the pre-specified sets of priors are $\mathcal{P}$ for the MEU representation of preferences, and $\mathcal{Q}$ for the α-MEU representation. In the MEU representation, the prior p_f used to evaluate an act f is selected so as to minimize the expected utility of f over $\mathcal{P}$. In the $\frac{3}{4}$-MEU representation, the act f is evaluated using the prior $\frac{3}{4}q_f + \frac{1}{4}Q_f$, where q_f minimizes the expected utility of f over the set $\mathcal{Q}$, and Q_f maximizes it.[4]

This interpretation applies to a broad class of decision models that includes CEU.[5] Moreover, it is fully consistent with the intuitive interpretation of priors discussed at the beginning of this Introduction: it portrays an individual who is willing to entertain more than one possible probabilistic description of the underlying uncertainty, and responds to perceived

[4]Note that $\frac{3}{4}q_f + \frac{1}{4}Q_f \in \mathcal{Q}$, so it is appropriate to say that the α-MEU representation "selects" an element of $\mathcal{Q}$ to evaluate each act.

[5]In particular, it applies to the "generalized α-MEU" class of preferences axiomatized by Ghirardato, Maccheroni and Marinacci [11]; in their representation, the coefficient α is a function of the act being evaluated. This class is characterized by the Gilboa-Schmeidler [14] axioms *other than* Uncertainty Aversion, and includes all CEU preferences.

ambiguity by evaluating different acts by means of different possible priors.

Now suppose that a multiple-prior representation of preferences is sought, but *no set of probabilities or decision criterion is specified a priori*: the objective is to derive both from preferences. In this respect, a plausible prior P is a strong candidate for inclusion in the set of probabilities that characterize the decision-maker's choices: by definition, the decision-maker behaves *as if* P was the prior "selected" to evaluate acts in a set C. In other words, although a set of probabilities is not specified a priori, the decision-maker behaves *as if* this set contained P, and her decision criterion specified that P be used for acts in C.

This interpretation would be less compelling if there were another probability $Q \neq P$ that also yielded the correct evaluation of acts in the set C: in this case, the decision-maker could also be said to behave as if Q, not P, was the prior "selected" for acts in the set C. The uniqueness requirement in the definition of plausible priors guarantees that this possibility does not arise; for this reason, this requirement is essential to the intended interpretation.

Finally, the above discussion suggests that, by repeatedly applying the proposed definition, it may be possible to derive a multiple-prior representation of a given preference relation, wherein the characterizing set of probabilities consists solely of plausible priors. As noted above, under the proposed axioms, this is indeed the case. In the three-color urn experiment, preferences admit a MEU representation, with plausible priors $\{P_1, P_2\}$.

I now briefly discuss additional important aspects of the definition of plausible prior. First, the proposed definition is fully behavioral, as intended: it does not rely upon any pre-specified functional representation of preferences. One consequence was noted above in the analysis of the three-color urn example: regardless of which representation of preferences one chooses to work with, the same plausible priors are obtained.

Second, the definition identifies plausible priors independently of the decision-maker's attitudes towards ambiguity (cf. Sec. 2.4). For instance, in the three-color-urn example, the plausible priors for a decision-maker with $\max max$-expected utility preferences and priors $\mathcal{P}$ are also P_1 and P_2, even though this decision-maker is ambiguity-loving.

Third, a possible alternative to Condition (i) in the definition of plausible priors might require that preferences on a subset C of acts be probabilistically sophisticated, but not necessarily consistent with SEU. This leads to the arguably interesting alternative notion of "plausible non-SEU prior". However, expected utility is the key building block of decision criteria such as MEU, α-MEU, and even CEU, as well as a central component of their intuitive interpretation. Since this paper is motivated by the interpretation of such decision models, it seems natural to adopt expected utility as basic building block, and defer non-SEU extensions to future research; see Sec. 6.3 in the Online Appendix for further discussion.

Finally, as demonstrated in Section 2.4, the axiomatic framework adopted in this paper is consistent with a variety of decision models, and accommodates a broad range of attitudes towards ambiguity. However, it does rule out certain interesting preferences, such as those consistent with the second-order-probability model axiomatized by Klibanoff, Marinacci and Mukerji [21]; see the discussion of Axiom 5 in Sec. 2.2.1. On the other hand, the definition of plausible prior proposed here, as well as the key behavioral assumption of this paper (Axiom 6 in Sec. 2.2.4), are not a priori inconsistent with such preferences; extending the axiomatic framework to accommodate them is another interesting direction for future research.

This paper is organized as follows. Section 2 introduces the decision framework, formulates and motivates the axioms, presents the main characterization result, and applies it to known decision models. Section 3 establishes the equivalence of probabilistic sophistication and SEU for plausible-priors preferences, and presents the characterization of Bayesian updating. Section 4 discusses the related literature. All proofs are in the Appendix.

2 Model and Characterization

2.1 Decision-Theoretic Setup

I adopt a slight variant of the Anscombe-Aumann [2] decision framework. Consider a set of states of nature S, endowed with a sigma-algebra Σ, a set X of consequences (prizes), the set Y of (finite-support) lotteries on X. For future reference, a *charge* is a finitely, but not necessarily countably additive measure on (S, Σ).

Acts are Σ-measurable maps from S into Y that are bounded in preference: cf. Gilboa and Schmeidler [14], Sec. 4. I assume that preferences are defined over all such acts. Formally, let $\succeq_0$ be a binary relation on Y; say that a function $f : S \to Y$ is Σ-measurable if, for all $y \in Y$, the sets $\{s : f(s) \succeq_0 y\}$ and $\{s : f(s) \succ_0 y\}$ belong to Σ; then, let L be the collection of all Σ-measurable maps $f : S \to Y$ for which there exist $y, y' \in Y$ such that $y \succeq_0 f(s)$ and $f(s) \succeq_0 y'$ for every $s \in S$. With the usual abuse of notation, denote by y the constant act assigning the lottery $y \in Y$ to each $s \in S$. Finally, denote by $\succeq$ a preference relation on L that extends $\succeq_0$: that is, for all $y, y' \in Y$, $y \succeq y'$ if and only if $y \succeq_0 y'$. Denote the asymmetric and symmetric parts of $\succeq$ by $\succ$ and $\sim$ respectively.

Mixtures of acts are taken pointwise: if $f, g \in L$ and $\alpha \in [0, 1]$, $\alpha f + (1 - \alpha)g$ is the act assigning the compound lottery $\alpha f(s) + (1 - \alpha)g(s)$ to each state $s \in S$.

Finally, a notion of uniform convergence for sequences of acts will be introduced. Let $\{f_k\} \subset L$ be a sequence of acts. Say that this sequence *converges uniformly in preference to the act* $f \in L$, written "$f_k \to f$", iff, for every $\alpha \in (0, 1)$ and $y, y' \in Y$ such that $y \succ y'$, there exists $K \geq 1$ such that, for all $k \geq K$, and for all $s \in S$,

$$\alpha f_k(s) + (1-\alpha)y' \preceq \alpha f(s) + (1-\alpha)y \quad \text{and} \quad \alpha f(s) + (1-\alpha)y' \preceq \alpha f_k(s) + (1-\alpha)y. \quad (3)$$

To interpret this definition, fix α, y and y' as above: this corresponds to fixing a neighborhood of a point in a metric space. Suppose that k is large enough so that the above relations hold. Consider the case $f_k(s) \succ f(s)$: then, loosely speaking, the first relation in Eq. 3 requires that the preference for $f_k(s)$ over $f(s)$ be weaker than the preference for y over y'. If instead $f_k(s) \prec f(s)$, then the second relation in Eq. 3 requires that the preference for $f(s)$ over $f_k(s)$ be weaker than the preference for y over y'. Thus, in either case, $f_k(s)$ and $f(s)$ are required to be "closer in preference" than y and y', and this must hold uniformly in s.[6]

A sequence of constant acts, or lotteries, $\{y_k\} \subset Y$ such that $y_k \to y \in Y$ according to the preceding definition will be said simply to "converge in preference".

The Anscombe-Aumann setup is adopted here merely for expository convenience. The analysis can be equivalently carried out in a "fully subjective" framework, following e.g. [5] or [12]. Specifically, let X be a "rich" (e.g. connected, separable topological) space of prizes; define acts as bounded, measurable maps from S to X. Then, under appropriate assumptions, preferences over prizes are represented by a utility function u such that $u(X)$ is a convex set; moreover, it is possible to define a "subjective" mixture operation $\oplus$ over prizes such that, for all $x, x' \in X$, and $\alpha \in [0,1]$, $u(\alpha x \oplus (1-\alpha)x') = \alpha u(x) + (1-\alpha)u(x')$. All mixture axioms stated below can then be reformulated by replacing "objective" lottery mixtures with subjective mixtures. Moreover, these axioms can be interpreted in a manner that is consistent with both their "objective" and "subjective" formulation.

2.2 Axioms and Interpretation

I begin by introducing a set of basic structural axioms (Axioms 1–5 in §2.2.1). Next, the notion of mixture neutrality is employed to provide a formal definition of plausible priors (§2.2.2). Then, I discuss the notion of hedging against ambiguity and define robust mixture neutrality (§2.2.3); the latter is finally employed to formulate the key axiom for preferences that admit plausible priors: "No Local Hedging" (Axiom 6 in §2.2.4).

2.2.1 Basic Structural Axioms

Axioms 1–5 will be applied both to the set L of all acts, and to certain subsets of L. For this reason, they are stated using intentionally vague expressions such as "for all acts f, g...".

[6] Lemma 5.2 Part 1 shows that, under the basic structural axioms considered below, Eq. 3 is equivalent to the condition $|u(f_k(s)) - u(f(s))| \leq \frac{1-\alpha}{\alpha}[u(y) - u(y')]$, where $u : Y \to \mathbb{R}$ represents $\succeq$ on Y.

Axioms 1–4 appear in "textbook" treatments of the Anscombe-Aumann characterization result, as well as in [14] and [32]; Axiom 5 was introduced by Gilboa and Schmeidler [14].

Axiom 1 (Weak Order) *$\succeq$ is transitive and complete.*

Axiom 2 (Non-degeneracy) *Not for all acts f, g, $f \succeq g$.*

Axiom 3 (Continuity) *For all acts f, g, h such that $f \succ g \succ h$, there exist $\alpha, \beta \in (0,1)$ such that $f \succ \alpha f + (1-\alpha)h \succ g$ and $g \succ \beta f + (1-\beta)h \succ h$.*

Axiom 4 (Monotonicity) *For all acts f, g, if $f(s) \succeq g(s)$ for all $s \in S$, then $f \succeq g$.*

Axiom 5 (Constant-act Independence) *For all acts f, g, all $y \in Y$, and all $\alpha \in (0,1)$: $f \succeq g$ if and only if $\alpha f + (1-\alpha)y \succeq \alpha g + (1-\alpha)y$.*

While Axiom 5 is standard, it warrants some discussion. Recall that the usual Independence axiom requires that the ranking of any two acts f and g be preserved under mixtures with *any* third act h; Constant-act Independence restricts this requirement to mixtures with lotteries, but allows for preference reversals when mixing with non-constant acts. This is motivated by the observation that mixing f and g with non-constant acts may allow for *hedging* against ambiguity, as was first suggested by Schmeidler [32]: also see §2.2.3 below. On the other hand, mixtures with constant acts corresponds to changing the "scale and location" of outcomes (in utility, or preference, terms), uniformly in each state and for both acts f and g; Constant-act Independence thus reflects a form of invariance of preferences in situations where no hedging can occur. However, it should be also noted that Constant-act Independence also incorporates a notion of "constant ambiguity aversion", as is discussed in Klibanoff, Marinacci and Mukerji [21].

Observe that the interpretation provided here does not involve objective randomizations in an essential way; the discussion is entirely stated in terms of changes in the outcome profiles of acts. Hence, this interpretation applies equally well to the current Anscombe-Aumann setup and to fully subjective settings, given a suitable mixture operation.

2.2.2 Mixture Neutrality and Plausible Priors

Recall that a prior P is deemed plausible if preferences are consistent with SEU on a subset $C \subset L$ of acts, and P is the only probability that yields a SEU representation of preferences on C. Under Axioms 1–5, consistency with SEU is characterized by one additional property, *Mixture Neutrality*.[7] This simplifies the formal definition of plausible priors.

Definition 2.1 (Mixture-neutral acts) Two acts $f, g \in L$ are *mixture-neutral* (denoted $f \simeq g$) iff,[8] for every pair of lotteries $y, y' \in Y$ such that $f \sim y$ and $g \sim y'$, and every $\alpha \in [0,1]$, $\alpha f + (1-\alpha)g \sim \alpha y + (1-\alpha)y'$.

The connection between mixture neutrality and ambiguity is discussed in §2.2.3. The following proposition confirms that mixture neutrality is the key property characterizing SEU preferences in the class of preferences that satisfy Axioms 1–5. This is the case whether axioms are applied to the entire set L of acts, or to an appropriate subset thereof.[9]

Proposition 2.2 *Consider a preference relation $\succeq$ on L and a convex subset C of L that contains all constant acts. Then the following statements are equivalent:*
1. $\succeq$ satisfies Axioms 1-5 on C; furthermore, for all $f, g \in C$, $f \simeq g$.
2. there exists a probability charge P on (S, Σ), and a non-constant, affine function $u : Y \to \mathbb{R}$, unique up to positive affine transformations, such that, for all acts $f, g \in C$, $f \succeq g$ if and only if $\int u(f(\cdot))\, dP \geq \int u(g(\cdot))\, dP$.

Relative to the usual characterizations of SEU, an essential feature of Proposition 2.2 is the fact that uniqueness of the probability charge P is *not* guaranteed for arbitrary sets C, even if preferences are non-degenerate (i.e. Axiom 2 holds). Instead, this is explicitly required in the formal definition of plausible prior, which can finally be stated.

Definition 2.3 (Plausible Prior) Consider a preference relation $\succeq$ that satisfies Axioms 1–5 on L. A probability charge P on (S, Σ) is a *plausible prior* for $\succeq$ iff there exists a convex subset C of L containing all constant acts and such that

(i) for all acts $f, g \in C$, $f \simeq g$;

(ii) P is the *unique* charge that provides a SEU representation of $\succeq$ on C.

[7]The Anscombe-Aumann characterization of SEU employs Axioms 1–4 plus the standard Independence axiom; however, under Axioms 1–4, a preference satisfies the latter if and only if it satisfies Axiom 5 and the Mixture Neutrality axiom, to be introduced momentarily.

[8]As usual, "iff" stands for "if and only if" in definitions.

[9]This result is standard if the set C in the statement equals L (or the collection of simple acts); for the general case, see the comments following the proof of Lemma 5.11 in the Appendix.

2.2.3 Hedging and Mixture Neutrality

Gilboa and Schmeidler [14] and Schmeidler [32] suggest that two acts f, g may fail to be mixture-neutral if their mixtures provide a *hedge* against perceived ambiguity;[10] conversely, a decision-maker for whom all acts are mixture-neutral either does not perceive ambiguity, or chooses not to respond to it.[11] Consistently with this intuition, invoking Proposition 2.2 with $C = L$ confirms that such an individual exhibits SEU preferences.

Decision models that depart from SEU to account for ambiguity differ widely in the violations of mixture neutrality they allow; as a result, they capture a variety of different attitudes towards ambiguity. Yet, as will be shown in §2.4, plausible priors exist for a rich class of such models. The key behavioral assumption of this paper, "No Local Hedging" (Axiom 6 in §2.2.4), is designed to identify preferences that admit plausible priors *without imposing a priori restrictions on ambiguity attitudes.*

The basic intuition underlying the No Local Hedging axiom is close in spirit to the logic behind Comonotonic Independence (Schmeidler [32]). As a first approximation, for a mixture of two acts f and g to provide an effective hedge against perceived ambiguity, f should yield "good" outcomes in states where g yields "bad" outcomes, and vice versa. If f and g yield nearly equivalent outcomes in each state, i.e. if they are uniformly close to one another, their mixtures arguably cannot provide an effective hedge. For this reason, *Axiom 6 requires that, loosely speaking, "nearby" act be mixture-neutral.* Since it imposes only "local" restrictions on preferences, Axiom 6 is compatible with a wide variety of "global" ambiguity attitudes, as the results in §2.4 indicate.

This subsection motivates the formulation of Axiom 6 by fleshing out these considerations. As noted above, the key observation is that the basic hedging intuition can lead to many different patterns of departure from mixture neutrality, even for MEU preferences; accommodating such diverse preference patters is a main desideratum of this paper.

In all examples, the set of prizes is $X = \{\$0, \$10\}$, and the state space S is finite. A lottery $y \in Y$ can thus be identified with the probability of receiving \$10, and an act f is represented by a tuple $(f(s_1), \ldots, f(s_{|S|})) \in [0,1]^S$: $f(s_n)$ is the probability of receiving \$10 in state s_n. By Axioms 1–5, the individual has EU preferences over Y, so such tuples can

[10] All considerations concerning MEU preferences in this subsection and the next apply to "maxmax EU", or 0-MEU, preferences as well (with the appropriate modifications).

[11] In the language of [14] and [32], mixture neutrality implies both "uncertainty aversion"and "uncertainty appeal", so this property may also be termed "uncertainty neutrality". But some authors (e.g. [9]) note that a preference for mixtures of acts vs. their certainty equivalents may not always be a good definition of "uncertainty aversion". For this reason, I adopt the less controversial terminology "mixture neutrality".

also be interpreted as utility profiles. Finally, $\Delta(S)$ is the probability simplex.

Example 1 (Ann) This is Ellsberg's three-color-urn experiment described in the Introduction, restated here for notational uniformity. Ann has MEU preferences, with priors $\mathcal{Q}_A = \{q \in \Delta(S) : p(r) = \frac{1}{3}, \frac{1}{6} \leq p(g) \leq \frac{1}{2}\}$ (denoted $\mathcal{P}$ in the Introduction).

Consider the acts $f_g = (0, 1, 0)$ and $f_b = (0, 0, 1)$ and the intuitively ambiguous events $\{g\}$ and $\{b\}$. Note that $f_g(g) \succ f_g(b)$ and $f_b(g) \prec f_b(b)$: that is, f_g and f_b respond in complementary ways to realizations of the underlying uncertainty. By mixing f and g, Ann can reduce variations in outcomes[12] across the ambiguous states g and b; indeed, $\frac{1}{2}f_g + \frac{1}{2}f_b$ is constant on $\{g, b\}$. Consistently with the hedging intuition, $f_g \not\succ f_b$.

Example 2 (Bob) (cf. Klibanoff [20], Ex. 1). A ball is drawn from an urn containing an equal, non-zero, but unspecified number of red and blue balls, and a non-zero, but unspecified number of green balls; thus, $S = \{r, g, b\}$. Bob has MEU preferences, with priors $\mathcal{Q}_B = \{q \in \Delta(S) : \epsilon \leq p(r) = p(b) \leq \frac{1-\epsilon}{2}\}$, for some $\epsilon > 0$.

Let $f = (.2, .3, .5)$ and $f' = (.1, .4, .6)$. Note that there are no $s, s' \in S$ with $f(s) \succ f(s')$ and $f'(s) \prec f'(s')$: that is, f and f' are *comonotonic* (cf. [32]). Yet, $f \not\succ f'$. This may be intuitively explained as follows. Since the urn contains an equal (albeit unknown) number of red and blue balls, it may be said that both f and f' yield the "expected" outcome .35 conditional upon the event $\{r, b\}$; thus, abusing notation, $f(g) \succ f(\{r, b\})$ and $f'(g) \prec f'(\{r, b\})$. By analogy with Example 1, this suggests that mixtures of f and f' provide hedging opportunities relative to the intuitively ambiguous *events* $\{r, b\}$ and $\{g\}$.

Informally, while Ann only cares about hedging outcomes across ambiguous *states*, Bob also cares about hedging "conditional expected outcomes" across disjoint, ambiguous *events*. Thus, to avoid restricting ambiguity attitudes, the proposed behavioral axiom should allow eventwise as well as statewise hedging.

Moreover, formally defining "eventwise hedging" necessarily involves the notion of "conditionally expected outcome". If the individual's conditional preferences are available, or can be derived from her unconditional preferences via some updating rule, then it seems sensible to stipulate that a lottery $y(f, E)$ is a conditional expected outcome, or *conditional evaluation*, of an act f given an event E if $y(f, E)$ and f are indifferent conditional upon E. Unfortunately, the literature considers several different updating rules, including prior-by-prior Bayesian updating for MEU preferences (see Section 3.1 below for references) and

[12]More precisely, mixtures of f and g reduce variations in *utilities*—equivalently, they reduce "preference variation" in outcomes.

"h-Bayesian updating" for ambiguity-averse CEU preferences (Gilboa and Schmeidler [15]). Each of these updating rules yields a potentially different notion of conditional evaluation.[13]

A characterization of plausible priors should arguably be robust to different choices of updating and conditional evaluation rules. This can be achieved by restricting attention to "nearby acts". The basic intuition is easiest to explain if the state space is finite. Assume that conditional preferences also satisfy Axioms 1–5, so that, in particular, if two acts f and g are uniformly close in preference, then so are $y(f, E)$ and $y(g, E)$ for every event E.[14] It is then clear that, if two acts f, g are sufficiently (uniformly) close in preference, there cannot be two events E, F such that $y(f, E) \succ y(f, F)$ and $y(g, E) \prec y(g, F)$. In other words, *regardless of the notion of conditional evaluation one adopts, if two acts are sufficiently close, they do not offer any opportunity for hedging across any pair of events.* This suggests the following assumption:

> *for any sequence $\{f_k\}$ of acts that converges to an act f uniformly in preference, there is a subsequence $\{f_{k(\ell)}\}$ such that $f_{k(\ell)} \simeq f_{k(\ell')}$ for all ℓ, ℓ'.*

Except for a modification that will be discussed in the next subsection, this is the content of Axiom 6. As intended, this formulation accommodates both statewise and eventwise hedging, and does not require committing to a specific conditional evaluation rule. It requires that mixture neutrality hold only for pairs of acts that arguably offer no hedging opportunities.

