A Behavioral Characterization of Plausible Priors

Marciano Siniscalchi*
Economics Department, Northwestern University, and
Economics Department, Princeton University.

May 2003

Abstract

Recent theories of choice under uncertainty represent ambiguity via multiple priors, informally interpreted as alternative probabilistic models of the uncertainty that the decision-maker considers equally plausible. This paper provides a robust behavioral foundation for this interpretation.

A prior P is deemed "plausible" if (i) preferences over a subset C of acts are consistent with subjective expected utility (SEU), and (ii) jointly with an appropriate utility function, P provides the *unique* SEU representation of preferences over C.

Under appropriate axioms, plausible priors can be elicited from preferences; moreover, if these axioms hold, (i) preferences are probabilistically sophisticated if and only if they are SEU, and (ii) under suitable consequentialism and dynamic consistency axioms, "plausible posteriors" can be derived from plausible priors via Bayes' rule. Several well-known decision models are consistent with the axioms proposed here.

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

---

*Email: `marciano@northwestern.edu`. 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 and Wolfgang Pesendorfer, as well as several RUD 2002 participants, for helpful discussion. All errors are my own.

1 Introduction

Multiple priors, i.e. 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. Consider for instance the following version of the Ellsberg [8] experiment: an urn contains 90 balls, of which 30 are red and 60 are green or blue, 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). While these preferences violate SEU, they are consistent with the maxmin expected-utility (MEU) decision model first axiomatized by Gilboa and Schmeidler [15]. According to this model, for all "acts" f, g mapping realizations to prizes, f is weakly preferred to g if and only if

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

where u is a utility index and $\mathcal{P}$ a set of "priors". In order to rationalize the above preferences, assume that $u(\$10) > u(\$0)$, and let $\mathcal{P}$ be 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).

The literature suggests that multiple priors may be interpreted as alternative probabilistic models of the underlying uncertainty, all equally plausible from the decision-maker's point of view. Indeed, a multiplicity of plausible priors is often interpreted as a direct consequence of ambiguity.[1] For conciseness, call this the *intuitive interpretation* of multiple priors.

This paper provides robust behavioral foundations for this interpretation. To this end, it identifies a formal notion of "plausible prior" that is both *robust* to different assumptions about the decision-maker's attitudes towards ambiguity, and *behavioral*, i.e. independent of the functional representation of overall preferences.

Simple, canonical examples of choice in the presence of ambiguity demonstrate the need for a robust, behavioral notion of "plausible prior". Return to the urn experiment; consider a decision-maker, henceforth referred to as "Ann", whose preferences are as described above: that is, they are consistent with the MEU model, with priors $\mathcal{P}$. Then, Ann's preferences are also consistent with other decision models, featuring different sets of priors. For instance,

[1] See e.g. Ellsberg [8, p 661]), Gilboa and Schmeidler [15, p. 142] and Schmeidler [30, p. 584]; also cf. Luce and Raiffa [22, pp. 304-305], and Bewley [3].

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

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

where $\mathcal{Q}$ is a set of priors and $\alpha \in [0,1]$.[2] 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 of Ann's preferences.[3]

The MEU representation of Ann's preferences might lend some support to the claim that $\mathcal{P}$ is the set of all priors she deems "plausible". But the $\frac{3}{4}$-MEU representation of Ann's preferences lends just as much support to the claim that $\mathcal{Q}$ is the set of plausible priors. Absent any formal behavioral criterion, deciding whether $\mathcal{P}$ or $\mathcal{Q}$ (or neither) is the "right" set reduces to a choice between alternative functional representations of the same preferences. But this diminishes the behavioral content and appeal of the intuitive interpretation of multiple priors. Hence the need for robust behavioral foundations.

Furthermore, the intuitive interpretation of priors is often invoked to suggest that specific functional representations of preferences reflect distinct attitudes towards ambiguity; for instance, the maxmin criterion may suggest "extreme pessimism". If sets of priors are identified only by the choice of a specific functional form, this intuition is questionable. But if a robust behavioral foundation for the notion of plausible priors is provided, this intuition can be made rigorous. This provides a secondary motivation for the present paper.