If the state space S, and hence the sigma-algebra Σ, are infinite, it is no longer possible to ensure that, in general, if two acts f and g are "close enough" they do not offer eventwise hedging opportunities. However, if conditional preferences satisfy Axioms 1–5, it turns out that the corresponding conditional evaluations converge *uniformly* in the conditioning event E.[15] As a consequence, if f and g are sufficiently close, but there exist disjoint events E, F such that $y(f, E) \succ y(f, F)$ and $y(g, E) \prec y(g, F)$, it must nevertheless be the case that $y(f, E)$ and $y(f, F)$ and, respectively, $y(g, E)$ and $y(g, F)$ are *nearly indifferent* (i.e. "close in preference": cf. Footnote 12). Thus, Axiom 6 can be interpreted as requiring that mixture neutrality hold when hedging opportunities are "small" (if they exist at all).

[13]In Example 2, the choice of $\mathcal{Q}_B$ ensures that 0.35 is the conditional evaluation of both f and f' given $\{r, b\}$ according to both prior-by-prior Bayesian updating, and to any h-Bayesian update rule.

[14] This is true for most updating rules, under Axioms 1–5; see Sec. 6.4 of the Online Appendix for details.

[15]The last assertion of Lemma 5.1 in the Appendix implies that, if preferences conditional upon E satisfy Axioms 1–5 (and agree with unconditional preferences on Y), $|u(y(f_k, E)) - u(y(f, E))| \leq \sup_{s \in S} |u(f_k(s)) - u(f(s))|$; this implies the claim.

2.2.4 Robust Mixture Neutrality and the No Local Hedging Axiom

A final issue must be addressed before Axiom 6 can be formally stated. While mixture neutrality is always associated with absence of hedging opportunities for MEU preferences, this is *not* the case for more general preferences that satisfy Axioms 1–5.

Example 3 (Chloe) Consider draws from a four-color urn of unknown composition; let $S = \{r, g, b, w\}$, where w is for "white". Chloe has α-MEU preferences (cf. the Introduction), with $\alpha = \frac{3}{4}$ and set of priors $\mathcal{Q}_C = \Delta(S)$. These preferences satisfy Axioms 1–5.

Now let $f = \left(1, \frac{2}{3}, \frac{1}{2}, \frac{1}{2}\right)$ and $f' = \left(\frac{1}{2}, \frac{1}{2}, 1, \frac{2}{3}\right)$. Under a mild "consequentialism" condition (cf. Axiom 7 in Sec. 3.1) that is satisfied by all the updating rules described after Ex. 2, $y(f, \{r, g\}) \succ y(f, \{b, w\})$ and $y(f', \{r, g\}) \prec y(f', \{b, w\})$; furthermore, the events $\{r, g\}$ and $\{b, w\}$ are intuitively ambiguous.[16] Hence, it seems plausible to expect violations of mixture neutrality; yet, as may be verified, $f \simeq f'$.

On the other hand, the mixture neutrality of f and f' is not "robust". Consider for instance a small perturbation of f, such as the act $f_\epsilon = \left(1 - \epsilon, \frac{2}{3}, \frac{1}{2}, \frac{1}{2}\right)$ for a suitable small $\epsilon > 0$. As above, $y(f_\epsilon, \{r, g\}) \succ y(f_\epsilon, \{b, w\})$ and $y(f', \{r, g\}) \prec y(f', \{b, w\})$; but now, consistently with the hedging intuition, it may be verified that $f_\epsilon \not\simeq f'$.

Example 3 indicates that non-MEU preferences allow for knife-edge instances of mixture-neutrality for acts that do provide hedging opportunities according to the preceding discussion. To rule out such instances, mixture neutrality must be *robustified*.

First of all, for two acts f and g to be "robustly" mixture-neutral, small perturbations of f should not affect mixture neutrality with g: if $h_k \to f$, then $h_k \simeq g$ for k large.

Furthermore, note that $f \simeq g$ implies that $f \simeq \gamma f + (1 - \gamma)g$ for all $\gamma \in (0, 1)$. This is consistent with the hedging interpretation: if mixtures of f and g provide no hedging opportunities, then neither do mixtures of f and $\gamma f + (1 - \gamma)g$. But if $\{h_k\}$ converges uniformly to f in preference, for k large, h_k and $\gamma f + (1 - \gamma)g$ should also offer no hedging opportunities; thus, it seems plausible to also require that $h_k \simeq \gamma f + (1 - \gamma)g$ for k large.

Definition 2.4 (Robustly mixture-neutral acts) Two acts $f, g \in L$ are *robustly mixture-neutral* (written $f \approx g$) iff, for every sequence $\{h_k\} \subset L$ such that $h_k \to f$ or $h_k \to g$, and for every $\gamma \in [0, 1]$, there exists K such that $h_k \simeq \gamma f + (1 - \gamma)g$ for all $k \geq K$.

Notice that $f \approx g$ implies $f \simeq g$ (consider the sequence $\{h_k\}$ given by $h_k = f$ for all k): that is, robust mixture neutrality is a strengthening of mixture neutrality, as intended.

[16] These events are ambiguous according to the definition provided by Ghirardato and Marinacci [13]. The example can be modified to ensure that they are also ambiguous in the sense of Epstein and Zhang [10].

The key behavioral axiom in this paper can finally be stated.

Axiom 6 (No Local Hedging) *For all sequences $\{f_k\} \subset L$ and acts $f \in L$ such that $f_k \to f$, there exists a subsequence $\{f_{k(\ell)}\}$ such that $f_{k(\ell)} \approx f_{k(\ell')}$ for all ℓ, ℓ'.*

Axiom 5 (Constant-act Independence), the notion of mixture neutrality, and hence Axiom 6 all involve forms of invariance of certain preference patterns to mixtures. However, Axiom 5 entails a *global* restriction on preferences, whereas, for the reasons discussed above, Axiom 6 has a distinctly *local* character.

2.3 The Main Result

One last definition is required before stating the main result. The utility profile of an act $f \in L$ is an element of the space $B(S,\Sigma)$ of bounded, Σ-measurable real functions on S. Theorem 2.6 states that, under Axioms 1–5 and 6, $B(S,\Sigma)$ can be covered by finitely many sets $C_1,\ldots,C_N$ that satisfy certain algebraic and topological properties; to each such set C_n is associated a unique probability charge P_n on (S,Σ); a representation of preferences is obtained by associating to every act f whose utility profile $u \circ f$ lies in C_n the integral $\int u \circ f \, dP_n$. For ease of reference, the properties of the sets $C_1,\ldots,C_n$ are listed in Def. 2.5.

Definition 2.5 A finite collection $C_1,\ldots,C_N$ of subsets of $B(S,\Sigma)$ is a *proper covering* iff

1. $\bigcup_n C_n = B(S,\Sigma)$;
2. for every $n = 1,\ldots,N$, C_n is non-empty and equal to the closure of its interior; furthermore, for $n, m \in \{1,\ldots,N\}$ such that $n \neq m$, $C_n \cap C_m$ has empty interior;
3. for every $n = 1,\ldots,N$, if $a \in C_n$, $\alpha, \beta \in \mathbb{R}$, and $\alpha \geq 0$, then $\alpha a + \beta \in C_n$;
4. for every $n = 1,\ldots,N$: every infinite subset $C \subset C_n$ contains a countably infinite collection $\{a_k\}$ such that, for all k and ℓ, and all $\gamma \in [0,1]$, there exists $\epsilon > 0$ such that, for all $b \in B(S,\Sigma)$,

$$\min\left(\|b-a_k\|, \|b-a_\ell\|\right) < \epsilon, b \in C_n \quad \Rightarrow \quad \forall \lambda \in (0,1), \lambda b + (1-\lambda)[\gamma a_k + (1-\gamma)a_\ell] \in C_n. \tag{4}$$

Consider Property 4 first. The sets C_n are *not* required to be convex; indeed, in general, for common multiple-priors decision models such as α-MEU with $\alpha \in (0,1)$, they are not. Property 4 instead requires a "local" version of convexity: every infinite subset of C_n contains a countable collection $\{a_k\}$ with the property that, for any k and ℓ, the segment joining points close to a_k or a_ℓ with any point between a_k and a_ℓ lies in C_n.

Two observations are in order. First, a simple sufficient condition for Property 4 to hold can be provided: see Section 5.1.2 in the Appendix for the proof.

Remark 1 If a set C_n is a union of finitely many closed convex subsets of $B(S, \Sigma)$, then it satisfies Property 4.

As will be demonstrated in the next subsection, many known decision models consistent with the axioms proposed here satisfy the simpler (stronger) sufficient condition in Remark 1.

Second, Property 4 is not particularly restrictive by itself. For instance, the set $B(S, \Sigma)$ can always be covered by uncountably many sets of the form $\{\alpha a + \beta : a \in B(S, \Sigma), \alpha, \beta \in \mathbb{R}, \alpha \geq 0\}$: that is, cones of affinely related functions. Since these sets are convex, Remark 1 implies that each such set satisfies Property 4 in Definition 2.5.

Property 2 implies that each C_n has non-empty interior, which ensures that unique probabilities can be identified. Also note that the interior of a convex set A in a linear topological space is dense in A, if it is non-empty (Holmes [17], Theorem 11.A); this again suggests that Def. 2.5 ensures that the sets C_n enjoy some of the properties of convex sets, even though they may not actually be convex.

Property 3 states that every C_n is closed under non-negative affine transformations; in particular, it contains all constant functions, and all non-negative multiples of its elements. This corresponds to constant-linearity of the representation, and hence to Axiom 5.

The main result of this paper can finally be stated.

Theorem 2.6 *Let $\succeq$ be a preference relation on L. The following statements are equivalent:*

1. $\succeq$ satisfies Axioms 1–5 and 6;

2. There exist an affine function $u : Y \to \mathbb{R}$, a proper covering $C_1, \ldots, C_N$, and a collection of probability charges $P_1, \ldots, P_N$ such that, for all $n, m \in \{1, \ldots, N\}$:

(i) $n \neq m$ implies $P_n \neq P_m$; however, for all $a \in C_n \cap C_m$, $\int a\, dP_n = \int a\, dP_m$;

(ii) for all $f, g \in L$, if $u \circ f \in C_n$ and $u \circ g \in C_m$, then

$$f \succeq g \iff \int u \circ f\, dP_n \geq \int u \circ g\, dP_m. \tag{5}$$

Furthermore, in Statement 2:

(a) u is unique up to positive affine transformations;

(b) for every $n \in \{1, \ldots, N\}$, if a probability charge Q is such that, for all $f, g \in L$ with $u \circ f, u \circ g \in C_n$, $f \succeq g$ iff $\int u \circ f dQ \geq \int u \circ g dQ$, then $Q = P_n$; and

(c) if a proper convering $D_1, \ldots, D_M$ satisfies (i) and (ii) jointly with a collection of probability charges $Q_1, \ldots, Q_M$, then $M = N$ and there is a permutation $\{\pi(1), \ldots, \pi(N)\}$ of $\{1, \ldots, N\}$ such that $D_n = C_{\pi(n)}$, hence $Q_n = P_{\pi(n)}$, for all $n = 1, \ldots, N$.

2.4 Corollaries and Interpretation

In light of Theorem 2.6, a proper covering may be viewed as a collection of "menus"; the decision-maker has standard SEU preferences when comparing items on the same menu (i.e. "locally"), but different considerations may guide her choices from different menus.[17] By claim (c) in Theorem 2.6, the proper covering $C_1, \ldots, C_N$ is uniquely determined (up to relabeling) by properties (i) and (ii) in Statement 2, and hence ultimately by preferences.

Each probability charge P_n appearing in Statement 2 is a plausible prior. Claim (b) in the Theorem asserts that every P_n is uniquely determined by preferences over acts whose utility profile lies in C_n; moreover, the set C_n need not be convex, but it does contain convex subsets that include all constant functions, as required by Def. 2.3. Furthermore, it can be shown that no other charge on (S, Σ) can be a plausible prior for $\succeq$:

Corollary 2.7 *Under the equivalent conditions of Theorem 2.6, the plausible priors for $\succeq$ are the charges $P_1, \ldots, P_N$ in Statement 2.*

Henceforth, I will employ the expression *plausible-priors preference* to indicate a binary relation $\succeq$ on L for which the equivalent statements of Theorem 2.6 are true.

As noted in the Introduction, a rich set of MEU, α-MEU and CEU preferences permit the elicitation of plausible priors. Conceptually, this suggests that, within the class of preferences that satisfy Axioms 1–5, *Axiom 6 does not restrict attitudes towards ambiguity*, and hence is compatible with a variety of decision models. The following corollaries provide the details.

Corollary 2.8 *Let $\succeq$ be an α-MEU preference, and let $\mathcal{Q}$ be the corresponding set of priors. If $\mathcal{Q}$ is the weak*-closed, convex hull of finitely many distinct probability charges $\{Q_1, \ldots, Q_N\}$, then $\succeq$ satisfies Axioms 1–5 and 6; the converse is also true if $\alpha \neq \frac{1}{2}$. Let $\mathcal{M} \subset \{1, \ldots, M\}^2$ be defined by*

$$\mathcal{M} = \left\{ (n, m) \ : \ Q_n \in \arg\min_k \int a\, dQ_k,\ Q_m \in \arg\max_k \int a\, dQ_k \text{ for some } a \in B(S, \Sigma) \right\} :$$

the set of plausible priors for $\succeq$ is $\{\alpha Q_n + (1-\alpha) Q_m : (n, m) \in \mathcal{M}\}$.

[17] I owe this interpretation to Mark Machina.

In particular, a MEU preference satisfy Axioms 1–6 if and only if its set of priors is the convex hull of finitely many charges $Q_1, \dots, Q_N$; the latter are its plausible priors.

CEU preferences (Schmeidler [32]) always satisfy the plausible-priors axioms, provided the state space S is finite (this assumption is not necessary for α-MEU preferences). Let $v : 2^S \to [0,1]$ be a capacity on S: that is, $A \subset B \subset S$ imply $v(A) \leq v(B)$, and $v(\emptyset) = 0 = 1 - v(S)$. Assume that $S = \{s_1, \dots, s_M\}$, and let Π_M be the set of all permutations $(\pi_1, \dots, \pi_M)$ of $\{1, \dots, M\}$. Recall that every permutation π identifies a maximal cone of comonotonic functions: $C_\pi = \{a \in B(S, \Sigma) : a(s_{\pi_1}) \geq \dots \geq a(s_{\pi_M})\}$.

Corollary 2.9 *Assume that S is finite and let $\Sigma = 2^S$. Let $\succeq$ be a CEU preference over L, and, for all permutations $\pi \in \Pi_M$, let P_π be the probability distribution defined by*

$$P_\pi(s_{\pi_i}) = v(\{s_{\pi_1}, \dots, s_{\pi_i}\}) - v(\{s_{\pi_1}, \dots, s_{\pi_{i-1}}\}).$$

Then $\succeq$ satisfies Axioms 1–5 and 6, and $\{P_\pi : \pi \in \Pi_M\}$ is its collection of plausible priors.

As noted after Def. 2.5, each set C_n has non-empty interior. This is *not* a necessary consequence of the definition of a plausible prior. On the other hand, it ensures that the plausible priors in Theorem 2.6 can be interpreted as the outcome of an *elicitation "procedure"*.

Fix $n \in \{1, \dots, N\}$, let $g \in L$ be such that $u \circ g$ is an interior point of C_n, and choose prizes $x, x' \in X$ such that $x \succ x'$. For every $E \in \Sigma$, let b_E be the binary act that yields prize x at states $s \in E$, and prize x' elsewhere. Since $u \circ g$ is in the interior of C_n, for $\alpha \in (0,1)$ sufficiently close to 1, $u \circ [\alpha g + (1-\alpha) b_E] \in C_n$; moreover, there exists $\pi_E \in [0,1]$ such that $\alpha g + (1-\alpha) b_E \sim \alpha g + (1-\alpha)[\pi_E x + (1-\pi_E) x']$. It is then easy to see that $\pi_E = P_n(E)$.

The "procedure" just described should be viewed merely as a thought experiment: in practice, identifying points in the interior of each set C_n seems non-trivial. This "procedure" does suggest, however, a sense in which plausible priors obtained in Theorem 2.6 exhibit familiar properties of standard SEU priors, even beyond the requirements of Definition 2.3.

The following Corollary confirms that, under the axioms proposed here, robust mixture neutrality reflects a strong notion of absence of hedging opportunities: loosely speaking, if $f \approx g$, then f and g belong to a set of acts over which preferences are consistent with SEU.

Corollary 2.10 *Under the equivalent conditions of Theorem 2.6, for all $f, g \in L$, $f \approx g$ implies $u \circ f$, $u \circ g \in C_n$ for some $n \geq 1$.*

Notice that the converse of this Corollary may be false: if some set C_n is not convex, then it is possible to find $f, g \in L$ such that $u \circ f, u \circ g \in C_n$, but the segment joining them does not lie in C_n. This, in turn, implies that $f \not\approx g$.

Finally, Theorem 2.6 also implies that *preferences are fully determined by plausible priors*. To clarify this point, it is useful to construct a functional representation of overall preferences on the basis of results in Theorem 2.6. Begin by noting the following Corollary:

Corollary 2.11 *Under the equivalent conditions of Theorem 2.6, for all acts $f, g \in L$, if $\int u \circ f \, dP_n \geq \int u \circ g \, dP_n$ for all $n = 1, \ldots, N$, then $f \succeq g$.*

Now let $\mathcal{R} = \{(\int a \, dP_n)_{n=1\ldots N} : a \in B(S, \Sigma)\}$ be the collection of N-vectors of integrals of functions with respect to each plausible prior obtained in Theorem 2.6. Notice that $\mathcal{R}$ is a vector subspace of $\mathbb{R}^N$ that includes the diagonal $\{(\gamma, \ldots, \gamma) : \gamma \in \mathbb{R}\}$.

Corollary 2.11 makes it possible to construct a representation of preferences that employs the plausible priors $P_1, \ldots, P_N$. Specifically, define a functional $V : \mathcal{R} \to \mathbb{R}$ by

$$\forall a \in B(S, \Sigma), \quad V\left(\left(\int a \, dP_n\right)_{n=1\ldots N}\right) = \int a \, dP_{n^*}, \tag{6}$$

where n^* is such that $a \in C_{n^*}$. Corollary 2.11 ensures that this definition is well-posed; furthermore, by Property (i) in Statement 2 of Theorem 2.6, V is uniquely determined.[18] Clearly, for all acts $f, g \in L$, $f \succeq g$ iff $\left(\left(\int u \circ f \, dP_n\right)_{n=1\ldots N}\right) \geq V\left(\left(\int u \circ g \, dP_n\right)_{n=1\ldots N}\right)$: that is, the functional V and the plausible priors $P_1, \ldots, P_n$ represent preferences.

In accordance with the discussion of multiple-prior decision rules in the Introduction, the map V can be thought of as "selecting" which of the priors $P_1, \ldots, P_N$ can be used to evaluate a given act f.[19] Thus, consistently with the intuitive interpretation of multiple-prior models, the selection criterion formalized by the map V can be viewed as reflecting the individual's attitudes towards ambiguity.

3 "SEU-like" properties of plausible-prior preferences

3.1 Prior-by-prior Bayesian updating

Plausible-priors preferences inherit a key property of SEU preferences: they are "closed under Bayesian updating". More precisely, consider an event $E \in \Sigma$; interpret it as information the decision-maker may receive in the dynamic context under consideration.[20] Correspondingly,

[18]Furthermore, V is *normalized*, i.e. $V(1 \ldots 1) = 1$; *monotonic*: $\varphi_n \geq \psi_n$ for all n implies $V(\varphi) \geq V(\psi)$; *c-linear*: for all $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$, and $\varphi \in \mathcal{R}$, $V(\alpha\varphi + \beta) = \alpha V(\varphi) + \beta$.

[19]In Examples 1 and 2, $N = \{1, 2\}$ and $V(\varphi) = \min_n \varphi_n$. In Example 3, $N = \{1, \ldots, 12\}$, and the functional V can be explicitly described by enumerating the possible orderings of the components of the vector φ, and associating with each such ordering the appropriate prior.

[20]For instance, E may correspond to the information that a given node in a decision tree has been reached.

consider a *conditional preference relation* $\succeq_E$ on the set L of acts; the ranking $f \succeq_E g$ is to be interpreted as stating that the decision-maker *would* prefer f to g, were she to learn that E has occurred. This section provides an axiomatic connection between conditional and unconditional preferences:[21] if unconditional preferences satisfy the equivalent conditions of Theorem 2.6, and if unconditional and conditional preferences satisfy two joint consistency requirements, then: (1) conditional preferences are uniquely determined, and also satisfy the equivalent conditions of Theorem 2.6; and (2) the set of "plausible posteriors" representing conditional preferences is related to the set of plausible priors via Bayesian updating.

Additional notation will be needed. Given any pair of acts $f, g \in L$, let

$$fEg(s) = \begin{cases} f(s) & \text{if} \quad s \in E; \\ g(s) & \text{if} \quad s \notin E. \end{cases} \tag{7}$$

As is the case for SEU preferences, updating is defined only for events that are "relevant" to the decision-maker's preferences. The following definition indicates the relevant restriction.

Definition 3.1 An event $E \in \Sigma$ is *non-null* iff, for all acts $f \in L$, all outcomes $y, y' \in Y$ such that $y \succ y'$, and all $\gamma \in (0,1)$, $\gamma f + (1-\gamma)[y\,E\,y'] \succ \gamma f + (1-\gamma)y'$.

Recall that an event E is *Savage-null* if, for all acts f, g, $f(s) \sim g(s)$ for all $s \notin E$ implies $f \sim g$. If E satisfies Def. 3.1, it is also not Savage-null; but the converse is false in general.[22]

Turn now to the key behavioral restrictions, stated as assumptions regarding an arbitrary conditional preference $\succeq_E$ and the unconditional preference $\succeq$. First, preferences conditional upon the event E are not affected by outcomes at states outside E:

Axiom 7 (Consequentialism) *For every pair of acts $f, h \in L$: $f \sim_E fEh$.*

Second, a weakening of the standard *dynamic consistency* axiom is imposed. Its interpretation (and the relationship with other consistency axioms) is discussed at length in [33]. Loosely speaking, Axiom 8 imposes consistency in situations where hedging considerations are arguably less likely to lead to preference reversals.

Axiom 8 (Dynamic c-Consistency) *For every act $f \in L$ and outcome $y \in Y$:*

$$\begin{aligned} f \succeq_E y, \quad f(s) \succeq y\ \forall s \in E^c \quad &\Rightarrow \quad f \succeq y; \\ f \preceq_E y, \quad f(s) \preceq y\ \forall s \in E^c \quad &\Rightarrow \quad f \preceq y. \end{aligned}$$

[21] The axioms and results are based on Siniscalchi [33].

[22] Let $S = \{s_1, s_2\}$ and $X = \{0,1\}$, and consider MEU preferences with priors $\Delta(S)$. Then $\{s_1\}$ is not Savage-null, because, in the notation of the preceding Section, $(0,1) \prec (1,1)$; however, $\frac{1}{2}(1,0) + \frac{1}{2}(1,0) = (1,0) \sim (\frac{1}{2}, 0) = \frac{1}{2}(1,0) + \frac{1}{2}(0,0)$, so $\{s_1\}$ does not satisfy Def. 3.1 for $f = (1,0)$, $y = 1$ and $y' = 0$.

Moreover, if the preference conditional on E is strict, then so is the unconditional preference.

The dominance conditions $f(s) \succeq y$ and $f(s) \preceq y$ are stated in terms of the *un*conditional preference; equivalently, one could assume that conditional and unconditional preferences agree on Y, and state the dominance conditions in terms of $\succeq_{E^c}$. Also note that strict preference conditional on the event E is required to imply strict unconditional preference.

Theorem 3.2 *Consider an event $E \in \Sigma$. Suppose the preference $\succeq$ satisfies Axioms 1–5 and 6, and assume that E is non-null. Let $\succeq$ be represented by u, $C_1, \ldots, C_n$ and $P_1, \ldots, P_N$ as in Theorem 2.6. Finally, assume that $\succeq_E$ satisfies Axiom 1. Then the following are equivalent:*

1. $\succeq_E$ satisfies Axiom 7, and $\succeq$, $\succeq_E$ jointly satisfy Axiom 8;

2. $\succeq_E$ satisfies the equivalent conditions of Theorem 2.6; in particular, there exists a proper covering $C_1^E, \ldots, C_K^E$ and a subset $\{n_1, \ldots, n_K\} \subset \{1, \ldots, N\}$ of indices such that, for all $k, \ell \in \{1, \ldots, K\}$ and $f, g \in L$ with $u \circ f \in C_k^E$ and $u \circ g \in C_\ell^E$,

$$f \succeq_E g \quad \Longleftrightarrow \quad \int u \circ f \, dP_{n_k}(\cdot|E) \geq \int u \circ g \, dP_{n_\ell}(\cdot|E).$$

Moreover, for every $k \in \{1, \ldots, K\}$ and $a \in C_k^E$,

$$\gamma = \int a \, dP_{n_k}(\cdot|E) \quad \Longrightarrow \quad \forall m \text{ s.t. } 1_E a + 1_{E^c}\gamma \in C_m, \int [1_E a + 1_{E^c}\gamma] \, dP_m = \gamma. \tag{8}$$

A few remarks are in order. First, observe that no restriction is imposed on the unconditional plausible-priors preference $\succeq$; furthermore, the unconditional preference $\succeq_E$ is only assumed to be a weak order. Thus, the Theorem ensures that *every* plausible-priors preference relation can be uniquely updated in a manner consistent with Axioms 7 and 8; moreover, the resulting conditional preference necessarily has an analogous "plausible-posteriors" representation. Conceptually, this is perhaps the most important part of Theorem 3.2, because it indicates that *the class of plausible-priors preferences is closed under updating.*

Second, every posterior is obtained by updating one of the priors $P_1, \ldots, P_N$. However, not every plausible prior generates a plausible posterior. Intuitively, certain ex-ante plausible probabilistic models of the underlying uncertainty might have to be discarded.

Third, the condition in Eq. (8) *characterizes the posterior evaluation of a function* in terms of the prior evaluation of a related function. To clarify, consider the set $\mathcal{R}$ and the functional $V : \mathcal{R} \to \mathbb{R}$ defined after Corollary 2.11; let $\mathcal{R}_E$ and $V_E : \mathcal{R}_E \to \mathbb{R}$ be the corresponding set and functional for the conditional preference $\succeq_E$. Thus, $V((\int a \, dP_n)_{n=1,\ldots,N})$ is the unconditional evaluation of the function a, and $V_E((\int a \, dP_{n_k}(\cdot|E))_{k=1,\ldots,K})$ is its evaluation conditional upon E. Then, Eq. (8) states that, for any function $a \in B(S, \Sigma)$,

$\gamma = V_E((\int a \, dP_{n_k}(\cdot|E))_{k=1,\dots,K})$ solves the equation

$$V\left(\left(\int [1_E a + 1_{E^c}\gamma] \, dP_n\right)_{n=1,\dots,N}\right) = \gamma \qquad (9)$$

(and, as shown in the Appendix, the solution is unique).

A similar "fixed point" condition has been used as a *definition* of posterior preferences in order to derive Bayesian updating for sets of priors (cf. Jaffray [18], Pires [30] and references therein). On the other hand, Theorem 3.2 shows that Eq. (8) is a *result* of consequentialism and consistency axioms on prior and posterior preferences.

3.2 Probabilistic Sophistication implies SEU

According to the intuitive interpretation discussed in the Introduction, a multiplicity of priors arises out of the decision-maker's perception of ambiguity. However, as is well-known, preferences that admit a non-degenerate multiple-prior representation may nevertheless be *probabilistically sophisticated* in the sense of Machina and Schmeidler [24].[23] This possibility suggests that, for some preferences, a multiplicity of priors may reflect something other than a concern for ambiguity—namely, a form of "probabilistic risk aversion".

This section shows that, under suitable regularity conditions, this possibility does not arise if the axioms proposed here hold: *a probabilistically sophisticated plausible-prior preference is consistent with SEU.* In other words, under the proposed axioms, a multiplicity of priors can be safely interpreted as reflecting perceived ambiguity.

The main result of this section is true regardless of whether the decision setting under consideration features (a) roulette lotteries à la Anscombe-Aumann and objective mixtures, defined as convex combinations of such lotteries, or (b) a rich outcome space and subjective mixtures, as discussed at the end of Sec. 2.1. However, the result is mainly of interest in a fully subjective setting, as in (b): within the objective/subjective Anscombe-Aumann decision framework, Axioms 1–5 imply that preferences over lotteries are consistent with EU maximization, so a multiplicity of priors necessarily reflects ambiguity.[24] Thus, throughout

[23] For example, let $S = [0, 1]$ and consider a CEU preference $\succeq$ represented by the capacity ν given by $\nu(E) = [\lambda(E)]^2$ for all Borel sets E, where λ denotes Lebesgue measure; since ν is convex, $\succeq$ also admits a MEU representation. Incidentally, λ might perhaps be viewed as a "plausible *non-SEU* prior" for $\succeq$ (i.e. a plausible prior for a decision-maker with non-SEU risk attitudes). Thus, this example confirms that, for reasons discussed in the Introduction, Def. 2.3 aims at capturing plausible *SEU* priors.

[24] Loosely speaking, a probabilistically-sophisticated decision-maker ranks acts by "reducing" them to lotteries, and then ordering the latter by means of some preference functional V (see [25] for details). In the Anscombe-Aumann setup, Axioms 1–5 imply that V is the EU functional.

this section, we focus on a fully subjective environment where objective lotteries are not available, and hence cannot be employed to pin down the decision-maker's risk preferences independently of her perception of and attitudes towards ambiguity. Formally:

Assumption 1 (i) Acts are maps from S to the set X of prizes; (ii) there exists a *convex-ranged* function $u : X \to \mathbb{R}$, unique up to positive affine transformations, such that, for all $x, x' \in X$, $x \succeq x'$ if and only if $u(x) \geq u(x')$; (iii) there exists a mixture operator $\oplus : X \times [0,1] \times X \to X$ such that, for all $\alpha \in [0,1]$ and $x, x' \in X$, $u(\alpha x \oplus (1-\alpha)x') = \alpha u(x) + (1-\alpha)u(x')$.

As discussed in Sec. 2.1, under these assumptions, a characterization of plausible-priors preferences is obtained simply by replacing objective mixtures with the subjective mixture operator $\oplus$ in Axioms 1–5 and 6. Also, it is possible to interpret these axioms in a manner consistent with both objective and subjective mixtures; in particular, this was explicitly done in Section 2 for Axioms 5 and 6. For basic assumptions on preferences leading to properties (ii) and (iii) in Assumption 1, see the references mentioned at the end of Sec. 2.1.

An act $f \in L$ is deemed *simple* if $\{x : \exists s \in S,\ f(s) = x\}$ is finite.

Definition 3.3 A preference relation $\succeq$ is *probabilistically sophisticated* (with respect to μ) iff there exists a probability charge μ on (S, Σ) such that, for all *simple* acts $f, g \in L$,

$$\Big[\forall x \in X,\ \mu(\{s : f(s) \preceq x\}) \leq \mu(\{s : g(s) \preceq x\})\Big] \quad \Rightarrow \quad f \succeq g,$$

with strict preference if strict inequality holds for at least one $x^* \in X$.

A probabilistically sophisticated decision-maker thus ranks acts in accordance with first-order stochastic dominance with respect to a charge μ. In particular, she is indifferent among acts that induce the same distribution over prizes given μ. Furthermore, the probability μ represents her "qualitative beliefs", as revealed by preferences over binary acts.

Finally, three regularity conditions are required. First, although Def. 3.3 does not require this, the axiomatization of probabilistic sophistication provided by Machina and Schmeidler [24] delivers a *convex-ranged* probability charge μ: that is, for every $E \in \Sigma$ and $\alpha \in [0,1]$, there exists $F \in \Sigma$ such that $F \subset E$ and $\mu(F) = \alpha\mu(E)$. Proposition 3.4 below requires that μ be convex-ranged.[25] Note that this implies that S is infinite.

[25] I emphasize that the assumption that μ is convex-ranged is essential for Proposition 3.4 to hold. However, to the best of my knowledge, the only characterization of probabilistically sophisticated preferences that does *not* deliver a convex-ranged charge is [25], which utilizes objective lotteries. As noted above, the claim is trivially true under Axioms 1–5 in that setup.

Second, it is necessary to ensure that both μ and the plausible priors $P_1, \dots, P_N$ be countably additive. To this end, a version of the standard *monotone continuity* axiom (cf. e.g. Epstein and Zhang [10]) is assumed to hold. Say that a sequence of acts $\{f_k\}_{k\geq 1} \subset L$ *converges monotonely in preference* to an act $f \in L$, denoted "$f_k \downarrow f$", if and only if (i) for all k, $f_k \succeq f_{k+1}$, and (ii) for all $y \in Y$ such that $y \succ f$, there is k such that $y \succ f_k$.

Axiom 9 (Monotone Continuity) *Consider a sequence of acts $\{f_k\}_{k\geq 1} \subset L$ and an act $f \in L$. If $f_k(s) \downarrow f(s)$ for all s, then $f_k \downarrow f$.*

Third, a structural assumption on the measurable space (S, Σ) is required. Specifically, (S, Σ) is assumed to be a *standard Borel space* (cf. e.g. Kechris [19], Def. 12.5): there exists a separable and completely metrizable topology τ on S such that Σ is the Borel sigma-algebra generated by τ. All Borel subsets of Euclidean space $\mathbb{R}^n$ are standard Borel spaces, as are many spaces of functions that arise in the theory of continuous-time stochastic processes.

The main result of this section can now be stated.

Proposition 3.4 *Suppose that (S, Σ) is a standard Borel space and Assumption 1 holds; let $\succeq$ be a plausible-priors preference that satisfies Axiom 9. If $\succeq$ is probabilistically sophisticated with respect to a convex-ranged probability charge μ, then μ is the only plausible prior for $\succeq$. Consequently, $\succeq$ is a SEU preference.*

Marinacci [26] provides a related result for α-MEU preferences that satisfy a version of Monotone Continuity. Specifically, he shows that, if *all* priors in the α-MEU representation assign the same probability $p \in (0, 1)$ to some event A, then preferences are probabilistically sophisticated if and only if they are SEU. Thus, "collapses to SEU" can obtain for other preferences that satisfy Axioms 1–5.

4 Discussion

4.1 Preferences without Plausible Priors

This subsection discusses an example of MEU preferences for which plausible priors cannot be elicited, because the uniqueness requirement in Def. 2.3 cannot be satisfied. Notation and assumptions about outcomes are as in the examples of Section 2.

Example 4 (Daphne) Let $S = \{s_1, s_2, s_3\}$; Daphne is a MEU decision-maker with priors $\mathcal{Q} = \{q \in \Delta(S) : \sum_{i=1,2,3}[q(s_i) - \frac{1}{3}]^2 \leq \varepsilon^2\}$ for $\varepsilon \in (0, \frac{1}{\sqrt{6}}]$. Graphically, $\mathcal{Q}$ is a circle of radius ε in the simplex in $\mathbb{R}^3$, centered at the uniform distribution on S. Corollary 2.8 implies that

$\succeq$ is *not* a plausible-priors preference, but a stronger statement is true: no plausible prior can be elicited. Note that, for this preference, $f \simeq g$ if and only if f and g are *affinely related*, i.e. iff $f(s) = \alpha g(s) + \beta$ for some $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$.[26] Now let C be any maximal collection of affinely related acts. Note that there is a unique prior $q_C \in \mathcal{Q}$ that minimizes $\int f\,dq$ over $\mathcal{Q}$ for all $f \in C$. It is clear that C satisfies Part (i) in Def. 2.3; however, it does *not* satisfy Part (ii): any probability q on S that satisfies $\int f\,dq = \int f\,dq_C$ for a non-constant act $f \in C$ also satisfies $\int g\,dq = \int g\,dq_C$ for any other act $g \in C$, because f and g are affinely related. Thus, there exists a continuum of probabilities that represent preferences on C.

In this example $f \simeq g$ (if and) only if f and g are affinely related, so it is easy to see that only constant acts are robustly mixture neutral. Hence, Axiom 6 is violated in a relatively trivial sense. However, the preferences in Example 4 also violate much weaker assumptions. For instance, Axiom 6 implies that, whenever $f_k \to f$, there is K such that $f_k \simeq f$ for $k \geq K$ (cf. Lemma 5.8 in the Appendix). Yet, Daphne's preferences do not satisfy this property.

Since the state space is finite, the discussion preceding Axiom 6 suggests that considerations other than hedging against ambiguity determine Daphne's violations of mixture neutrality. In any case, Daphne behaves very differently from a SEU decision-maker, even "locally": mixture neutrality is violated for *any* pair of acts that are not affinely related, no matter how close in preference. By way of contrast, a plausible prior can only be elicited if the individual behaves "as if" he had unique "local" SEU preferences.

I emphasize that, even if a preference does not admit plausible priors, it may allow for alternative, behaviorally-based interpretations of sets of probabilities appearing in its representation. For instance, Wang [34] axiomatizes an entropy-based multiple-priors model. Other decision models (e.g. CEU) may have natural interpretations that are unrelated to probabilistic priors, and as such are not affected by the considerations in the Introduction.

4.2 Related Literature

4.2.1 Probabilistic Representations of Ambiguity

Sets of probabilities provide an intuitively appealing representation of ambiguity in the α-MEU decision model. Ghirardato, Maccheroni and Marinacci [11, GMM henceforth] and Nehring [29] formalize this key insight, and show that it applies to a broader class of preferences. GMM take as primitive a preference relation over acts that satisfies Axioms 1–5, and

[26] In general, f and g are affinely related if $u \circ f = \alpha u \circ g + \beta$, with α, β as above. But recall that, for all examples, $X = \{\$0, \$10\}$, so Y can be identified with $[0,1]$ and it is w.l.o.g. to assume that $u(y) = y$.

derive from it an auxiliary, incomplete relation $\succeq^*$ that is intended to capture "unambiguous" comparisons of acts; they then show that $\succeq^*$ admits a representation à la Bewley [3]: there exists a set $\mathcal{Q}$ of probability charges such that, for all acts $f, g \in L$,

$$f \succeq^* g \quad \Leftrightarrow \quad \forall Q \in \mathcal{Q}, \int u \circ f \, dQ \geq \int u \circ g \, dQ. \tag{10}$$

Loosely speaking, Nehring takes as primitive both a preference relation $\succeq$ on acts, and an incomplete unambiguous likelihood relation $\trianglerighteq$ on events; he then axiomatically relates the two, and provides a Bewley-style representation of $\trianglerighteq$ analogous to Eq. (10). Both papers suggest that a non-singleton set $\mathcal{Q}$ is associated with ambiguity; GMM and Nehring then develop these ideas in several, complementary directions.

Thus, both GMM and Nehring identify a set of probabilities that, *as a whole*, provides a *specific* representation of "unambiguous" preferences and beliefs. This is appropriate for their purposes, but does not achieve the objectives of the present paper: it is not intended to deliver priors that can be deemed "plausible" according to the stringent behavioral criteria set forth in Def. 2.3. Specifically, the identification issues highlighted in the Introduction for MEU priors apply *verbatim* to sets of probabilities in the representation of Bewley preferences such as $\succeq^*$ (and, by analogy, $\trianglerighteq$). Such sets are identified by the "functional-form" assumption that they represent $\succeq^*$ (or $\trianglerighteq$) according to Eq. (10); but, just like a MEU preference, a Bewley preference admits alternative representations, characterized by different sets of priors.[27]

These considerations do not invalidate the insight that ambiguity can be represented via sets of probabilities, or the related developments that are the main focus of GMM and Nehring. Moreover, it can be shown that, under the additional axioms provided in the present paper, the sets identified by GMM and Nehring can be obtained as the weak* closed, convex hull of the set of plausible priors delivered by Theorem 2.6. However, as in the case of MEU preferences, if Axiom 6 does not hold, the intuitive interpretation of the elements of $\mathcal{Q}$ as possible probabilistic models of the underlying uncertainty may be problematic.

Also, note that a probabilistically sophisticated preference may give rise to a non-singleton set $\mathcal{Q}$ in the GMM setup. By Proposition 3.4, this is never the case if Axiom 6 and the regularity conditions in Sec. 3.2 hold.

4.2.2 Other Related Literature

Castagnoli and Maccheroni [6] (see also [7]) explicitly assume that preferences satisfy the Independence axiom when restricted to *exogenously specified* convex sets of acts, and derive a

[27]Section 6.2 in the Online Appendix discusses Bewley preferences and provides examples.

representation analogous to Eq. (5); the corresponding probabilities are *not unique*. By way of contrast, the approach adopted here entails *deriving* a proper covering from preferences, and ensuring that the corresponding probabilities are unique.

Machina [23] investigates the robustness of "the *analytics* of the classical [i.e. SEU] model... to behavior that departs from the probability-theoretic nature of the classical paradigm." [23, p. 1; italics added for emphasis]. Among other results, Machina shows (Theorem 4, p. 34) that it is sometimes possible to associate with a specific act f_0 a *local probability measure* μ_{f_0} that represents the decision-maker's "local revealed likelihood rankings" and, jointly with a *local utility function* U_{f_0}, her response to *event-differential* changes in the act being evaluated. However, he is careful to point out that "the existence of a local probability measure μ_{f_0} at each f_0 should *not* be taken to imply the individual has conscious probabilistic beliefs that somehow depend upon the act(s) being evaluated." (p. 35; italics in the original). This is fully consistent with the point of view advocated in the present paper: a probability μ can be a useful *analytical* tool to model certain properties (e.g. responses to differential changes) of the mathematical representation of preferences; however, for μ to be deemed a "plausible prior", additional behavioral conditions must be met.

5 Appendix

5.1 Proof of Theorem 2.6.

5.1.1 Numerical Representation of preferences and restatement of the axioms

Most proofs for this subsection are in the Online Appendix.

Lemma 5.1 *The preference relation $\succeq$ satisfies Axioms 1, 2, 3, 4 and 5 if and only if there exists a non-constant affine function $u : Y \to \mathbb{R}$, unique up to positive linear transformations, and a unique, normalized, monotonic and c-linear functional $I : B(S, \Sigma) \to \mathbb{R}$, such that, for all $f, g \in L$, $f \succeq g$ iff $I(u \circ f) \geq I(u \circ g)$. Furthermore, u can be chosen so $u(Y) \supset [-1, 1]$. Finally, for all $a, b \in B(S, \Sigma)$, $|I(a) - I(b)| \leq \|a - b\|$.*

Throughout the remainder of the appendix, u and I denote a utility function and, respectively, a functional, with the properties indicated in Lemma 5.1.

Abusing notation, for functions $a, b \in B(S, \Sigma)$, $a \simeq b$ iff $I(\alpha a + (1 - \alpha)b) = \alpha I(a) + (1 - \alpha)I(b)$ for all $\alpha \in [0, 1]$. Similarly, $a \approx b$ iff, for every sequence $c_k \subset B(S, \Sigma)$ that supnorm-converges to either f or g, and all $\gamma \in [0, 1]$, there exists K such that $c_k \simeq \gamma f + (1 - \gamma)g$.

Lemma 5.2 *Suppose $\succeq$ satisfies Axioms 1, 2, 3, 4 and 5, and let I, u be its representation.*

1. *For all $\{f_k\} \subset L$ and $f \in L$, $f_k \to f$ iff $u \circ f_k \to u \circ f$ in $B(S, \Sigma)$.*
2. *For all $f, g \in L$, $f \simeq g$ iff $u \circ f \simeq u \circ g$.*
3. *For all $a, b \in B(S, \Sigma)$, and $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$: $a \simeq b$ implies $a \simeq \alpha b + \beta$.*
4. *For all $a, b \in B(S, \Sigma)$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha, \beta \geq 0$: $a \simeq b$ implies $a \simeq \alpha a + \beta b$.*
5. *For all sequences $\{a_k\}, \{b_k\} \subset B(S, \Sigma)$ such that $a_k \to a$ and $a_k \to b$ for $a, b \in B(S, \Sigma)$: $a_k \simeq b_k$ for all k implies $a \simeq b$.*
6. *For all $f, g \in L$, $f \approx g$ iff $u \circ f \approx u \circ g$.*
7. *For all $a, b \in B(S, \Sigma)$: $a \approx b$ implies $a \simeq b$.*
8. *For all $a, b \in B(S, \Sigma)$, and $\alpha, \beta \in \mathbb{R}$ with $\alpha > 0$: $a \approx b$ implies $a \approx \alpha b + \beta$.*
9. *For all $a, b \in B(S, \Sigma)$, and $\lambda \in (0, 1)$: $a \approx b$ implies $\lambda a + (1 - \lambda)b \approx b$.*
10. *For all $a, b \in B(S, \Sigma)$: $a \approx b$ iff, for every $\gamma \in [0, 1]$, there exists $\epsilon > 0$ such that $\|c - a\| < \epsilon$ or $\|c - b\| < \epsilon$ imply $c \simeq \gamma a + (1 - \gamma)b$.*

Corollary 5.3 *Under the conditions of Lemma 5.2, the relation $\succeq$ satisfies Axiom 6 if and only if, for all $\{a_k\} \subset B(S, \Sigma)$ and $a \in B(S, \Sigma)$ such that $a_k \to a$, there exists a subsequence $\{a_{k(\ell)}\}$ such that $a_{k(\ell)} \approx a_{k(\ell')}$ for all ℓ, ℓ'.*

Proof. Suppose $f_k \to f$ in L; then, by Part 1 in Lemma 5.2, $u \circ f_k \to u \circ f$ in $B(S, \Sigma)$; if the property in the Corollary holds, there is a subsequence such that $u \circ f_{k(\ell)} \approx u \circ f_{k(\ell')}$ for all ℓ, ℓ'; by Part 6 in the Lemma, this implies $f_{k(\ell)} \approx f_{k(\ell')}$.