An alternative response to these observations is to regard the intuitive interpretation of multiple priors simply as a loose and informal "rationale" for certain functional representations of preferences. But this interpretation has played a central role in motivating both theoretical and applied research on ambiguity (cf. Footnote 1). Furthermore, it underlies and facilitates the economic interpretation of several results in applications featuring multiple priors.[4] These considerations alone justify an attempt to identify conditions under which the intuitive interpretation is essentially correct.

To complement the behavioral definition of plausible priors, this paper proposes an axiomatic framework wherein preferences satisfy several important properties:

---

[2] An axiomatization is provided by Ghirardato, Maccheroni and Marinacci [12].

[3] α-MEU-type representations featuring *arbitrarily small subsets* of $\mathcal{P}$ can also be constructed. Moreover, Section 6.1 in the Online Appendix shows that similar constructions are feasible for all MEU preferences.

[4] See e.g. Mukerji [26, p. 1209], Epstein and Wang [9, p. 289], Hansen, Sargent and Tallarini [17, p. 878], Billot, Chateauneuf, Gilboa and Tallon [4, p.686].

- The main result of this paper shows that, under the proposed axioms, plausible priors may be uniquely derived from preferences.
- The axioms I consider are compatible with a variety of known decision models, so the proposed definition of plausible prior is indeed robust (cf. Section 2.4). *A fortiori*, it is independent of any specific functional representation of preferences.
- A preference relation that satisfies the proposed axioms ("plausible-priors preference" henceforth) is probabilistically sophisticated in the sense of Machina and Schmeidler [23] *if and only if* it is consistent with SEU—and hence admits a single plausible prior.
- The class of plausible-priors preferences is closed under Bayesian updating. Consider an "unconditional" plausible-priors preference, and a "conditional" preference that is also characterized by "plausible posteriors". I identify necessary and sufficient consistency axioms for plausible posteriors to be related to plausible priors via Bayesian updating.
- Finally, the collection of plausible priors fully characterizes preferences: the ordering of any two acts is fully determined by their expected utilities computed with respect to each plausible prior (as is the case e.g. for MEU preferences).

The proposed definition of "plausible priors".

This paper adopts the Anscombe-Aumann [1] framework (although this is not essential: see Section 2); thus, acts map states to lotteries over prizes.

I deem a probability measure P over states of nature a *plausible prior* if there exists a (suitable) subset C of acts that satisfies the following two conditions:

(i) the decision-maker's preferences among acts in C satisfy the Anscombe-Aumann [1] axioms (and hence are consistent with SEU); and

(ii) a *unique* subjective probability can be derived from preferences among acts in C, and coincides with P.

More succinctly, P is a plausible prior if it uniquely rationalizes the decision-maker's choices over a subset of acts. Informally, a prior is deemed plausible if the decision-maker behaves as if she "actually used it" to rank a sufficiently rich subset of acts.

The uniqueness requirement in Condition (ii) is relatively demanding, but essential. The axiomatic framework adopted here ensures that this requirement can be met.

To illustrate the definition, return to the three-color-urn experiment. The set of priors that Ann considers plausible is $\{P_1, P_2\}$, where $P_1(r) = P_2(r) = \frac{1}{3}$ and $P_1(g) = P_2(b) = \frac{1}{6}$.

Let C_1 be the set of acts f such that $f(b)$ is weakly preferred to $f(g)$: thus, $f \in C_1$ if and only if $u(f(g)) \leq u(f(b))$. Regardless of which representation of Ann's preferences one chooses to employ, it is easy to verify that acts in C_1 are ranked according to their expected utility with respect to the probability P_1;[5] hence, preferences over C_1 are consistent with SEU. Moreover, P_1 *is* the only probability distribution on $\{r, g, b\}$ that is compatible with Ann's preferences over C_1.Similar arguments hold for the set C_2 of acts f such that $f(g)$ is weakly preferred to $f(b)$; preferences over C_2 uniquely identify the plausible prior P_2. Finally, while there exist other sets C of acts for which Condition (i) holds (a trivial example is the set of all constant acts), Condition (ii) fails for such sets (for example, any probability distribution provides a "SEU representation" of preferences over constant acts).