In the opposite direction, consider $a_k \to a$ in $B(S, \Sigma)$. Since $\{a_k\}$ converges and a is bounded, there exist $\gamma, \gamma' \in \mathbb{R}$ such that $\gamma \geq a_k(s) \geq \gamma'$ for all k and s, and similarly $\gamma \geq a(s) \geq \gamma'$ for all s. There exists $\alpha > 0$ such that $\alpha\gamma, \alpha\gamma' \in [-1, 1]$; for this α, there exists $\{f_k\} \subset L$ and $f \in L$ such that $u \circ f_k = \alpha a_k$ for all k, and $u \circ f = \alpha a$. Clearly, $\alpha a_k \to \alpha a$; by Part 1 in the Lemma, this implies that $f_k \to f$. Now Axiom 6 implies that there is a subsequence such that $f_{k(\ell)} \approx f_{k(\ell')}$, hence $\alpha a_{k(\ell)} \approx \alpha a_{k(\ell')}$ by Part 6, for all ℓ, ℓ'. Now Part 8 yields the required conclusion. ■

In light of the above Lemma and Corollary, the analysis will henceforth focus on the properties and representation of the functional I on $B(S, \Sigma)$. To streamline the exposition, expressions such as “by Axiom 6 and the Corollary to Lemma 5.2, there exists a subsequence $\{a_{k(\ell)}\}$ such that …” will be shortened to “by Axiom 6, there exists a subsequence …”.

5.1.2 Proof of Remark 1. Necessity of the Axioms: preliminaries

Definition 5.4 (cf. Property 4 of Def. 2.5) A set $C \subset B(S,\Sigma)$ is **minimally convex** iff every infinite subset $C' \subset C$ contains a countable collection $\{a_k\}_{k\geq 1} \subset C'$ with the property that, for all $k, \ell \geq 1$ and $\gamma \in [0,1]$, there exists $\epsilon > 0$ such that, for all $b \in B(S,\Sigma)$,

$$\min(\|b - a_k\|, \|b - a_\ell\|) < \epsilon,\ b \in C\ \Rightarrow\ \forall \lambda \in (0,1),\ \lambda b + (1-\lambda)[\gamma a_k + (1-\gamma) a_\ell] \in C.$$

Proof of Remark 1. Let $C_n = C_{n,1} \cup \ldots \cup C_{n,M}$, where each $C_{n,m}$ is closed and convex. Fix an infinite $C' \subset C_n$. The collection

$$\left\{ \bigcap_{m\in\mathcal{M}} C_{n,m} \cap \bigcap_{m\in\{1,\ldots,M\}\setminus\mathcal{M}} B(S,\Sigma) \setminus C_{n,m}\ :\ \mathcal{M} \subset \{1,\ldots,M\} \right\}$$

is a finite partition of $B(S,\Sigma)$, so there is $\mathcal{M} \subset \{1,\ldots,N\}$ such that the set $\{a \in C' : a \in C_{n,m} \Leftrightarrow m \in \mathcal{M}\}$ is infinite. In turn, this set contains a countably infinite collection $\{a_k\}$.

Now fix k, k' and consider $a_k, a_{k'}$. There exists $\epsilon > 0$ such that $\|c - a_k\| < \epsilon$ and $c \in C_n$ implies $c \in C_{n,m}$ for some $m \in \mathcal{M}$. To see this, suppose that, for all ℓ, there is c_ℓ such that $\|c_\ell - a_k\| < \frac{1}{\ell}$ and $c_\ell \in C_{n,m(\ell)}$ for some $m(\ell) \notin \mathcal{M}$. Then there is a subsequence of $\{c_\ell\}$ that lies in some $C_{n,m}$ with $m \notin \mathcal{M}$, and this subsequence converges to $a_k \notin C_{n,m}$; this contradicts the fact that $C_{n,m}$ is closed.

Fix such $\epsilon > 0$ and c such that $\|c - a\| < \epsilon$ and $c \in C_n$, so $c \in C_{n,m}$ for some $m \in \mathcal{M}$. By construction, $a_k, a_{k'} \in C_{n,m}$, so for all $\gamma \in [0,1]$, $\gamma a_k + (1-\gamma) a_{k'} \in C_{n,m}$ as $C_{n,m}$ is convex; hence, for the same reason, $\lambda c + (1-\lambda)[\gamma a_k + (1-\gamma) a_{k'}] \in C_{n,m} \subset C_n$ for all $\lambda \in [0,1]$. ∎

Lemma 5.5 *Suppose that $C_1, \ldots, C_N$ is a proper covering of $B(S,\Sigma)$. Let $a, b \in B(S,\Sigma)$. Then, for some $K \geq 1$, there exists a finite collection $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_K = 1$ such that, for each $k = 0, ..., K-1$, there exists $n_k \in \{1,\ldots,N\}$ such that $\alpha a + (1-\alpha) b) \in C_{n_k}$ for all $\alpha \in [\alpha_k, \alpha_{k+1}]$.*

Proof. Let $\alpha_0 = 0$. Proceeding by induction, assume that $\alpha_0, \ldots, \alpha_{k-1}$ as above have been defined, for some $k > 0$, and that $\alpha_{k-1} < 1$. For every $n = 1, \ldots, N$, let

$$A(n, k-1) = \{\alpha' \in [\alpha_{k-1}, 1] : \forall \alpha \in [\alpha_{k-1}, \alpha'], \alpha a + (1-\alpha) b \in C_n\}.$$

For every n, if $A(n, k-1)$ is non-empty, it is of the form $[\alpha_{k-1}, \alpha_{n,k}]$ for some $\alpha_{n,k} \geq \alpha_{k-1}$.[28]

[28] It is clear that, if $\alpha' \in A(n, k-1)$, then $[\alpha_{k-1}, \alpha'] \subset A(n, k-1)$. Furthermore, since C_n is closed, $\sup A(n, k-1) \in A(n, k-1)$.

There exists $n_k \in \{1,\dots,N\}$ such that $A(n_k,k-1)$ is non-empty, and indeed $\alpha_{n,k} > \alpha_{k-1}$. To see this, consider the sequence $\{\beta_\ell\}$ defined by $\beta_\ell = \frac{1}{\ell} + (1-\frac{1}{\ell})\alpha_{k-1}$, so $\beta_\ell \to \alpha_{k-1}$. Then, for some subsequence $\{\beta_{\ell(r)}\}$ and some $n_k \in \{1,\dots,N\}$, $\beta_{\ell(r)}a + (1-\beta_{\ell(r)})b \in C_{n_k}$ for all r. Minimal convexity implies that there is a further subsequence $\{\beta_{\ell(r(\rho))}\}$ such that, in particular, for all $\lambda \in [0,1]$, and all ρ,ρ', $\lambda[\beta_{\ell(r(\rho))}a + (1-\beta_{\ell(r(\rho))})b] + (1-\lambda)[\beta_{\ell(r(\rho'))}a + (1-\beta_{\ell(r(\rho'))})b] \in C_{n_k}$.[29] For any $\lambda \in [0,1]$, fixing $\rho = 1$ and letting $\rho' \to \infty$, since C_{n_k} is closed and $\beta_{\ell(r(\rho'))} \to \alpha_{k-1}$, one obtains $\lambda[\beta_{\ell(r(1))}a+(1-\beta_{\ell(r(1))})b]+(1-\lambda)[\alpha_{k-1}a+(1-\alpha_{k-1}b] \in C_{n_k}$.

Hence, $A(n_k,k-1) \supset [\alpha_{k-1},\beta_{\ell(r(1))}]$, so there exists $\alpha_{n,k} \ge \beta_{\ell(r(1))} > \alpha_{k-1}$ such that $A(n_k,k-1) = [\alpha_{k-1},\alpha_{n,k}]$. Finally, define $\alpha_k = \max_{n:A(n,k-1)\neq\emptyset} A(n,k-1)$; the argument just given shows that $\alpha_k > \alpha_{k-1}$. Now suppose that $\alpha_k < 1$ for all k; then $\alpha_k \uparrow \bar\alpha \in [0,1]$, so $\alpha_k a + (1-\alpha_k)b \to \bar\alpha a + (1-\bar\alpha)b$. As above, there is a subsequence $\{\alpha_{k(\ell)}\}$ and an index $n \in \{1,\dots,N\}$ such that $\alpha_{k(\ell)}a + (1-\alpha_{k(\ell)})b \in C_n$ for all ℓ, and minimal convexity yields a further subsequence $\{\alpha_{k(\ell(r))}\}$ such that, for all $\lambda \in [0,1]$ and all r,r', $\lambda[\alpha_{k(\ell(r))}a + (1-\alpha_{k(\ell(r))})b] + (1-\lambda)[\alpha_{k(\ell(r'))}a + (1-\alpha_{k(\ell(r'))})b] \in C_n$; again, letting $r=1$ and $r' \to \infty$, since C_n is closed, one gets $\lambda[\alpha_{k(\ell(1))}a + (1-\alpha_{k(\ell(1))})b] + (1-\lambda)[\bar\alpha a + (1-\bar\alpha)b] \in C_n$ for all λ; but this contradicts the fact that $\alpha_{k(\ell(1))+1} < \bar\alpha$. This proves the claim. ■

5.1.3 Necessity of the Axioms: Completing the argument

Let $u, C_1,\dots,C_N$ and $P_1,\dots,P_N$ be as in Statement 2 of Theorem 2.6. Then $u \circ f \in C_n \cap C_m$ implies that $\int u \circ f\, dP_n = \int u \circ f\, dP_m$. Since every C_n is closed under non-negative affine transformations ("affine" henceforth), this holds for all $a \in B(S,\Sigma)$. Hence, one can define $I : B(S,\Sigma) \to \mathbb{R}$ by letting $I(a) = \int a\, dP_n$ for $a \in C_n$. Then (I,u) represent $\succeq$. It is possible, of course, to assume that $u(Y) \supset [-1,1]$. Furthermore, since each P_n is unique, so is I.

I is c-linear. Let $a \in B(S,\Sigma)$, $\beta \in \mathbb{R}_+$ and $\gamma \in \mathbb{R}$. Since each C_n is affine, $\beta a + \gamma \in C_n$ implies $a = \frac{\beta a+\gamma}{\beta} - \frac{\gamma}{\beta} \in C_n$; hence, $I(\beta a + \gamma) = \int (\beta a + \gamma)\, dP_n = \beta \int a\, dP_n + \gamma = \beta I(a) + \gamma$.

I is monotonic. Let $a,b \in B(S,\Sigma)$ be such that $a(s) \ge b(s)$ for all s; then $\int a\, dP_n \ge \int b\, dP_n$ for all $n \in \{1,\dots,N\}$. Let K, α_k and n_k be as in Lemma 5.5; then for $k = 0,\dots,K-1$, $\int [\alpha_k a + (1-\alpha_k)b]\, dP_{n_k} \le \int [\alpha_{k+1}a + (1-\alpha)_{k+1}b]\, dP_{n_k} = \int [\alpha_{k+1}a + (1-\alpha)_{k+1}b]\, dP_{n_{k+1}}$: the inequality follows from $\int a\, dP_{n_k} \ge \int b\, dP_{n_k}$ and $\alpha_k < \alpha_{k+1}$, and the equality holds because $\alpha_{k+1}a + (1-\alpha_{k+1})b \in C_{n_k} \cap C_{n_{k+1}}$. Thus, $I(\alpha_k a + (1-\alpha_k)b) \le I(\alpha_{k+1}a + (1-\alpha_{k+1})b)$ for all $k = 0,\dots,K-1$; since $\alpha_0 = 0$ and $\alpha_K = 1$, $I(b) \le I(a)$. Clearly, I is also *normalized*, so Lemma 5.1 implies that I satisfies Axioms 1, 2, 3, 4, and 5.

[29]To clarify: in Def. 5.4, take "a_k" and "a_ℓ" to be the mixtures corresponding to ρ and ρ', let "b" be the mixture corresponding to ρ, and let $\gamma = 0$. Then minimal convexity implies the claim for all $\lambda \in (0,1)$, and since C_{n_k} is closed, the claim is also true for $\lambda = 0, 1$.

To see that Axiom 6 holds, suppose $a_k \to a$ in $B(S,\Sigma)$. As in the proof of Remark 1, since there are only finitely many intersections of distinct elements of the collection $\mathcal{C} = \{C_1, \ldots, C_N\}$, there exists a subsequence $\{a_{k_0(\ell)}\}$ such that, for all $n = 1, \ldots, N$ and ℓ, ℓ', $a_{k_0(\ell)} \in C_n$ iff $a_{k_0(\ell')} \in C_n$. Without loss of generality, suppose that $a_{k_0(1)} \in C_1 \cap \ldots \cap C_M$, for some $M \leq N$. Define further subsequences by induction, as follows. For $n = 1, \ldots, N$, observe that $\{a_{k_{n-1}(\ell)}\}$ is an infinite subset of C_n; since C_n is minimally convex, there is a subsequence $\{a_{k_n(\ell)}\} \subset \{a_{k_{n-1}(\ell)}\}$ that satisfies the condition in Definition 5.4. After M steps, this procedure defines a sequence $\{a_{k_M(\ell)}\}$—a subsequence of $\{a_k\}$. Now fix ℓ, ℓ'; let $b = a_{k_M(\ell)}$ and $b' = a_{k_M(\ell')}$. To complete the proof, it must be shown that $b \approx b'$.

Fix $\gamma \in [0,1]$ arbitrarily: by Lemma 5.2 Part 10, it is sufficient to show that there exists $\epsilon > 0$ such that $\|c - b\| < \epsilon$ implies $c \simeq \gamma b + (1-\gamma)b'$ [the argument for c such that $\|c - b'\| < \epsilon$ is identical]. Notice first that there exists $\epsilon_0 > 0$ such that $\|c - b\| < \epsilon_0$ implies $c \in C_n$ for some $n \in \{1, \ldots, M\}$; otherwise, as in the proof of Remark 1, one could find $m > M$ and a sequence $\{c_k\} \subset C_m$ such that $c_k \to b$; since $c_k \notin C_m$, this would contradict the fact that C_m is closed. Furthermore, due to the above construction, for every $n = 1, \ldots, M$, there exists $\epsilon_n > 0$ for which the property in Definition 5.4 holds for $C = C_n$. Let $\epsilon = \min(\epsilon_0, \ldots, \epsilon_M) > 0$.

Finally, consider any $c \in B(S,\Sigma)$ such that $\|c-b\| < \epsilon$. Since $\|c-b\| < \epsilon_0$, $c \in C_n$ for some $n \in \{1, \ldots, M\}$; and since $\|c-b\| < \epsilon_n$, it is the case that $\lambda c + (1-\lambda)[\gamma b + (1-\gamma)b'] \in C_n$ for all $\lambda \in (0,1)$, and therefore also for $\lambda = 0, 1$ because C_n is closed; that is, $\gamma b + (1-\gamma)b' \in C_n$. But this implies that

$$
\begin{aligned}
&I(\lambda c + (1-\lambda)[\gamma b + (1-\gamma)b']) = P_n(\lambda c + (1-\lambda)[\gamma b + (1-\gamma)b']) = \\
&= \lambda P_n(c) + (1-\lambda)P_n(\gamma b + (1-\gamma)b') = \lambda I(c) + (1-\lambda)I(\gamma b + (1-\gamma)b'),
\end{aligned}
$$

i.e. $c \simeq \gamma b + (1-\gamma)b'$. Thus, Axiom 6 holds.

5.1.4 Sufficiency: first covering of $B(S,\Sigma)$ and other implications of Axiom 6

Lemma 5.6 *Assume that Axioms 1–5 and 6 hold. There exists a finite collection $\{c_1, \ldots, c_{N_1}\} \subset B(S,\Sigma)$ such that 1. for all $n, m \in \{1, \ldots, N_1\}$ such that $n \neq m$, $c_n \not\approx c_m$; and 2. for all $a \in B(S,\Sigma)$, there exists $n \in \{1, \ldots, N_1\}$ such that $a \approx c_n$.*

Proof. Consider the following procedure. At step 1, let c_1 be an arbitrary point of $B(S,\Sigma)$. Now consider step $n > 1$ and assume that $c_1, \ldots, c_{n-1}$ such that Property 1 above holds have been defined. If, for all $a \in B(S,\Sigma)$, there is $m \in \{1, \ldots, n-1\}$ such that $a \approx c_m$, then stop; otherwise, let $c_n \in B(S,\Sigma)$ be such that $c_n \not\approx c_m$ for all $m \in \{1, \ldots, n-1\}$. This procedure must stop in finitely many steps. Suppose not: then the procedure yields a sequence $\{c_n\}$; letting $c'_n = \frac{1}{n\|c_n\|}h_n$ if $c_n \neq 0$, and $c'_n = 0$ otherwise, yields a sequence converging to 0. Axiom 6 implies in particular that there are n, m such that $c'_n \approx c'_m$; by Lemma 5.2 Part 8, this implies $c_n \approx c_m$. But this contradicts the construction of the sequence $\{c_n\}$, so the above procedure must terminate in finitely many steps, thereby yielding a collection with the required properties. ■

Corollary 5.7 *Let $C_n^1 = \{a \in B(S,\Sigma) : a \approx c_n\}$, $n = 1, \ldots, N_1$: then $\bigcup_{n=1}^{N_1} C_n^1 = B(S,\Sigma)$.*

Lemma 5.8 *Assume that Axioms 1–5 and 6 hold. Then, if $\{a_k\} \to a$ in $B(S,\Sigma)$, there exists $K \geq 1$ such that $k \geq K$ implies $a_k \simeq a$. Consequently, for all a, b, there is $\alpha \in (0,1]$ such that $\alpha a + (1-\alpha)b \simeq b$; furthermore, for all $a \in B(S,\Sigma)$, $a \approx a$.*

Proof. Arguing by contradiction, suppose that, for every $\ell \geq 1$, there is $k(\ell) \geq \ell$ such that $a_{k(\ell)} \not\simeq a$. This yields a subsequence $a_{k(\ell)} \to a$. Axiom 6 implies that there is a further subsequence $\{a_{k(\ell(r))}\}$ such that, for all r, r', $a_{k(\ell(r))} \approx a_{k(\ell(r))}$. In particular, $a_{k(\ell(1))} \approx a_{k(\ell(r))}$ for all $r > 1$. By Part 7 in Lemma 5.2, this implies $a_{k(\ell(1))} \simeq a_{k(\ell(r))}$ for all $r > 1$. Since $a_{k(\ell(r))} \to a$, by Part 5 in the same Lemma, this implies $a_{k(\ell(1))} \simeq a$. This is a contradiction, because $\{a_{k(\ell(r))}\}$ is a subsequence of $\{a_{k(\ell)}\}$, which was chosen so that $a_{k(\ell)} \not\simeq a$ for all ℓ.

The second claim follows by considering $a_k = \frac{1}{k}a + \frac{k-1}{k}b \to b$. The third claim is clear from the definition of $\approx$. ■

Remark 2 Assume that Axioms 1–5 and 6 hold. Then every set C_n^1 defined in Corollary 5.7 satisfies the following properties.

1. If $a \in C_n^1$, $\alpha, \beta \in \mathbb{R}$, and $\alpha > 0$, then $\alpha a + \beta \in C_n^1$.
2. $c_n \in C_n^1$; furthermore, for every $a \in C_n^1$ and $\alpha \in [0,1]$, $\alpha a + (1-\alpha)c_n \in C_n^1$.

Proof. Part 1 follows from Lemma 5.2 Part 8. In Part 2, $c_n \in C_n^1$ follows from Lemma 5.8, and the other claim follows from $a \approx c_n$ and Lemma 5.2 Part 9. ■

5.1.5 Sufficiency: Representation of I on C_n^1

The next step is to show that the restriction of I to each set C_n^1 coincides with the restriction to the same set of some linear functional P_n. The following Lemma provides the key step.

Lemma 5.9 *Assume $\succeq$ satisfies Axioms 1—5 and 6. For every n, if $b_1, \ldots, b_M \in C_n^1$ and $a = \sum_{m=1}^M \lambda_m b_m$ for weights $\lambda_m > 0$ such that $\sum_{m=1}^M \lambda_m = 1$, then there exists $\alpha \in (0,1]$ such that $\alpha a + (1-\alpha)c_n \simeq c_n$ and $I(\alpha a + (1-\alpha)c_n) = \sum_{m=1}^M \lambda_m I(\alpha b_m + (1-\alpha)c_n)$.*

Corollary 5.10 *If $\alpha' \in (0,\alpha)$, then $\alpha' a + (1-\alpha')c_n \simeq c_n$ and $I(\alpha' a + (1-\alpha')c_n) = \sum_{m=1}^M \lambda_m I(\alpha' b_m + (1-\alpha')c_n)$.*

Proof. Begin with the Corollary; let $\gamma = \frac{\alpha'}{\alpha} \in (0,1)$. Then $\alpha' a + (1-\alpha')c_n = \gamma[\alpha a + (1-\alpha)c_n] + (1-\gamma)c_n$ and $\alpha' b_m + (1-\alpha')c_n = \gamma[\alpha b_m + (1-\alpha)c_n] + (1-\gamma)c_n$; hence, $\alpha' a + (1-\alpha')c_n \simeq c_n$ by Lemma 5.2 Part 4. Furthermore,

$$\begin{aligned} & I(\alpha' a + (1-\alpha')c_n) = \gamma I(\alpha a + (1-\alpha)c_n) + (1-\gamma)I(c_n) = \\ & = \gamma \sum_{m=1}^M \lambda_m I(\alpha b_m + (1-\alpha)c_n) + (1-\gamma)I(c_n) = \\ & = \sum_{m=1}^M \lambda_m [\gamma I(\alpha b_m + (1-\alpha)c_n) + (1-\gamma)I(c_n)] = \sum_{m=1}^M \lambda_m I(\alpha' b_m + (1-\alpha')c_n), \end{aligned}$$

where the last equality follows from $b_m \approx c_n$, hence $b_m \simeq c_n$ by Lemma 5.2 Part 7, hence $\alpha b_m + (1-\alpha)c_n \simeq c_n$ by Part 4 in the same Lemma.

Now turn to the proof of Lemma 5.9. The claim is true for $M = 1$: in this case, it must be the case that $\lambda_1 = 1$, so $a \in C_n^1$, and hence $a \approx c_n$ by definition; Lemma 5.2 Part 7 then implies that $a \simeq c_n$, and the second claim in the Lemma is trivially true. Arguing by induction, consider $M > 1$ and assume that the claim is true for $M-1$. Consider $b_1, \ldots, b_M \in C_n^1$ and $a = \sum_{m=1}^M \lambda_m b_m$ as above; also, let $b_{-1} = \sum_{m=2}^M \frac{\lambda_m}{1-\lambda_1} b_m$; by the induction hypothesis, there exists $\alpha \in (0,1]$ such that $\alpha b_{-1} + (1-\alpha)c_n \simeq c_n$ and $I(\alpha b_{-1} + (1-\alpha)c_n) = \sum_{m=2}^M \frac{\lambda_m}{1-\lambda_1} I(\alpha b_m + (1-\alpha)c_n)$. By Corollary 5.10, for every $\alpha' \in (0,\alpha]$, $\alpha' b_{-1} + (1-\alpha')c_n \simeq c_n$ and $I(\alpha' b_{-1} + (1-\alpha')c_n) = \sum_{m=2}^M \frac{\lambda_m}{1-\lambda_1} I(\alpha' b_m + (1-\alpha')c_n)$; furthermore, since $b_1 \approx c_n$, for any such α', by Lemma 5.2 Part 9, $\alpha' b_1 + (1-\alpha')c_n \approx c_n$. Finally, note that, for every α',

$$\begin{aligned} & \lambda_1[\alpha' b_1 + (1-\alpha')c_n] + (1-\lambda_1)[\alpha' b_{-1} + (1-\alpha')c_n] = \\ & = \alpha'[\lambda_1 b_1 + (1-\lambda_1)b_{-1}] + (1-\alpha')c_n = \alpha' a + (1-\alpha')c_n; \end{aligned}$$

therefore, by Lemma 5.8, for some sufficiently small such α', $\lambda_1[\alpha' b_1 + (1-\alpha')c_n] + (1-\lambda_1)[\alpha' b_{-1} + (1-\alpha')c_n] \simeq c_n$. Thus, fix one such small $\alpha^* \in (0,\alpha)$. It is convenient to make the following definitions to simplify the notation: $B_1 = \alpha^* b_1 + (1-\alpha^*)c_n$, $B_2 = \alpha^* b_{-1} + (1-\alpha^*)c_n$, and $A = \lambda_1 B_1 + (1-\lambda_1)B_2 = \alpha^* a + (1-\alpha^*)c_n$. Thus, we have $B_2 \simeq c_n$, $A \simeq c_n$, and $B_1 \approx c_n$. Hence, for this α^*, the first claim of the Lemma holds, i.e. $\alpha^* a + (1-\alpha^*)c_n = A \simeq c_n$.

Similarly,

$$\lambda_1 I(B_1) + (1-\lambda_1)I(B_2) = \lambda_1 I(\alpha^* b_1 + (1-\alpha^* c_n) + (1-\lambda_1)I(\alpha^* b_{-1} + (1-\alpha^*)c_n) =$$

$$=\lambda_1 I(\alpha^* b_1 + (1-\alpha^* c_n) + (1-\lambda_1)\sum_{m=2}^{M} \frac{\lambda_m}{1-\lambda_1} I(\alpha^* b_m + (1-\alpha^*)c_n) =$$

$$=\sum_{m=1}^{M} \lambda_m I(\alpha^* b_m + (1-\alpha^*)c_n)$$

where the third equality follows from the induction hypothesis, as above; thus, to complete the proof, it suffices to show that $I(A) = \lambda_1 I(B_1) + (1-\lambda_2)I(B_2)$.

Since $B_1 \approx c_n$, there is $\beta \in (0,1)$ such that $\beta B_2 + (1-\ \beta)c_n \simeq B_1$; that is,

$$\begin{aligned} I(\gamma B_1 + (1-\gamma)[\beta B_2 + (1-\beta)c_n]) &= \gamma I(B_1) + (1-\gamma)I(\beta B_2 + (1-\beta)c_n) = \\ &= \gamma I(B_1) + (1-\gamma)\beta I(B_2) + (1-\gamma)(1-\beta)I(c_n) \end{aligned}$$

for all $\gamma \in [0,1]$, where the second equality uses the fact that α^* above was chosen so that $B_2 \simeq c_n$. Now consider γ such that $\frac{\gamma}{\gamma+(1-\gamma)\beta} = \lambda_1$, i.e. $\gamma = \frac{\lambda_1\beta}{1-\lambda_1+\lambda_1\beta} \in (0,1)$; then, letting $\Gamma = \gamma + (1-\gamma)\beta \in (0,1)$, we can rewrite the above equation as

$$I(\Gamma[\lambda_1 B_1 + (1-\lambda_1)B_2] + (1-\Gamma)c_n) = \Gamma[\lambda_1 I(B_1) + (1-\lambda_1)I(B_2)] + (1-\Gamma)I(c_n);$$

but, since $\lambda_1 B_1 + (1-\lambda_1)B_2 = A \simeq c_n$, we also have

$$I(\Gamma[\lambda_1 B_1 + (1-\lambda_1)B_2] + (1-\Gamma)c_n) = \Gamma I(\lambda_1 B_1 + (1-\lambda_1)B_2) + (1-\Gamma)I(c_n) :$$

thus, $I(A) = I(\lambda_1 B_1 + (1-\lambda_1)B_2) = \lambda_1 I(B_1) + (1-\lambda_1)I(B_2)$, as required. ∎

Remark 3 Assume that $\succeq$ satisfies Axioms 1–6. Consider $a, b \in B(S,\Sigma)$ such that $a = \sum_{\ell=1}^{M_a} \alpha_\ell a_\ell$ and $b = \sum_{m=1}^{M_b} \beta_m b_m$, where all $a_\ell, b_m, \alpha_\ell, \beta_m$ are as in Lemma 5.9, and $a \geq b$. Then $\sum_{\ell=1}^{M_a} \alpha_\ell I(a_\ell) \geq \sum_{m=1}^{M_b} \beta_m I(b_m)$.

Proof. By Lemma 5.9, there exist $\alpha, \beta \in (0,1]$ such that $I(\alpha a + (1-\alpha)c_n) = \sum_{\ell=1}^{M_a} \alpha_\ell I(\alpha a_\ell + (1-\alpha)c_n) = \alpha \sum_{\ell=1}^{M_a} \alpha_\ell I(a_\ell) + (1-\alpha)I(c_n)$, because as usual $a_\ell \simeq c_n$, and similarly $I(\beta b + (1-\beta)c_n) = \beta \sum_{m=1}^{M_b} \beta_m I(b_m) + (1-\beta)I(c_n)$.

Suppose $\alpha > \beta$. Then, by Corollary 5.10, it is also the case that $I(\beta a + (1-\beta)c_n) = \beta \sum_{\ell=1}^{M_a} \alpha_\ell I(a_\ell) + (1-\beta)I(c_n)$. Since $a \geq b$, $\beta a + (1-\beta)c_n \geq \beta b + (1-\beta)c_n$; since I is monotonic, $I(\beta a + (1-\beta)c_n) \geq I(\beta b + (1-\beta)c_n)$, and therefore $\beta \sum_{\ell=1}^{M_a} \alpha_\ell I(a_\ell) + (1-\beta)I(c_n) \geq \sum_{m=1}^{M_b} \beta_m I(b_m) + (1-\beta)I(c_n)$: and since $\beta \in (0,1)$, $\sum_{\ell=1}^{M_a} \alpha_\ell I(a_\ell) \geq \sum_{\ell=1}^{M_b} \beta_m I(b_m)$.

The case $\alpha \leq \beta$ is handled symmetrically; if $\alpha = \beta$, Corollary 5.10 is not needed. ■

For any set $C \subset B(S, \Sigma)$, let $\operatorname{conv} C$ and $\operatorname{cl} C$ denote the convex hull and sup-norm closure of C respectively. It is now possible to state the main result of this subsection.

Lemma 5.11 *Assume that $\succeq$ satisfies Axioms 1–5 and 6. For every $n = 1, \ldots, N_1$, there exists a probability charge P_n on (S, Σ) such that, for all $a \in C^1_n$, $I(a) = \int a\, dP_n$.*

In the following, it will be convenient to denote the integral $\int a\, dP_n$ simply by $P_n(a)$.

Proof. For $n = 1, \ldots, N_1$, let $C^2_n = \{\gamma : \gamma \in \mathbb{R}\} \cup \operatorname{conv} C^1_n$, where, with the usual abuse of notation, constant functions are identified with scalars. Note that C^2_n is convex: in particular, suppose that $a \in \operatorname{conv} C^1_n$, so $a = \sum_m \lambda_m b_m$ for suitable points $b_m \in C^1_n$ and positive weights λ_m, with $\sum_m \lambda_m = 1$; then, for all $\alpha \in (0,1)$ and $\gamma \in \mathbb{R}$, $\alpha a + (1-\alpha)\gamma = \alpha\left(\sum_m \lambda b_m\right) + (1-\alpha)\gamma = \sum_m \lambda_m[\alpha b_m + (1-\alpha)\gamma] \in \operatorname{conv} C^1_n$, because $\alpha b_m + (1-\alpha)\gamma \in C^1_n$ by Remark 2 Part 1. The latter result, together with the fact that $0 \in C^2_n$, also implies that C^2_n is closed under multiplication by a non-negative scalar, i.e. it is a wedge; therefore, $C^2_n - C^2_n = \{a - b : a, b \in C^2_n\}$ is a linear subspace of $B(S, \Sigma)$ (cf. Holmes [17], §5.A).

Now define a functional $I_n : C^2_n \to \mathbb{R}$ by $I_n(\gamma) = \gamma$ for all $\gamma \in \mathbb{R}$ and

$$\forall b_1, \ldots, b_M \in C^1_n,\ \lambda_1, \ldots, \lambda_M \in (0,1]\ \text{s.t.} \sum_m \lambda_m = 1, \quad I_n(\sum_m \lambda_m b_m) = \sum_m \lambda_m I(b_m). \quad (11)$$

The functional I_n is *well-defined*. First, Remark 3 ensures that $\sum_{\ell=1}^{M_a} \alpha_\ell I(a_\ell) = \sum_{m=1}^{M_b} \beta_m I(b_m)$ whenever $\sum_{\ell=1}^{M_a} \alpha_\ell a_\ell = \sum_{m=1}^{M_b} \beta_m b_m$ and all $a_\ell, b_m, \alpha_\ell, \beta_m$ are as in Lemma 5.9. Second, suppose $\sum_m \lambda_m b_m = \gamma$ (a constant function) for λ_m, b_m as in Eq. 11; by Lemma 5.9, there is $\alpha \in (0,1]$ such that $I(\alpha\gamma + (1-\alpha)c_n) = \sum_m \lambda_m I(\alpha b_m + (1-\alpha)c_n) = \alpha \sum_m \lambda_m I(b_m) + (1-\alpha) I(c_n)$, where the last equality follows from $b_m \simeq c_n$, which is implied by $b_m \approx c_n$; but by c-linearity of I, this is readily seen to imply that $\sum_m \lambda_m I(b_m) = \gamma$.

The functional I_n is *positively homogeneous*: for $\gamma \in \mathbb{R}$ and $\alpha \geq 0$, $I_n(\alpha\gamma) = \alpha\gamma = \alpha I_n(\gamma)$; for $a = \sum_m \lambda_m b_m$ (λ_m, b_m as above) and $\alpha > 0$, $\alpha a = \sum_m \lambda_m \alpha b_m$ and $\alpha b_m \in C^1_n$ by Remark 2 Part 1, so $I_n(\alpha a) = \sum_m \lambda_m I(\alpha b_m) = \alpha \sum_m \lambda_m I(b_m) = \alpha I_n(a)$, because I is positively homogeneous. Finally, for a as above and $\alpha = 0$, $\alpha a = 0$, so $I_n(\alpha a) = I_n(0) = 0 = \alpha I_n(a)$.

The functional I_n is also *additive*. By definition, for $\gamma, \delta \in \mathbb{R}$, $I_n(\gamma + \beta) = \gamma + \beta = I_n(\gamma) + I_n(\beta)$. For $a = \sum_m \lambda_m b_m$ (λ_m, b_m as above) and $\gamma \in \mathbb{R}$, $a + \gamma = \sum_m \lambda_m(b_m + \gamma)$ and $b_m + \gamma \in C^1_n$ by Remark 2 Part 1, so $I_n(a + \gamma) = \sum_m \lambda_m I(b_m + \gamma) = \sum_m \lambda_m I(b_m) + \gamma = I_n(a) + \gamma$, because I is c-linear. Finally, if $a = \sum_\ell \alpha_\ell a_\ell$ and $b = \sum_m \beta_m b_m$ for suitable $\alpha_\ell, a_\ell, \beta_m, b_m$, then $\frac{1}{2}a + \frac{1}{2}b = \sum_\ell \frac{1}{2}\alpha_\ell a_\ell + \sum_m \frac{1}{2}\beta_m b_m$ (where some a_ℓ may be equal to some b_m) and therefore

$I_n(\frac{1}{2}a+\frac{1}{2}b) = \sum_\ell \frac{1}{2}\alpha_\ell I(a_\ell) + \sum_m \frac{1}{2}\beta_m I(b_m) = \frac{1}{2}\sum_\ell \alpha_\ell I(a_\ell) + \frac{1}{2}\sum_m I(b_m) = \frac{1}{2}I_n(a) + \frac{1}{2}I_n(b)$. Since I_n is positively homogeneous, $I_n(a+b) = I_n(a) + I_n(b)$.

Finally, the functional I_n is *monotonic*. Remark 3 implies that $I_n(a) \geq I_n(b)$ for $a, b \in \operatorname{conv} C_n^1$, and by definition this is also true if both a and b are constant acts. So, suppose $a = \sum_m \lambda_m b_m$, for λ_m, b_m as in Eq. 11, and $a \geq \gamma$ for some $\gamma \in \mathbb{R}$. Then $a \geq \frac{1}{2}a + \frac{1}{2}\gamma$; since $\frac{1}{2}a + \frac{1}{2}\gamma = \sum_m \lambda_m(\frac{1}{2}b_m + \frac{1}{2}\gamma) \in \operatorname{conv} C_n^1$, invoking monotonicity on $\operatorname{conv} C_n^1$, and additivity and positive homogeneity, one obtains $I_n(a) \geq I_n(\frac{1}{2}a + \frac{1}{2}\gamma) = I_n(\frac{1}{2}a) + I_n(\frac{1}{2}\gamma) = \frac{1}{2}I_n(a) + \frac{1}{2}\gamma$; this implies $I_n(a) \geq \gamma$, as required. The case $a \leq \gamma$ is analogous.

To summarize, I_n is well-defined, positively homogeneous, additive and monotonic on the convex wedge C_n^2. Therefore, it has a unique extension to a positive (hence monotonic and sup-norm continuous) linear functional J_n on the linear subspace $C_n^2 - C_n^2$, given by $J_n(a-b) = J_n(a) - J_n(b)$ for $a, b \in C_n^2$. It follows that J_n can be extended to a (not necessarily unique) *positive* linear functional P_n on $B(S,\Sigma)$ (cf. e.g. Holmes [17], §6.B; observe that the constant function 1 belongs to the subspace $C_n^2 - C_n^2$ and is a core point of the cone of non-negative functions). Furthermore, since $P_n(1) = J_n(1) = I_n(1) = 1$, $\|P_n\| = 1$; that is, P_n can be represented by a probability charge on (S,Σ) (cf. e.g. [1], Theorem 11.32), henceforth also denoted P_n. Clearly, for all $a \in C_n^1$, $P_n(a) = J_n(a) = I_n(a) = I(a)$. ■

Observation. The last paragraph provides the key step in the proof of Proposition 2.2. If $f \simeq g$ for all $f, g \in C \subset L$, then $a \simeq b$ for all $a, b \in D \equiv \{\alpha u \circ f : \alpha \geq 0, f \in C\}$. Thus I is positively homogeneous, additive and monotonic on the convex wedge D; as above, the restriction of I to D has a unique positive linear extension J to the linear space $D - D$, which in turn has a positive linear extension to all of $B(S,\Sigma)$. Thus, there exists a probability charge P such that $I(a) = \int a\, dP$ for all $a \in D$. The converse is obvious.

5.1.6 Sufficiency: Uniqueness of the charges P_n

Henceforth, $\operatorname{int} C$ denotes the interior of the generic set $C \subset B(S,\Sigma)$. For $n = 1, \ldots, N_1$, let

$$C_n^2 = \operatorname{cl} \operatorname{int} \operatorname{cl} C_n^1; \tag{12}$$

note that,C_n^2 has non-empty interior if $\operatorname{cl} C_n^1$ does, and is empty otherwise; moreover, since $\operatorname{int} \operatorname{cl} C_n^1 \subset C_n^2$, $\operatorname{int} \operatorname{cl} C_n^1 \subset \operatorname{int} C_n^2$, so C_n^2 is the closure of its interior.

Lemma 5.12 *Assume that $\succeq$ satisfies Axioms 1–5 and 6. Then at least one of the sets $\operatorname{cl} C_1^1, \ldots, \operatorname{cl} C_{N_1}^1$ has non-empty interior. Furthermore, assume w.l.o.g. that* $\operatorname{int} \operatorname{cl} C_n^1 \neq \emptyset$, *hence* $\operatorname{int} C_n^2 \neq \emptyset$ *and $C_n^2 = \operatorname{cl} \operatorname{int} C_n^2$, for $n = 1, \ldots, N_2 \leq N_1$; then:*

1. *$B(S,\Sigma) = \bigcup_{n=1}^{N_2} C_n^2$;*

2. *For every $n = 1, \dots, N_2$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$: $a \in C^2_n$ implies $\alpha a + \beta \in C^2_n$; in particular, if $\alpha > 0$ and $a \in \operatorname{int} C^2_n$, then $\alpha a + \beta \in \operatorname{int} C^2_n$.*
3. *For every $n = 1, \dots, N_2$, $P_n(a) = I(a)$ for all $a \in C^2_n$.*

Proof. Since $B(S, \Sigma) = \bigcup_{n=1}^{N_1} C^1_n$, a fortiori $B(S, \Sigma) = \bigcup_{n=1}^{N_1} \operatorname{cl} C^1_n$. That is, $B(S, \Sigma)$ is the union of finitely many closed sets; therefore,[30] $\bigcup_{n=1}^{N_1} \operatorname{int} \operatorname{cl} C^1_n$ is dense in $B(S, \Sigma)$. Hence, some $\operatorname{cl} C^1_n$'s have empty interior; assume w.l.o.g. that these are the first $N_2 \leq N_1$.

For Part 1, consider $a \in B(S, \Sigma)$ and $\{a_k\} \subset \bigcup_{n=1}^{N_2} \operatorname{int} \operatorname{cl} C^1_n$ such that $a_k \to a$. Since there are finitely many sets $\operatorname{cl} C^1_n$ with non-empty interior, there is $n \in \{1, \dots, N_2\}$ and a subsequence $\{a_{k(\ell)}\}$ such that $a_{k(\ell)} \in \operatorname{int} \operatorname{cl} C^1_n$ for all ℓ; thus, $a \in \operatorname{cl} \operatorname{int} \operatorname{cl} C^1_n = C^2_n$. Thus, $B(S, \Sigma) = \bigcup_{n=1}^{N_2} C^2_n$.

For Part 2, consider first the following subclaim: *if $C \subset B(S, \Sigma)$ is such that $a \in C$, $\alpha, \beta \in \mathbb{R}$, and $\alpha > 0$ imply $\alpha a + \beta \in C$, then in particular $a \in \operatorname{int} C$ implies $\alpha a + \beta \in \operatorname{int} C$.* To see this, fix α, β and C as stated, and choose $a \in \operatorname{int} C$. Then there exists $\epsilon > 0$ such that $\|b - a\| < \epsilon$ implies $b \in C$. Consider $c \in B(S, \Sigma)$ such that $\|c - [\alpha a + \beta]\| < \alpha\epsilon$: then $\left\|\frac{c-\beta}{\alpha} - a\right\| = \frac{1}{\alpha}\|c - \beta - \alpha a\| < \frac{1}{\alpha}\alpha\epsilon = \epsilon$, so $\frac{c-\beta}{\alpha} \in C$, and therefore $c \in C$.

Now consider $\alpha, \beta \in \mathbb{R}$ with $\alpha > 0$ and recall that, by Remark 2 Part 1, $a \in C^1_n$ implies $\alpha a + \beta \in C^1_n$; hence, the same is true for $a \in \operatorname{cl} C^1_n$. The above subclaim applied to $C = \operatorname{cl} C^1_n$ implies that $a \in \operatorname{int} \operatorname{cl} C^1_n$ implies $\alpha a + \beta \in \operatorname{int} \operatorname{cl} C^1_n$; consequently, $a \in C^2_n = \operatorname{cl} \operatorname{int} \operatorname{cl} C^1_n$ implies $\alpha a + \beta \in C^2_n$. The subclaim, applied to $C = C^2_n$, implies the last statement in Part 2. Finally, pick any $a \in C^2_n$; for any $\beta \in \mathbb{R}$, $\frac{1}{k}a + \beta \in C^2_n$ for all $k \geq 1$ and $\frac{1}{k}a + \beta \to \beta$. Since C^2_n is closed, it follows that $\beta \in C^2_n$. Hence $\alpha a + \beta \in C^2_n$ for $\alpha = 0$ as well.

For Part 3, recall that $P_n(a) = I(a)$ for all $a \in C^1_n$. Consider $a \in B(S, \Sigma)$ and $\{a_k\} \subset C^1_n$ such that $a_k \to a$: then $P_n(a) = \lim_k P_n(a_k) = \lim_k I(a_k) = I(a)$, because both P_n and I are sup-norm continuous functionals. Thus, P_n and I also agree on on $\operatorname{cl} C^1_n$, hence a fortiori on $C^2_n = \operatorname{cl}[\operatorname{int} \operatorname{cl} C^1_n] \subset \operatorname{cl}[\operatorname{cl} C^1_n] = \operatorname{cl} C^1_n$. ■

Lemma 5.13 *Assume that $\succeq$ satisfies Axioms 1–5 and 6. Then, for every $n = 1, \dots, N_2$, P^2_n is the only continuous linear functional that agrees with I on C^2_n. In particular, every charge P_n, $n = 1, \dots, N_2$, is a plausible prior.*

Proof. Fix n as above. Lemma 5.12 Part 3 states that P_n agrees with I on C^2_n; furthermore, the latter set has non-empty interior. Thus, fix $c \in \operatorname{int} C^2_n$; by the second claim in Part 2, it is w.l.o.g. to assume that $\|c\| \leq \frac{1}{2}$ (if this is not the case, replace c by $\frac{1}{2\|c\|}c \in \operatorname{int} C^2_n$).

[30] In any topological space T, the intersection of *finitely* many open dense sets is dense; this (cf. e.g. [1], Theorem 3.34) implies that, if $T = \bigcup_{i=1}^{I} F_i$ and each F_i is closed, then $\bigcup_{i=1}^{I} \operatorname{int} F_i$ is dense in T.

Also, there exists $\epsilon' > 0$ such that $\|a - c\| < \epsilon'$ implies $a \in C_n^2$. Finally, let $\epsilon = \min(\epsilon, \frac{1}{2})$: then $\|a - c\| < \epsilon$ implies $a \in C_n^2$ and $\|a\| < 1$, so that there exists $f \in L$ with $u \circ f = a$.

Now consider the set of acts $C = \{\alpha f + (1-\alpha)y \ : \ \|u \circ f - c\| < \epsilon,\ \alpha \in [0,1],\ y \in Y\}$. By construction, C contains all lotteries; moreover, it is closed under mixtures. Consequently, the set $u \circ C = \{u \circ f : f \in C\}$ is convex; furthermore, it contains the open ϵ–ball around c, and by Lemma 5.12 Part 2, $u \circ C \subset C_n^2$. Therefore P_n agrees with I on $u \circ C$, so $f \succeq g$ iff $P_n(u \circ f) \geq P_n(u \circ g)$ for all $f, g \in C$; in particular, $\succeq$ satisfies Mixture Neutrality on C.

To complete the proof, it will be shown that P_n is the unique linear functional that agrees with I on $u \circ C$ (hence on C_n^2); this implies that P_n is a plausible prior. Consider another linear functional Q such that $Q(a) = I(a)$ for all $a \in u \circ C$. Fix an arbitrary $a \in B(S, \Sigma)$; then there exists $\alpha \in (0,1)$ such that $\|\alpha a + (1-\alpha)c - c\| = \alpha\|a - c\| < \epsilon$, and therefore $\alpha a + (1-\alpha)c \in u \circ C$. This implies that $\alpha Q(a) + (1-\alpha)Q(c) = Q(\alpha a + (1-\alpha)c) = I(\alpha a + (1-\alpha)c) = P_n(\alpha a + (1-\alpha)c) = \alpha P_n(a) + (1-\alpha)P_n(c)$. Furthermore, $Q(c) = I(c) = P_n(c)$: therefore, $Q(a) = P_n(a)$. Since a was arbitrary, $Q = P_n$, so P_n is a plausible prior. ■

Lemma 5.14 *Assume that $\succeq$ satisfies Axioms 1–5 and 6. Let $D \subset B(S, \Sigma)$ be such that, for some continuous linear functional Q on $B(S, \Sigma)$, $I(a) = Q(a)$ for all $a \in D$. Then there exists $n \in \{1, \ldots, N_2\}$ such that $Q(a) = P_n(a)$ for all $a \in D$. Hence, the charges $P_1, \ldots, P_n$ are the only plausible priors for $\succeq$.*

Proof. Since I and Q are both norm-continuous, $I(a) = Q(a)$ for all $a \in \mathrm{cl}\, D$, so it is w.l.o.g. to assume that D is itself closed. Note that $D = \bigcup_{n=1}^{N_2}(D \cap C_n^2)$, and every set $D \cap C_n^2$ is closed in the relative topology on D inherited from $B(S, \Sigma)$. Furthermore, D is a complete subspace of $B(S, \Sigma)$. Therefore, arguing as in the proof of Lemma 5.12, not all sets $D \cap C_n^2$ have empty relative interior.

Thus, consider n such that $D \cap C_n^2$ has non-empty relative interior; that is, there exists $c \in D \cap C_n^2$ and $\epsilon > 0$ such that $\|a - c\| < \epsilon$ and $a \in D$ imply $a \in D \cap C_n^2$.

Consider an arbitrary $a \in D$; then there exists $\alpha \in (0,1)$ such that $\|\alpha a + (1-\alpha)c - c\| = \alpha\|a - c\| < \epsilon$, and hence $\alpha a + (1-\alpha)c \in D \cap C_n^2$: this implies that, for this α, $Q(\alpha a + (1-\alpha)c) = I(\alpha a + (1-\alpha)c) = P_n(\alpha a + (1-\alpha)c)$. Since in particular $Q(c) = P_n(c)$, it follows by linearity of Q and P_n that $Q(a) = P_n(a)$. ■

5.1.7 Sufficiency: Construction of the proper covering

In general, a set C_n^2 may fail to be minimally convex; however, consider the following construction. First, assume w.l.o.g. that, for some $N \leq N_1$, (i) $P_n \neq P_m$ for all $n, m \in \{1, \ldots, N\}$ with $n \neq m$, and that (ii) for every $m \in \{N+1, \ldots, N_2\}$ (if any), there exists $n \in \{1, \ldots, N\}$

such that $P_m = P_n$. Then, for $n = 1, \ldots, N$, define

$$C_n = \bigcup \{C_m^2 : P_m = P_n\}. \tag{13}$$

Thus, C_n is the union of C_n^2 and any other set C_m^2 for which $P_m = P_n$.

Lemma 5.15 *Assume that $\succeq$ satisfies Axioms 1–5 and 6. Then, for all $a, b \in B(\Sigma)$ such that $a \approx b$, and for every $\gamma \in [0,1]$, there is $\epsilon > 0$ such that $\|c - a\| < \epsilon$ and $c \in C_n$ for some $n = 1, \ldots, N$ implies $\lambda c + (1-\lambda)[\gamma a + (1-\gamma)b] \in C_n$ for all $\lambda \in (0,1)$.*

Since each set C_n is closed, the conclusion of the Lemma also holds for $\lambda = 0$ (it is true by assumption for $\lambda = 1$). By symmetry of $\approx$, an analogous statement is true for points $c \in B(S, \Sigma)$ such that $\|c - b\| < \epsilon$. This result also implies that $a \approx b$ and $a \in C_n$ implies that $\lambda a + (1-\lambda)b \in C_n$ for all $\lambda \in [0,1]$.

Proof. Consider $a \approx b$ and suppose that the assertion fails for some $\gamma \in [0,1]$. Let $b' = \gamma a + (1-\gamma)b$ for notational simplicity. Then, in particular, for all integer $k \geq 1$, there is $c_k \in B(S, \Sigma)$ such that $\|c_k - a\| < \frac{1}{k}$, $c_k \in C_{n(k)}$, and $\lambda_k c_k + (1-\lambda_k)b' \notin C_{n(k)}$ for some $\lambda_k \in (0,1)$. Now $c_k \in C_\ell^2$ for some $\ell \in \{1, \ldots, N_2\}$ such that $P_\ell = P_{n(k)}$, so c_k is the limit of points in $\operatorname{int} \operatorname{cl} C_\ell^1 \subset \operatorname{int} C_\ell^2 \subset \operatorname{int} C_{n(k)}$. Note that, if $c \in B(S, \Sigma)$ satisfies $\|c - c_k\| < \frac{1}{k} - \|c_k - a\|$, then $\|c - a\| < \frac{1}{k}$; furthermore, since the complement of $C_{n(k)}$ is open and $\lambda_k c + (1-\lambda_k)b' \to \lambda_k c_k + (1-\lambda_k)b' \notin C_{n(k)}$ as $c \to c_k$, it is possible to find $c'_k \in \operatorname{int} C_{n(k)}$ such that $\|c'_k - a\| < \frac{1}{k}$ and $\lambda_k c'_k + (1-\lambda_k)b' \in C_{m(k)}$, with $m(k) \neq n(k)$.

Now $\lambda_k c'_k + (1-\lambda_k)b'$ is itself the limit of points in $\operatorname{int} \operatorname{cl} C_\ell^1 \subset \operatorname{int} C_\ell^2 \subset \operatorname{int} C_{m(k)}$ for some (different) $\ell \in \{1, \ldots, N_2\}$ with $P_\ell = P_{m(k)}$. Also, for any $c \in B(S, \Sigma)$, the act $c'' = \frac{c - (1-\lambda_k)b'}{\lambda_k}$ satisfies $\lambda_k c'' + (1-\lambda_k)b' = c$, and as c approaches $\lambda_k c'_k + (1-\lambda_k)b'$, c'' approaches c'_k. Therefore, there exists $c \in \operatorname{int} C_{m(k)}$ close enough to $\lambda_k c'_k + (1-\lambda_k)b'$ so that the act c''_k defined by $c = \lambda_k c''_k + (1-\lambda_k)b'$ satisfies $c''_k \in \operatorname{int} C_{n(k)}$ and $\|c''_k - a\| < \frac{1}{k}$.

To summarize, there exists $c''_k \in \operatorname{int} C_{n(k)}$ such that $\|c''_k - a\| < \frac{1}{k}$ and $\lambda_k c''_k + (1-\lambda_k)b' \in \operatorname{int} C_{m(k)}$, where $m(k) \neq n(k)$. Next, observe that there exist $\epsilon_{m(k)} > 0$ such that $\|c - [\lambda c''_k + (1-\lambda_k)b']\| < \epsilon_{m(k)}$ implies $c \in \operatorname{int} C_{m(k)}$, and $\epsilon_{n(k)} > 0$ such that $\|c - c''_k\| < \epsilon_{n(k)}$ implies $c \in \operatorname{int} C_{n(k)}$. Thus, let $\epsilon_k = \min(\frac{1}{k} - \|c''_k - a\|, \epsilon_{n(k)}, \epsilon_{m(k)})$; then $\|c - c''_k\| < \epsilon_k$ implies both $c \in \operatorname{int} C_{n(k)}$ and $\|c - a\| < \frac{1}{k}$, and furthermore

$$\|\lambda_k c + (1-\lambda_k)b' - \lambda_k c''_k - (1-\lambda_k)b'\| = \lambda_k \|c - c''_k\| < \lambda_k \epsilon_{m(k)} < \epsilon_{m(k)},$$

hence $\lambda_k c + (1-\lambda_k)b' \in \operatorname{int} C_{m(k)}$.

Now consider the following preliminary subclaim: *Suppose that $a \simeq b$ and, for distinct $\lambda, \lambda' \in [0,1]$, $\lambda a + (1-\lambda)b, \lambda' a + (1-\lambda')b \in C_n$. Then $P_n(a) = I(a)$ and $P_n(b) = I(b)$.*

To prove the subclaim, note that $\lambda P_n(a) + (1-\lambda)P_n(b) = P_n(\lambda a + (1-\lambda)b) = I(\lambda a + (1-\lambda)b) = \lambda I(a) + (1-\lambda)I(b)$ and similarly $\lambda' P_n(a) + (1-\lambda')P_n(b) = \lambda' I(a) + (1-\lambda')I(b)$. Subtracting the second equation from the first yields $(\lambda - \lambda')[P_n(a) - P_n(b)] = (\lambda-\lambda')[I(a) - I(b)]$, hence $P_n(a) = P_n(b) + I(a) - I(b)$; substituting in the first equation now yields $\lambda P_n(b) + \lambda I(a) - \lambda I(b) + (1-\lambda)P_n(b) = \lambda I(a) + (1-\lambda)I(b)$, i.e. $P_n(b) = I(b)$, and therefore also $P_n(a) = P_n(b)$. [Note that the proof is trivial if one of λ, λ' is zero or one].

Continuing with the proof of the main claim, observe that, since $a \approx b$, there exists K such that, for $k \geq K$, it is the case that $\|c - a\| < \frac{1}{k}$ implies $c \simeq b'$. Thus, choose $k \geq K$ and consider any $c \in B(S,\Sigma)$ such that $\|c - c''_k\| < \epsilon_k$; then, by the choice of k and ϵ_k, $\|c-a\| < \frac{1}{k}$, so $c \simeq b'$, and $c \in \operatorname{int} C_{n(k)}$, so there exists $\lambda < 1$ such that $\lambda c + (1-\lambda)b' \in C_{n(k)}$. Hence, invoking the subclaim with the values 1 and λ, $P_{n(k)}(c) = I(c)$ and $P_{n(k)}(b') = I(b')$. Similarly, for such c, there exists $\lambda \neq \lambda_k$ such that $\lambda c + (1-\lambda)b' \in C_{m(k)}$, and the subclaim implies that $P_{m(k)}(c) = I(c)$ and $P_{m(k)}(b') = I(b')$.

Finally, consider an arbitrary $c \in B(S,\Sigma)$ and the mixture $\alpha c + (1-\alpha)c''_k$. Since $\|\alpha c + (1-\alpha)c''_k - c''_k\| = \alpha\|c - c''_k\|$, by choosing $\alpha > 0$ small one can ensure that this quantity is smaller than ϵ_k; thus, $P_{n(k)}(\alpha c + (1-\alpha)c''_k) = I(\alpha c + (1-\alpha)c''_k) = P_{m(k)}(\alpha c + (1-\alpha)c''_k)$; since clearly also $P_{n(k)}(c''_k) = P_{m(k)}(c''_k)$, it follows that $P_{n(k)}(c) = P_{m(k)}(c)$. Therefore, $P_{n(k)} = P_{m(k)}$, which contradicts the construction of the sets $C_{n(k)}$ and $C_{m(k)}$. ■

Lemma 5.16 *Assume that $\succeq$ satisfies Axioms 1–5 and 6. Then the sets $C_1, \dots, C_N$ are minimally convex, and constitute a proper covering.*

Proof. It is clear that every set C_n is closed and has non-empty interior; that $B(S,\Sigma) = \bigcup_{n=1}^N C_n$, that $a \in C_n$ implies $\alpha a + \beta \in C_n$ for $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$, and that $P_n(a) = I(a)$ for all $a \in C_n$, because the corresponding C^2_n's satisfy these properties. Suppose $a \in C_n$, so $a \in C^2_\ell$ for some $\ell \in \{1, \dots, N_2\}$ with $P_\ell = P_n$; then $a \in \operatorname{cl}\operatorname{int} C^2_\ell \subset \operatorname{cl}\operatorname{int} C_n$, i.e. $C_n = \operatorname{cl}\operatorname{int} C_n$. Furthermore, suppose $C_n \cap C_m$ has non-empty interior, and consider $a \in \operatorname{int} C_n \cap C_m$. For any $b \in B(S,\Sigma)$, there is $\alpha \in (0,1)$ such that $\alpha b + (1-\alpha a) \in \operatorname{int} C_n \cap C_m$; therefore, $P_n(\alpha b + (1-\alpha)a) = I(\alpha b + (1-\alpha)a) = P_m(\alpha b + (1-\alpha)a)$; since $P_n(a) = P_m(a)$ as well, $P_n(b) = P_m(b)$. Thus, $P_n = P_m$, so $n = m$; hence, if $n \neq m$, $C_n \cap C_m$ has empty interior.

It remains to be shown that each set C_n is minimally convex. Consider an infinite subset $C' \subset C_n$. This subset contains a countably infinite subset $\{a_k\}$. Consider the sequence $\{b_k\}$ defined by $b_k = 0$ if $a_k = 0$ and $b_k = \frac{1}{k\|a_k\|}a_k$ otherwise. Then $b_k \to 0$, so Axiom 6 implies that there is a subsequence $\{b_{k(\ell)}\}$ for which $b_{k(\ell)} \approx b_{k(\ell')}$ for all ℓ, ℓ'. Since, for every k, b_k is either equal to a_k or to a positive multiple of a_k, Lemma 5.2 Part 8 implies that $a_{k(\ell)} \approx a_{k(\ell')}$ for every ℓ. Lemma 5.15 shows that, for this subsequence, and for every $\gamma \in [0,1]$, there is $\epsilon > 0$ such that $\|c - a_{k(\ell)}\| < \epsilon$ and $c \in C_m$ implies $\lambda c + (1-\lambda)[\gamma a_{k(\ell)} + (1-\gamma)a_{k(\ell')}]$ for all

$\lambda \in (0,1)$; in particular, this is the case for $m = n$. ■

To summarize: By Lemma 5.16, $C_1, \ldots, C_N$ is a proper covering; the associated probabilities are all distinct, by construction, so property (i) in Statement 2 of Theorem 2.6 holds; moreover, $I(a) = P_n(a)$ for all $a \in C_n$, which (jointly with Lemma 5.1) implies that property (ii) also holds. Uniqueness of u guaranteed by Lemma 5.1; uniqueness of each P_n is established in Lemma 5.13. Finally, consider another proper covering $D_1, \ldots, D_{N'}$, with associated priors $Q_1, \ldots, Q_{N'}$, that satisfy properties (i) and (ii) in Statement 2 of Theorem 2.6: then Lemma 5.14 implies that every $Q_{n'}$ corresponds to some P_n, and vice versa, so $N = N'$ and it is wlog to assume that each D_n is associated with P_n. Moreover, if $\operatorname{int} D_n \cap C_m \neq \emptyset$ for $n \neq m$, the argument in the proof of Lemma 5.16 implies that $P_n = P_m$, a contradiction; thus, $\operatorname{int} D_n \subset C_n$, and similarly $\operatorname{int} C_n \subset D_n$. Thus, $\operatorname{int} C_n = \operatorname{int} D_n$. Since $C_n = \operatorname{cl} \operatorname{int} C_n$ and similarly for D_n, $C_n = D_n$. This **completes the proof of Theorem 2.6.**

Turn now to the Corollaries in the text. Consider first Corollary 2.7: Lemma 5.13 shows that every P_n is a plausible prior, and Lemma 5.14 ensures that there are no other plausible priors. Also, Lemma 5.15 implies that $f \approx g$ only if $u \circ f, u \circ g \in C_n$ for some n.

To prove Corollary 2.8, it is sufficient to construct the proper covering corresponding to the charges $\{\alpha Q_n + (1-\alpha) Q_m : (n,m) \in \mathcal{M}\}$. Consider the sets of the form $C(n,m) = \{a \in B(S,\Sigma) : Q_n \in \arg\min_k \int a\, dQ_k,\ Q_m \in \arg\max_k \int a\, dQ_k\}$. Define a relation R on the collection of such sets by stipulating that $C(n,m) \, R \, C(n',m')$ iff $\alpha Q_n + (1-\alpha) Q_m = \alpha Q_{n'} + (1-\alpha) Q_{m'}$. Then every element of the proper covering for α-MEU preferences is the (finite) union of sets $C(n,m)$ in the same equivalence class for R. In particular, since every set $C(n,m)$ is closed and convex, Remark 1 ensures that Property 4 in Def. 2.5 holds. Necessity is proved in the Online Appendix, §6.6.3 (which also discusses the case $\alpha = \frac{1}{2}$).

The construction of the proper covering in Corollary 2.9 is analogous to that of Corollary 2.8: each element is a union of maximal comonotonic cones associated with the same probability distribution. As above, since each such cone is convex, Remark 1 applies.

Finally, the fact that $\int u \circ f dP_n \geq \int u \circ g dP_n$ for all n implies $f \succeq g$ is established in the proof of necessity (see the argument for monotonicity of I). A related fact is used below.

Lemma 5.17 *Under the equivalent conditions of Theorem 2.6, for all $a, b \in B(S,\Sigma)$: if $\int a\, dP_n = \int b\, dP_n$ for all $n \in \{1, \ldots, N\}$ and $b \in C_m$ for some $m \in \{1, \ldots, N\}$, then $a \in C_m$. Furthermore, if $b \in \operatorname{int} C_m$, then $a \in \operatorname{int} C_m$.*

Proof. If $P_n(a) = P_n(b)$ for all n, then for all $\lambda \in [0,1]$, $P_n(a) = P_n(\lambda a + (1-\lambda) b) = P_n(b)$ for all n; hence $I(a) = I(\lambda a + (1-\lambda) b) = I(b)$.

Furthermore, for any $\gamma \in [0,1]$, $\lambda \in [0,1]$, and $c \in B(S,\Sigma)$, and for any $n \in \{1,\dots,N\}$,

$$P_n(\gamma a + (1-\gamma)c) = \gamma P_n(a) + (1-\gamma)P_n(c) = \gamma P_n(\lambda a + (1-\lambda)b) + (1-\gamma)P_n(c) = \\ = P_n(\gamma[\lambda a + (1-\lambda)b] + (1-\gamma)c),$$

and similarly $P_n(\gamma b + (1-\gamma)c) = P_n(\gamma[\lambda a + (1-\lambda)b] + (1-\gamma)c)$, which, as above, implies that $I(\gamma a + (1-\gamma)c) = I(\gamma[\lambda a + (1-\lambda)b] + (1-\gamma)c) = I(\gamma b + (1-\gamma)c)$. Therefore, $a \simeq c$ or $b \simeq c$ imply $\lambda a + (1-\lambda)b \simeq c$ for all $\lambda \in [0,1]$.

Now suppose $c_k \to a$; then, by Lemma 5.8, $c_k \simeq a$ for large k; for such k, the argument just given implies that also $c_k \simeq \lambda a + (1-\lambda)b$ for all $\lambda \in [0,1]$. The same argument applies if $c_k \to b$, so $a \approx b$. Since $b \in C_m$, as noted above, Lemma 5.15 implies that $a \in C_m$ as well.

Finally, if $b \in \operatorname{int} C_m$ but $a \notin \operatorname{int} C_m$, there is $n \neq m$ such that $a \in C_n$.[31] Since $a \approx b$, Lemma 5.15 implies $b \in C_n$; but then $C_n \cap C_m$ has non-empty interior,[32] a contradiction. ■

5.2 Proof of Theorem 3.2

5.2.1 Notation and Preliminary results

Let $u, C_1, \dots, C_N$ and $P_1, \dots, P_N$ represent $\succeq$; as in §5.1.2, let $I(a) = \int a\,dP_n$ for all $a \in C_n$. Recall that I is monotonic, normalized, and c-linear. Finally, assume that $u(Y) \supset [-1,1]$, and define $aEb = 1_E a + 1_{E^c} b$ for $a, b \in B(S,\Sigma)$. Note that $E \in \Sigma$ is non-null iff, for all $a, b \in B(S,\Sigma)$, $a(s) = b(s)$ for $s \in S \setminus E$ and $a(s) > b(s)$ for all $s \in E$ imply $I(a) > I(b)$.

Lemma 5.18 *An event $E \in \Sigma$ is non-null for $\succeq$ if and only if, for all $n \geq 1$, $P_n(E) > 0$.*

Proof. Clearly, by c-linearity of I, E is non-null iff, for all $a \in B(S,\Sigma)$, $x, x' \in \mathbb{R}$ with $x > x'$, and $\lambda > 0$, $I(a + \lambda[x\,E\,x']) > I(a + \lambda x')$, i.e. iff $I(a + \lambda 1_E(x - x')) > I(a)$, i.e. iff $I(a + \lambda 1_E) > I(a)$ for all $\lambda > 0$.

Suppose E is non-null; pick $n \in \{1,\dots,N\}$ and $a \in \operatorname{int} C_n$. Then there is $\epsilon > 0$ such that $a + \epsilon 1_E \in C_n$. Thus, $P_n(a + \epsilon 1_E) > P_n(a)$, so $P_n(E) > 0$. Conversely, assume $P_n(E) > 0$ for all $n \in \{1,\dots,N\}$, and consider $a \in B(S,\Sigma)$ and $\epsilon > 0$. Let $p = \min_n P_n(E)$: then, for each n, $P_n(a + \epsilon 1_E) \geq P_n(a) + \epsilon p = P_n(a + \epsilon p)$, so $I(a + \epsilon 1_E) \geq I(a + \epsilon p) = I(a) + \epsilon p > I(a)$. ■

[31] Every neighborhood of a contains a point not in C_m; form a sequence, and note that there is n such that a subsequence lies entirely in C_n. Hence, so does its limit a.

[32] Suppose that $b \in C^2_\ell = \operatorname{cl}\operatorname{int}\operatorname{cl} C^1_\ell$ for some ℓ such that $P_\ell = P_n$, so that $C^2_\ell \subset C_n$. Then there is a subsequence $\{b_k\} \subset \operatorname{int}\operatorname{cl} C^1_\ell \subset \operatorname{int} C^2_\ell \subset \operatorname{int} C_n$ such that $b_k \to b$. Since $b \in \operatorname{int} C_m$, there is K such that $b_k \in \operatorname{int} C_m$ for all $k \geq K$; thus, there is ϵ_m such that $\|c - b_K\| < \epsilon_m$ implies $c \in C_m$. Furthermore, since $b_K \in \operatorname{int} C_n$, there is $\epsilon_n > 0$ such that $\|c - b_K\| < \epsilon_n$ implies $c \in C_n$. Hence, the open ball $\{c : \|c - b_K\| < \min(\epsilon_n, \epsilon_m)\}$ is a subset of $C_n \cap C_m$.

Lemma 5.19 *Assume that $\succeq$ satisfies Axioms 1–5, and suppose that $E \in \Sigma$ is non-null. Then, for every $a \in B(S, \Sigma)$, there exists a unique solution $x \in \mathbb{R}$ to the equation*

$$x = I(aEx). \tag{14}$$

The map $J : B(S, \Sigma) \to \mathbb{R}$ associating to each $a \in B(S, \Sigma)$ the unique solution to Eq. (14) is monotonic, c-linear and normalized.

Now define the relations $\simeq_E$ and $\approx_E$ on $B(S, \Sigma)$ by $a \simeq_E b$ iff $J(\gamma a + (1-\gamma)b) = \gamma J(a) + (1-\gamma)J(b)$ and $a \approx b$ iff, for any sequence $\{c_k\}$ such that $c_k \to a$ or $c_k \to b$, and for any $\gamma \in [0, 1]$, there is K such that $k \geq K$ implies $c_k \simeq_E \gamma a + (1-\gamma)b$.

Corollary 5.20 *If $\succeq$ additionally satisfies Axiom 6, then for all $\{a_k\} \subset B(S, \Sigma)$ and $a \in B(S, \Sigma)$ such that $a_k \to a$, there exist $\{k(\ell)\}$ such that $a_{k(\ell)} \approx_E a_{k(\ell')}$ for all ℓ, ℓ'.*

Proof. (Lemma 5.19): Let $x_1 = \sup_{s \in E} a(s)$, $x_0 = \inf_{s \in E} a(s)$; by monotonicity, $I(aEx_1) - x_1 \leq 0$ and $I(aEx_0) - x_0 \geq 0$. By norm-continuity, there exists $x \in [x_0, x_1]$ such that $x = I(aEx)$. Furthermore, suppose there are two such solutions x, x', with $x > x'$. Then $I(aEx) - x = I(aEx') - x'$, i.e. $I(1_E(a - x)) = I(1_E(a - x')) = 0$. But this contradicts the fact that E is non-null, because $1_E(s)[a(s) - x] = 1_E(s)[a(s) - x'] = 0$ for $s \in S \setminus E$ and $1_E(s)[a(s) - x] = a(s) - x < a(s) - x' = 1_E(s)[a(s) - x']$ for $s \in E$.

The other properties are easy to prove, so the arguments are omitted.

(Corollary:) Now suppose that $a_k \to a$; then $J(a_k) \to J(a)$, which implies that $a_k\, E\, J(a_k) \to a\, E\, J(a)$ as well. If $\succeq$ satisfies Axiom 6, then there exists a subset of indices $\{k(\ell)\}$ such that $a_{k(\ell)}\, E\, J(a_{k(\ell)}) \approx a_{k(\ell')}\, E\, J(a_{k(\ell')})$ for all ℓ, ℓ'. Thus, to complete the proof of the last claim, it is sufficient to show that, for any $a, b \in B(S, \Sigma)$, if $a\, E\, J(a) \approx b\, E\, J(b)$, then $a \approx_E b$.

First, it will be shown that, for all $a, b \in B(S, \Sigma$, $a\, E\, J(a) \simeq b\, E\, J(b)$ implies $a \simeq_E b$. To see this, note that, for all $\gamma \in [0, 1]$,

$$\begin{aligned} &I([\gamma a + (1-\gamma)b]\, E\, [\gamma J(a) + (1-\gamma)J(b)]) = I(\gamma[a\, E\, J(a)] + (1-\gamma)[b\, E\, J(b)]) = \\ &= \gamma I(a\, E\, J(a)) + (1-\gamma)I(b\, E\, J(b)) = \gamma J(a) + (1-\gamma)J(b); \end{aligned}$$

since $x = J(\gamma a + (1-\gamma)b)$ is the only solution to the fixpoint equation $I([\gamma a + (1-\gamma)b]\, E\, x) = x$, this implies that $J(\gamma a + (1-\gamma)b) = \gamma J(a) + (1-\gamma)J(b)$, i.e. $a \simeq_E b$.

Now assume $a\, E\, J(a) \approx b\, E\, J(b)$. Recall that this implies $a\, E\, J(a) \simeq b\, E\, J(b)$, and hence $J(\gamma a + (1-\gamma)b) = \gamma J(a) + (1-\gamma)J(b)$ for all $\gamma \in [0, 1]$, as was just shown. Consider $c_k \to a$; as above, this implies $J(c_k) \to J(a)$, and hence $c_k\, E\, J(c_k) \to a\, E\, J(a)$. Then, for every

$\gamma \in [0,1]$, there exists K such that $k \geq K$ implies

$$c_k\, E\, J(c_k) \simeq \gamma[a\, E\, J(a)] + (1-\gamma)[b\, E\, J(b)] = [\gamma a + (1-\gamma)b]\, E\, [\gamma J(a) + (1-\gamma)J(b)] = \\ = [\gamma a + (1-\gamma)b]\, E\, J(\gamma a + (1-\gamma)b)),$$

and therefore $c_k \simeq_E \gamma a + (1-\gamma)b$. Thus, $a \approx_E b$, as claimed. ■

5.2.2 Necessity of the Axioms

Now turn to the proof of Theorem 3.2. To show that Statement 2 implies Statement 1, consider a non-null $E \in \Sigma$ and assume that $\succeq_E$ is represented by u and the Bayesian updates $P_{n_k}(\cdot|E)$, for $k = 1, \ldots, K$ and each $n_k \in \{1, \ldots, N\}$ such that Eq. (8) holds; conditional probabilities are well-defined by Lemma 5.18. Since $P_{n_k}(S \setminus E|E) = 0$ for all $k \in \{1, \ldots, K\}$, $\succeq_E$ satisfies Axiom 7. It remains to be shown that $\succeq, \succeq_E$ jointly satisfy Axiom 8.

Fix an act $f \in L$ such that $u \circ f \in C^E_k$; then a lottery $y \in Y$ satisfies $f \sim_E y$, i.e. $u(y) = \int u \circ f\, dP_{n_k}(\cdot|E)$, if and only if $fEy \sim y$. "Only if": assume $f \sim_E y$ and $u \circ [fEy] \in C_m$ for some $m \in \{1, \ldots, N\}$; then, by Eq. (8), $\int u \circ [fEy]\, dP_m = \int u \circ f\, E\, u(y)\, dP_m = u(y)$, i.e. $fEy \sim y$. "If": suppose $fEy \sim y$ and $u \circ [fEy] \in C_m$, so $u(y)$ solves the equation $I([u \circ f]Ex) = x$; if $f \not\sim_E y$, then $f \sim_E y'$ for some $y' \not\sim_E y$. By the "only if" part, assuming $u \circ [fEy'] \in C_{m'}$, $\int u \circ [fEy'] dP_{m'} = u(y')$, i.e. $I([u \circ f]Eu(y')) = u(y')$; since $u = u^E$, $u(y') \neq u(y)$, so there are two distinct solutions to $I(u \circ fEx) = x$, which contradicts Lemma 5.19. Thus, $fEy \sim y$ implies $f \sim_E y$. It follows that $f \succeq_E g$ iff $y \succeq y'$, where $fEy \sim y$ and $gEy' \sim y'$.

Dynamic c-Consistency can now be verified. Suppose $f \succeq_E y'$ and $f(s) \succeq y'$ for $s \in E^c$; by Monotonicity of $\succeq$, $f \succeq fEy'$. Also, if $y \sim fEy$, then $y \succeq y'$; thus, by monotonicity again, since $I(1_E[u \circ f - u(y)]) = 0$, $I(1_E[u \circ f - u(y')]) \geq 0$, or equivalently $I(u \circ fEu(y')) \geq u(y')$, i.e. $fEy' \succeq y'$. Thus, $f \succeq y'$, as needed. If instead $f \succ_E y'$, then $y \succ y'$; as above, $I(1_E[u \circ f - u(y')]) \geq 0$, but since, by Lemma 5.19, the solution to Eq. (14) is unique, it must be the case that actually $I(1_E[u \circ f - u(y')]) > 0$, or $fEy' \succ y'$. Thus, $f \succ y'$, as needed. The cases $f \preceq_E y'$ and $f \prec_E y'$ are treated similarly.

5.2.3 Sufficiency of the Axioms

Claim 1: For all acts f and outcomes y, $f \succeq_E y \Leftrightarrow fEy \succeq y$ and $f \preceq_E y \Leftrightarrow fEy \preceq y$.

Proof: suppose $f \succeq_E y$. By Axiom 7, $fEy \sim_E f \succeq_E y$. Clearly, $fEy(s) \sim y$ for all $s \in E^c$. Thus, by Axiom 8, $fEy \succeq y$. If instead $f \prec_E y$, the same argument shows that $fEy \prec y$, which proves the first part of the claim. The second is proved similarly.

Claim 2: For all outcomes y, y', $y \succeq_E y' \Leftrightarrow y \succeq y$.

Proof: The preceding claim implies that $y \succeq_E y'$ iff $yEy' \succeq y'$; that is, for some $n \geq 1$, $u(y)P_n(E)+u(y')P_n(E^c) \geq u(y')$. Since E is non-null, $P_n(E) > 0$, so the preceding expression reduces to $u(y) \geq u(y')$. This implies the claim.

Now, by Claim 2, u represents $\succeq_E$ on Y. Also, by Claims 1 and 2, $f \succeq_E g$ iff $y \succeq y'$ for all y, y' such that $fEy \sim y$ and $gEy' \sim y'$. To see this, note that, by Claim 1, $f \sim_E y$ and $g \sim_E y'$; hence, $f \succeq_E g$ iff $y \succeq_E y'$; by Claim 2, this is equivalent to $y \succeq y'$, as required.

Thus, the unique, monotonic, c-linear, and normalized fixpoint map J defined in Lemma 5.19 represents $\succeq^E$: for all $f, g \in L$, $f \succeq_E g$ iff $J(u \circ f) \geq J(u \circ g)$. Furthermore, Corollary 5.20 implies that $\succeq$ also satisfies Axiom 6; therefore, there exists a proper covering $C_1^E, \ldots, C_K^E$ of $B(S, \Sigma)$, and probability charges $P_1^E, \ldots, P_K^E$ such that, for all $k = 1, \ldots, K$ and $a \in C_k^E$, $J(a) = \int a\, dP_k^E \equiv P_k^E(a)$.

Clearly, $P_k^E(E) = 1$ for all k. To see this, consider $a \in \operatorname{int} C_k^E$; then, for $\epsilon > 0$ small, $a + 1_{S \setminus E}\epsilon \in \operatorname{int} C_k^E$, so $J(a + 1_{S\setminus E}\epsilon) = P_k^E(a + 1_{S\setminus E}\epsilon) = P_k^E(a) + \epsilon P_k^E(S \setminus E)$; since $J(a) = J(a + 1_{S\setminus E}\epsilon)$, it follows that $P_k^E(S \setminus E) = 0$.

It must now be verified that, for every $k \in \{1, \ldots, K\}$, Eq. (8) holds, and $P_k^E = P_{n_k}(\cdot|E)$ for some $n_k \in \{1, \ldots, N\}$. Fix k and consider the set $D_k = \{1_E[a - J(a)] \ : a \in C_k^E\}$. Then, for all $a \in C_k^E$, $I(1_E[a - J(a)]) = 0 = P_k^E(1_E[a - J(a)])$, so by Lemma 5.14, there exists $n_k \in \{1, \ldots, N\}$ such that $0 = I(1_E[a - J(a)]) = \int 1_E[a - J(a)]\, dP_{n_k}$ for every $a \in C_k^E$. Therefore, for each such a, adding $J(a)$ to each term yields

$$J(a) = I(aEJ(a)) = \int a\, E\, J(a)\, dP_{n_k} = P_{n_k}(E) \int a\, dP_{n_k}(\cdot|E) + [1 - P_{n_k}(E)]J(a);$$

since $P_{n_k}(E) > 0$, $J(a) = \int a\, dP_{n_k}(\cdot|E)$. Also, for all $a \in C_k^E$, if $aEJ(a) \in C_m$, then

$$\int aEJ(a)\, dP_m = I(aEJ(a)) = J(a),$$

i.e Eq. (8) holds; finally, since P_k^E is the unique measure representing $\succeq_E$ on C_k^E, $P_k^E = P_{n_k}(\cdot|E)$, and the proof of Theorem 3.2 is complete.

5.3 Proof of Proposition 3.4

Throughout this section, assume that $\succeq$ satisfies Axioms 1–5 and 6; to remind the reader of this fact, the expression "Under the maintained assumptions" will be used in the statement of intermediate results. Let I, u, $C_1, \ldots, C_N$ and $P_1, \ldots, P_N$ be as in Section 5.1. As in

§5.1.2, let $I(a) = \int a\, dP_n$ for all $a \in C_n$ and $n \in \{1, \ldots, N\}$; also write $P_n(a)$ for $\int a\, dP_n$. By Assumption 1, $u(X)$ is convex; assume w.l.o.g. that $u(X) \supset [-1, 1]$, as in §5.1.2.

By assumption, (S, Σ) is a standard Borel space, and μ is convex-valued. Hence, singleton sets are measurable, S is uncountable, and μ is continuous, i.e. $\mu(\{s\}) = 0$ for all $s \in S$. With reference to Axiom 9, it is clear that $f_k \downarrow f$ monotonely iff $I(u \circ f_k) \downarrow I(u \circ f)$.

5.3.1 Countable Additivity of μ; Borel Isomorphisms

Lemma 5.21 *Under the maintained assumptions, if $\succeq$ is probabilistically sophisticated with respect to μ and satisfies Axiom 9, μ is continuous at $\emptyset$, hence countably additive.*

Proof. Consider a sequence of events $\{A_k\}_{k\geq 1}$ such that $A_k \supset A_{k+1}$ and $\bigcap_{k\geq 1} A_k = \emptyset$. Let $x_1, x_0 \in X$ be such that $u(x_1) = 1$, $u(x_0) = 0$. Then, by Axioms 4 and 9, for every $x \in X$ such that $x \succ x_0$, there exists $K \geq 1$ such that $k \geq K$ implies $x \succ x_1\, A_k\, x_0$; moreover, clearly $x_1\, A_k\, x_0 \succeq x_0$. Now suppose $\mu(A_k) \downarrow \epsilon > 0$. Since μ is convex-ranged, there exists an event E such that $\mu(E) = \epsilon$; by Def. 3.3, since $\mu(\{s : x_1 A_k x_0(s) \preceq x\}) = 1 - \mu(A_k) \leq 1 - \mu(E) = \mu(\{s : x_1 E x_0(s) \preceq x\})$ for $x_1 \succ x \succeq x_0$, $x_1 A_k x_0 \succeq x_1 E x_0$. Similarly, for $x_1 \succ x \succeq x_0$, $\mu(\{s : x_1 E x_0(s) \preceq x\}) = 1 - \mu(E) < 1 = \mu(\{s : x_0(s) \preceq x\})$, so $x_1 E x_0 \succ x_0$. Since $u(X)$ is convex and $x_1 \succeq x_1\, E\, x_0 \succ x_0$, there exists x_ϵ such that $x_\epsilon \sim x_1\, E\, x_0$, and hence $x_1 A_k x_0 \succeq x_\epsilon \succ x_0$ for all $k \geq 1$: contradiction. Thus, $\mu(A_k) \downarrow 0$. ∎

Since μ is countably additive and continuous, the Borel isomorphism theorem for measures [19, Theorem 17.41] yields a bijection $\varphi : S \to [0, 1]$ such that φ and φ^{-1} are both Borel measurable, and the Borel measure m on $[0, 1]$ defined by $m(E) = \mu(\varphi^{-1}(E))$ for all Borel sets $E \subset [0, 1]$ is Lebesgue measure on $[0, 1]$. This implies that it is sufficient to prove Proposition 3.4 for the case $S = [0, 1]$, with Σ its Borel sigma-algebra.[33]

5.3.2 Countable additivity of $P_1, \ldots, P_N$; Continuous functions in $\operatorname{int} C_1, \ldots, \operatorname{int} C_N$

Lemma 5.22 *Under the maintained assumptions, $\succeq$ satisfies Axiom 9 if and only if, for every $n \in \{1, \ldots, N\}$, P_n is countably additive, and $I(a) = P_n(a)$ for every $a \in B(S, \Sigma)$ that*

[33]Suppose (S, Σ) is any (uncountable) standard Borel space; given φ as above, consider the map $T_\varphi : B(S, \Sigma) \to B([0, 1], \Sigma_{[0,1]})$, where $\Sigma_{[0,1]}$ is the Borel sigma-algebra on $[0, 1]$, given by $T_\varphi a = a \circ \varphi^{-1}$. Then T_φ is an isometric isomorphism between $B(S, \Sigma)$ and $B([0, 1], \Sigma_{[0,1]})$. Hence, if $C_1, \ldots, C_N$ is a proper covering of $B(S, \Sigma)$, then $T_\varphi C_1, \ldots, T_\varphi C_N$ is a proper covering of $B([0, 1], \Sigma_{[0,1]})$. Also, for any probability charge P on (S, Σ), consider the probability charge $Q = P \circ \varphi^{-1}$ on $([0, 1], \Sigma_{[0,1]})$; then, for any $b \in B([0, 1], \Sigma_{[0,1]})$,$\int_{[0,1]} b\, dQ = \int_S b \circ \varphi\, d(Q \circ \varphi) = \int_S a \circ \varphi\, dP$. Finally, let L' be the set of acts from $[0, 1]$ to X, and define $\succeq'$ over L' by $f' \succeq g'$ iff $f' \circ \varphi \succ g' \circ \varphi$. Then $\succeq'$ admits a representation as in Theorem 2.6; its plausible priors $Q_1, \ldots, Q_N$ are defined by $Q_n = P_n \circ \varphi^{-1}$. Finally, if $Q_n = \mu \circ \varphi^{-1}$, then $P_n = \mu$.

is the pointwise limit of a monotonically decreasing sequence of elements of C_n.

Proof. (Only if): Fix n and $a \in \operatorname{int} C_n$ such that $\|a\| < 1$; thus, there exists $\epsilon > 0$ such that $\|b - a\| < \epsilon$ implies $b \in C_n$ and $\|b\| < 1$, so there exists $g \in L$ such that $b = u \circ g$.

Now consider a sequence of events $\{A_k\}$ such that $A_k \supset A_{k+1}$ for all k, and $\bigcap_k A_k = \emptyset$. For each k, let $a_k = a + \frac{\epsilon}{2} 1_{A_k}$; then $\|a_k - a\| = \frac{\epsilon}{2} < \epsilon$, so $a_k \in C_n$, and furthermore there exists a sequence $\{f_k\} \subset L$ such that $a_k = u \circ f_k$ for all k. Clearly, $a_k(s) \geq a_{k+1}(s)$ for all k and s, so $I(a_k) \geq I(a_{k+1})$; also, $f_k(s) \succeq f_{k+1}(s)$ for all s. Furthermore, for every $s \in S$, there is $K(s)$ such that $k \geq K(s)$ implies $a_k(s) = a(s)$; thus, $a_k(s) \downarrow a(s)$ for all s, hence $f_k(s) \downarrow f(s)$ for all s, and Axiom 9 implies that $f_k \downarrow f$, or equivalently $I(a_k) \downarrow I(a)$. Therefore,

$$P_n(A_k) = P_n(1_{A_k}) = \frac{2}{\epsilon} P_n(a_k - a) = \frac{2}{\epsilon}[P_n(a_k) - P_n(a)] = \frac{2}{\epsilon}[I(a_k) - I(a)] \downarrow 0,$$

i.e. P_n is continuous; thus, P_n is countably additive.

Now consider a sequence $\{a_k\} \subset C_n$ such that $a_k(s) \downarrow a(s)$ for all s. Then $I(a) = \lim_k I(a_k) = \lim_k P(a_k) = P(a)$, where the first equality follows from Axiom 9, and the last from Monotone Convergence.

(If): omitted (not required for the proof of Proposition 3.4). ■

Lemma 5.23 *Under the maintained assumptions, if $\succeq$ satisfies Axiom 9, then, for every $n \in \{1, \ldots, N\}$, the interior of C_n contains a continuous function.*

Proof. Since every C_n is affine, wlog restrict attention to $C_n \cap B^1$, where B_1 denotes the closed unit ball of $B(S, \Sigma)$, viewed as the set of all Borel-measurable functions $a : S \to [-1, 1]$. Also let $C^1 \subset B^1$ denote the continuous functions in B^1. Begin with two preliminaries.

1. Let $\mathcal{B}_1$ be the set of pointwise limits of functions in C^1. Then, by Kechris [19, Theorem 24.10 and Exercise 24.13], $\mathcal{B}_1$ is the set of functions of Baire class 1. Next, for any ordinal ξ such that $1 < \xi < \omega_1$ (where ω_1 denotes the first uncountable ordinal), let $\mathcal{B}_\xi$ be the set of functions of Baire class ξ, i.e. pointwise limits of sequences $\{a_k\} \subset B^1$, where for each k, $a_k \in \mathcal{B}_{\xi_k}$ for some $\xi_k < \xi$. Then, by Kechris [19, Theorem 24.3], $B^1 = \bigcup_{\xi < \omega_1} \mathcal{B}_\xi$. It is easy to show by induction that every Baire class is closed under multiplication by a scalar.

2. Consider the linear operator $T : B(S, \Sigma) \to \mathbb{R}^N$ defined by $T(a) = (P_1(a), \ldots, P_N(a))$ for all $a \in B(S, \Sigma)$. Clearly, $T(B(S, \Sigma)) \equiv \mathcal{R}$ is a normed linear subspace of $\mathbb{R}^N$. Then (e.g. Megginson [27, Exercise 1.46]) T is an open mapping, so for every $n = 1, \ldots, N$, $T(\operatorname{int}(C_n \cap B^1))$ is open in $\mathcal{R}$.

Now consider $n \in \{1, \ldots, N\}$. For notational simplicity, let $V_n \equiv \operatorname{int}(C_n \cap B^1)$; observe that that $V_n \neq \emptyset$, and that $T(V_n)$ is open in $\mathcal{R}$. Now suppose V_n does not contain any continuous function. It will be shown that then $V_n = \emptyset$, a contradiction.

Suppose first that there exists $a \in \mathcal{B}_1 \cap V_n$; then there is a sequence $\{a_k\} \subset C^1$ such that $a_k(s) \to a(s)$ for all $s \in S$. Since $\|a_k\| \le 1$, by Dominated Convergence $P_m(a_k) \to P_m(a)$ for all $m \in \{1, \ldots, N\}$; that is, $T(a_k) \to T(a)$. Since $T(V_n)$ is open in $\mathcal{R}$ and $\{T(a_k)\} \subset \mathcal{R}$, there is K such that $T(a_k) \in T(V_n)$ for all $k \ge K$. In other words, for every such k, there is $b_k \in V_n$ (not necessarily also in $\mathcal{B}_1$) such that $T(a_k) = T(b_k)$, i.e. $P_m(a_k) = P_m(b_k)$ for all m. Lemma 5.17 then implies that also $a_k \in \operatorname{int} C_n$; hence, either a_k or, if $\|a_k\| \ge 1$, e.g. $\frac{a_k}{2\|a_k\|}$ lie in $\operatorname{int}(C_n \cap B^1) = V_n$ (cf. Lemma 5.12 Part 2) which contradicts the assumption that V_n does not contain any continuous function. Thus, $V_n \cap \mathcal{B}_1 = \emptyset$.

By induction, consider an ordinal ξ such that $1 < \xi < \omega_1$ and suppose that $V_n \cap \mathcal{B}_{\xi'} = \emptyset$ for all $1 \le \xi' < \xi$. Suppose that $a \in V_n \cap \mathcal{B}_\xi$, so there is a sequence $a_k \to a$ such that $a_k \in \mathcal{B}_{\xi_k}$ and $1 \le \xi_k < \xi$ for each k. As above, $T(a_k) \in T(V_n)$ for large k, so by Lemma 5.17 either a_k or e.g. $\frac{a_k}{2\|a_k\|}$ lie in V_n. But $a_k \in \mathcal{B}_{\xi_k}$, and similarly, if $\|a_k\| \ge 1$, $\frac{a_k}{2\|a_k\|} \in \mathcal{B}_{\xi_k}$, which contradicts the assumption that $V_n \cap \mathcal{B}_{\xi_k} = \emptyset$. Thus also $V_n \cap \mathcal{B}_\xi = \emptyset$.

It follows that $V_n \cap \bigcup_{\xi<\omega_1} \mathcal{B}_\xi = \bigcup_{\xi<\omega_1}(V_n \cap \mathcal{B}_\xi) = \emptyset$; since $\bigcup_{\xi<\omega_1} \mathcal{B}_\xi = B^1$, it follows that $V_n = \emptyset$, as claimed: contradiction. ■

5.3.3 Main Result

Fix $n \in \{1, \ldots, N\}$. By Lemma 5.23, the interior of C_n contains a continuous function, denoted c; by Lemma 5.12 Part 2, it is w.l.o.g. to assume that $\inf_s c(s) = 0$ and $\sup_s c'(s) = 1$ [e.g. consider any continuous $c' \in \operatorname{int} C_n$; if c' is constant, let $c = c'$; otherwise, let $c = \frac{c' - \inf_s c'}{\sup_s c' - \inf_s c'}$; the Lemma guarantees that this point will also lie in the interior of C_n]. Note also that, since $S = [0, 1]$, minima and maxima are attained.

If c is constant, then $\succeq$ is easily seen to be a SEU preference.[34] In particular, it admits a unique plausible prior, and it is straightforward to show that this prior must coincide with μ. Thus, assume c is nonconstant. Since $c \in \operatorname{int} C_n$, there exists $\epsilon > 0$ such that $\sup_{s\in T} |a(s) - c(s)| = \|a - c\| < 2\epsilon$ imply $a \in C_n$; fix such an $\epsilon > 0$ throughout.

Also, since $0 \le c(s) \le 1$ for all $s \in S$, c is the uniform limit of the sequence of step functions $\{a_M\}_{M\ge1}$ defined by

$$a_M(s) = \begin{cases} \frac{1}{M}(m-1) & s \in E_m \equiv \{s : c(s) \in [\frac{m-1}{M}, \frac{m}{M})\}, \text{ for } m = 1, \ldots, M-1 \\ \frac{1}{M}(M-1) & s \in E_M \equiv \{s : c(s) \in [\frac{M-1}{M}, 1]\}. \end{cases}$$

For $M > \frac{1}{\epsilon}$, $\|a_M - c\| = \frac{1}{M} < \epsilon$ (hence, $a_M \in C_n$) and furthermore $\min\{a_M(s) - a_M(t) : s, t \in T,\, a_M(s) > a_M(t)\} = \frac{1}{M} < \epsilon$. Fix such a value of M throughout, and let $f \in L$ be a simple act such that $u \circ f = a_M$; write $f = (x_1, E_1; \ldots x_M, E_M; -\frac{1}{2}, S \setminus T)$, where $u(x_m) = \frac{1}{M}(m-1)$.

[34] Suppose $c = 1_S\gamma$ for some $\gamma \in \mathbb{R}$; fix $a \in B(S, \Sigma)$: since $\gamma \in \operatorname{int} C_n$, for some $\alpha \in (0, 1]$, $\alpha a + (1-\alpha)\gamma \in C_n$, so $\alpha I(a) + (1-\alpha)\gamma = I(\alpha a + (1-\alpha)\gamma) = P_n(\alpha a + (1-\alpha)\gamma) = \alpha P_n(a) + (1-\alpha)\gamma$, i.e. $I = P_n$.

Since $S = [0,1]$ is connected and c is continuous, $c([0,1])$ is connected; and since $\max_s c(s) = 0$ and $\min_s c(s) = 1$, $c([0,1]) = [0,1]$. Thus, for every m, the open set $c^{-1}(\frac{m-1}{M}, \frac{m}{M}) \subset E_m$ is non-empty; since μ has full support, $\mu(E_m) > 0$.

The remainder of the proof consists of two claims.

Claim 1. For any $m \in \{1, \ldots, M\}$, $P_n(E_m) > 0$ and $P_n(F) = \frac{\mu(F)}{\mu(E_m)} P_n(E_m)$ for all $F \in \Sigma$ such that $F \subset E_m$.

Proof: Fix m, and let $x \in X$ be such that $u(x) = u(x_m) + \frac{1}{M}$. Define the act f' by $f'(s) = f(s)$ for $s \notin E_m$, and $f'(s) = x$ for $s \in E_m$. Note that $\|u \circ f' - c\| \leq \|u \circ f' - u \circ f\| + \|u \circ f - c\| < 2\epsilon$, so $u \circ f' \in C_n$.

Then Def. 3.3 implies that $f' \succ f$, because, for x' such that $x_m \preceq x' \prec x$, $\mu(\{s : f'(s) \preceq x'\}) = \mu(\bigcup_{\ell=1}^{m-1} E_\ell) < \mu(\bigcup_{\ell=1}^{m} E_\ell) = \mu(\{s : f(s) \preceq x'\})$, and equality holds for all other x'. Hence, $P_n(u \circ f') = I(u \circ f') > I(u \circ f) = P_n(u \circ f)$, so $P_n(E_m) > 0$ as needed.

Next, by range convexity of μ, for every $K \geq 1$ there exists a partition $\{E_m^1, ..., E_m^K\}$ of E_m such that $\mu(E_m^k) = \frac{1}{K}\mu(E_m)$ for all $k = 1, \ldots, K$. For each such k, construct acts f^k such that $f^k(s) = f(s)$ for all $s \in S \setminus E_m^k$, and $f^k(s) = x$ for $s \in E_m^k$. Arguing as above, $u \circ f^k \in C_n$; furthermore, Def. 3.3 implies that $f^k \sim f^h$, hence $P_n(u \circ f^k) = I(u \circ f^k) = I(u \circ f^h) = P_n(u \circ f^h)$, for all $k, h \in \{1, \ldots, K\}$. Since f^k and f^h only differ on E_m^k and E_m^h, a simple calculation shows that $P_n(E_m^k) = P_n(E_m^h)$, so $P_n(E_m^k) = \frac{1}{K} P_n(E_m)$. Hence, the second part of the claim is true for all events $F \subset E_m$ such that $\frac{\mu(F)}{\mu(E_m)}$ is rational.

Now assume $\frac{\mu(F)}{\mu(E_m)}$ is irrational, and consider $r \in \mathbb{Q} \cap (\frac{\mu(F)}{\mu(E_m)}, 1]$. By range convexity of μ, there exists $F_r \in \Sigma$ such that $F_r \subset E_m \setminus F$ and $\frac{\mu(F)+\mu(F_r)}{\mu(E_m)} = r$,[35] so $P_n(F \cup F_r) = r P_n(E_m)$. Thus, $P_n(F) \leq r P_n(E_m)$ for all $r \in \mathbb{Q} \cap (\frac{\mu(F)}{\mu(E_m)}, 1]$, which implies that $P_n(F) \leq \frac{\mu(F)}{\mu(E_m)} P_n(E_m)$. Similarly, $P_n(F) \geq \frac{\mu(F)}{\mu(E_m)} P_n(E_m)$, so Claim 1 holds for all Borel $F \subset E$.

Claim 2. For any $m \in \{1, \ldots, M\}$, $P_n(F) = \frac{\mu(F)}{\mu(\bigcup_{\ell=1}^m E_\ell)} P_n(\bigcup_{\ell=1}^m E_\ell)$ for all $F \in \Sigma$ such that $F \subset \bigcup_{\ell=1}^m E_\ell$. Thus, in particular, $P_n = \mu$.

Proof: arguing by induction, the assertion follows from Claim 1 for $m = 1$; thus, assume that it holds for $m - 1 \geq 1$. Recall that $\mu(E_{m-1}) > 0$ and $\mu(E_m) > 0$; since μ is convex-ranged, there exist events $G_{m-1} \subset E_{m-1}$ and $G_m \subset E_m$ such that $\mu(G_{m-1}) = \mu(G_m) > 0$ [e.g. if $\mu(E_{m-1}) \leq \mu(E_m)$, let $G_{m-1} = E_{m-1}$ and choose G_m so $\mu(G_m) = \mu(E_{m-1})$, which is possible by range convexity; similarly for $\mu(E_{m-1}) > \mu(E_m)$.]

Now define an act f' by $f'(s) = f(s)$ for $s \in S \setminus (G_{m-1} \cup G_m)$, $f'(s) = x_m$ for $s \in G_{m-1}$, and $f'(s) = x_{m-1}$ for $s \in G_m$. Note that, by construction, $u(x_m) - u(x_{m-1}) = \frac{1}{M} < \epsilon$,

[35] Equivalently, F_r must satisfy $\mu(F_r) = r\mu(E_m) - \mu(F) \leq \mu(E_m) - \mu(F) = \mu(E_m \setminus F)$; so range convexity implies that such F_r can be found.

so $\|u \circ f' - c\| \leq \|u \circ f' - u \circ f\| + \|u \circ f - c\| < 2\epsilon$, hence $f' \in C_n$. Furthermore, $\mu(\{s : f'(s) = x_\ell\}) = \mu(\{s : f(s) = x_\ell\})$ for all $\ell = 1, ..., M$. This is obvious for $\ell < m-1$ or $\ell > m$; moreover, for $\ell = m-1$, by the choice of G_{m-1} and G_m,

$$\mu(\{s : f'(s) = x_{m-1}\}) = \mu([E_{m-1} \setminus G_{m-1}] \cup G_m) = \mu(E_{m-1}) - \mu(G_{m-1}) + \mu(G_m) = \mu(E_{m-1}),$$

and similarly for $\ell = m$. Therefore, $f \sim f'$, which implies $P_n(u \circ f) = P_n(u \circ f')$; since f, f' only differ on $G_{m-1} \cup G_m$, a simple calculation shows that $P_n(G_m) = P_n(G_{m-1})$. By Claim 1, $P_n(G_m) = \frac{\mu(G_m)}{\mu(E_m)} P_n(E_m)$; by the induction hypothesis, $P_n(G_{m-1}) = \frac{\mu(G_{m-1})}{\mu(\bigcup_{\ell=1}^{m-1} E_\ell)} P_n(\bigcup_{\ell=1}^{m-1} E_\ell)$. Conclude that $\frac{P_n(E_m)}{\mu(E_m)} = \frac{P_n(\bigcup_{\ell=1}^{m-1} E_\ell)}{\mu(\bigcup_{\ell=1}^{m-1} E_\ell)} \equiv \alpha$; thus,

$$\alpha = \frac{\mu(E_m)}{\mu(\bigcup_{\ell=1}^{m} E_\ell)} \frac{P_n(E_m)}{\mu(E_m)} + \frac{\mu(\bigcup_{\ell=1}^{m-1} E_\ell)}{\mu(\bigcup_{\ell=1}^{m} E_\ell)} \frac{P_n(\bigcup_{\ell=1}^{m-1} E_\ell)}{\mu(\bigcup_{\ell=1}^{m-1} E_\ell)} = \frac{P_n(E_m)}{\mu(\bigcup_{\ell=1}^{m} E_\ell)} + \frac{P_n(\bigcup_{\ell=1}^{m-1} E_\ell)}{\mu(\bigcup_{\ell=1}^{m} E_\ell)} = \frac{P_n(\bigcup_{\ell=1}^{m} E_\ell)}{\mu(\bigcup_{\ell=1}^{m} E_\ell)}.$$

Finally, consider an arbitrary $F \subset \bigcup_{\ell=1}^{m} E_\ell$. Then

$$\begin{aligned} P_n(F) &= P_n(F \cap \bigcup_{\ell=1}^{m-1} E_\ell) + P_n(F \cap E_m) = \frac{\mu(F \cap \bigcup_{\ell=1}^{m-1} E_\ell)}{\mu(\bigcup_{\ell=1}^{m-1} E_\ell)} P_n(\bigcup_{\ell=1}^{m-1} E_\ell) + \frac{\mu(F \cap E_m)}{\mu(E_m)} P_n(E_m) = \\ &= \mu(F \cap \bigcup_{\ell=1}^{m-1} E_\ell) \cdot \alpha + \mu(F \cap E_m) \cdot \alpha = \mu(F) \cdot \alpha = \frac{\mu(F)}{\mu(\bigcup_{\ell=1}^{m} E_\ell)} P_n(\bigcup_{\ell=1}^{m} E_\ell). \end{aligned}$$

References

[1] C. Aliprantis, K. Border, Infinite Dimensional Analysis, Springer Verlag, Berlin, 1994.

[2] F. J. Anscombe, R. J. Aumann, A definition of subjective probability, Annals of Mathematical Statistics 34 (1963) 199–205.

[3] T. Bewley, Knightian decision theory: Part I, Decisions in Economics and Finance 25 (2) (2002) 79–110, (first version 1986).

[4] A. Billot, A. Chateauneuf, I. Gilboa, J.-M. Tallon, Sharing beliefs: Between agreeing and disagreeing, Econometrica 68 (2000) 685–694.

[5] R. Casadesus-Masanell, P. Klibanoff, E. Ozdenoren, Maxmin expected utility over savage acts with a set of priors, J. Econ. Theory 92 (2000) 33–65.

[6] E. Castagnoli, F. Maccheroni, Restricting independence to convex cones, J. Math. Econ. 34 (2000) 215–223.

[7] E. Castagnoli, F. Maccheroni, M. Marinacci, Expected utility with multiple priors, mimeo, Università Bocconi and Università di Torino (July 2003).

[8] D. Ellsberg, Risk, ambiguity, and the Savage axioms, Quart. J. Econ. 75 (1961) 643–669.

[9] L. G. Epstein, A definition of uncertainty aversion, Rev. Econ. Stud. 66 (1999) 579–608.

[10] L. G. Epstein, J. Zhang, Subjective probabilities on subjectively unambiguous events, Econometrica 69 (2001) 265–306.

[11] P. Ghirardato, F. Maccheroni, M. Marinacci, Ambiguity from the differential viewpoint, Social Science Working Paper 1130, Caltech, (2002).
URL `http://www.hss.caltech.edu/ paolo/differential.pdf`

[12] P. Ghirardato, F. Maccheroni, M. Marinacci, M. Siniscalchi, A subjective spin on roulette wheels, Econometrica 71 (6) (2003) 1897–1908.

[13] P. Ghirardato, M. Marinacci, Ambiguity made precise: A comparative foundation, J. Econ. Theory 102 (2002) 251–289.

[14] I. Gilboa, D. Schmeidler, Maxmin expected utility with a non-unique prior, J. Math. Econ. 18 (1989) 141–153.

[15] I. Gilboa, D. Schmeidler, Updating ambiguous beliefs, J. Econ. Theory 59 (1993) 33–49.

[16] L. P. Hansen, T. Sargent, T. D. Tallarini, Robust permanent income and pricing, Rev. Econ. Stud. 66 (1999) 873–907.

[17] R. Holmes, Geometric Functional Analysis and its Applications, New York, Springer-Verlag, 1975.

[18] J.-Y. Jaffray, Dynamic decision making with belief functions, in: Y. et al. (Ed.), Advances in the Dempster-Shafer Theory of Evidence, Wiley, 1994.

[19] A. Kechris, Classical Descriptive Set Theory, New York, Springer-Verlag, 1995.

[20] P. Klibanoff, Characterizing uncertainty aversion through preference for mixtures, Soc. Choice Welfare 18 (2001) 289–301.

[21] P. Klibanoff, M. Marinacci, S. Mukerji, A smooth model of ambiguity aversion, mimeo, Northwestern University (2003).

[22] R. D. Luce, H. Raiffa, Games and Decisions, Wiley, New York, 1957.

[23] M. Machina, Robustifying the classical model of risk preferences and beliefs, Economics Discussion Paper 2002-06, UCSD (2002).

[24] M. J. Machina, D. Schmeidler, A more robust definition of subjective probability, Econometrica 60 (1992) 745–780.

[25] M. J. Machina, D. Schmeidler, Bayes without Bernoulli: Simple conditions for probabilistically sophisticated choice, J. Econ. Theory 67 (1995) 106–128.

[26] M. Marinacci, Probabilistic sophistication and multiple priors, Econometrica 70 (2002) 755–764.

[27] R. Megginson, An Introduction to Banach Space Theory, Springer-Verlag, New York, 1998.

[28] S. Mukerji, Ambiguity aversion and incompleteness of contractual form, Amer. Econ. Rev. 88 (1998) 1207–1231.

[29] K. Nehring, Ambiguity in the context of probabilistic beliefs, mimeo, UC Davis (November 2001).

[30] C. P. Pires, A rule for updating ambiguous beliefs, Theory Dec. 53 (2) (2002) 137–152.

[31] L. J. Savage, The Foundations of Statistics, Wiley, New York, 1954.

[32] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.

[33] M. Siniscalchi, Bayesian updating for general maxmin expected utility preferences, mimeo, Princeton University (September 2001).

[34] T. Wang, A class of multi-prior preferences, mimeo, University of British Columbia (September 2003).