Condition (i) is expressed solely in terms of preferences: whether or not a set C has the required properties is independent of specific functional representations of the decision-maker's overall preferences. Condition (ii) then establishes the connection between the plausible prior P and preferences. Thus, the proposed definition is indeed fully behavioral.

The uniqueness requirement in Condition (ii) is essential in supporting the claim that a candidate prior is considered plausible *regardless of any specific functional representation of overall preferences*. As noted above, if uniqueness were dropped in Condition (ii), one would conclude that any prior is "plausible" in the three-color-urn example.[6] Further restrictions might of course be imposed by *arbitrarily* adopting a specific representation of Ann's overall preferences (e.g. MEU) and insisting that "plausible priors" be elements of the corresponding set of priors (e.g. $\mathcal{P}$). But these restrictions would be motivated solely by non-behavioral, functional-form considerations. The uniqueness requirement in Condition (ii) eliminates this difficulty and preserves the behavioral character of the proposed definition.

Organization of the paper

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.

[5] For the $\frac{3}{4}$-MEU representation, consider an act f such that $u(f(g)) < u(f(b))$; then $\arg\min_{q\in\mathcal{Q}} \int u(f(s))\, q(ds) = \{q_1\}$, where $q_1(r) = \frac{1}{3}$, $q_1(g) = \frac{2}{3}$; also, $\arg\max_{q\in\mathcal{Q}} \int u(f(s))\, q(ds) = \{q_2\}$, with $q_2(r) = \frac{1}{3}$, $q_2(g) = 0$. Hence, f is effectively evaluated using the prior $\frac{3}{4}q_1 + \frac{1}{4}q_2 = P_1$.

[6] For a less extreme example, consider the collection $C_1 \cap C_2$ of acts f such that Ann is indifferent between $f(g)$ and $f(b)$. Dropping the uniqueness requirement, any probability q such that $q(r) = \frac{1}{3}$ could be deemed a "plausible prior".

2 Model and Characterization

2.1 Decision-Theoretic Setup

I adopt a slight variant of the Anscombe-Aumann [1] decision framework. Consider a set of states of nature S, endowed with a σ-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).

A regularity condition on the σ-algebra Σ is required. I assume that (S, Σ) is a *standard Borel space* (cf. Kechris [20], Def. 12.5). That is, there exists a Polish (separable, completely metrizable) topology $\mathcal{T}$ on S such that Σ is the Borel σ-algebra generated by $\mathcal{T}$. This assumption is unlikely to be restrictive in applications: examples of standard Borel spaces include all finite and countable sets, Euclidean n-space $\mathbb{R}^n$ and any Borel-measurable subset thereof, as well as spaces of sample paths in the theory of continuous-time stochastic processes. The Polish topology $\mathcal{T}$ mentioned above plays no role in the axioms.

As in the standard Anscombe-Aumann [1] setup, acts are Σ-measurable maps from S into Y that are bounded in preference. I assume that preferences are defined over all such acts. Formally, let $\succeq_0$ be a binary relation on Y, and 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 pointwise convergence for appropriate sequences of acts will be introduced: cf. e.g. Epstein and Zhang [10]. Say that a sequence $\{f_n\}_{n\geq 1} \subset L$ of acts is *uniformly bounded* ("u.b.") if there exist $y, y' \in Y$ such that, for all $s \in S$ and $n \geq 1$, $y \preceq f_n(s) \preceq y'$. Say that the sequence *converges in preference* to an act $f \in L$ (written "$f_n \to f$") if, for all $y \in Y$ with $y \prec f$ (resp. $y \succ f$), there is $N \geq 1$ such that $n \geq N$ implies $y \prec f_n$ (resp. $y \succ f_n$). Finally, say that $\{f_n\}$ *converges pointwise in preference* to f (written "$f_n \xrightarrow{\forall s} f$") if, for all $s \in S$, $f_n(s) \to f(s)$.

It should be emphasized that the Anscombe-Aumann setup is adopted here merely as a matter of expository convenience. The analysis can be equivalently carried out in the "fully subjective" framework proposed by Ghirardato, Maccheroni, Marinacci and Siniscalchi [14]. A brief sketch of their approach is as follows. 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 (cf. [14], Preference Assumption A and Remark 4, and references therein), 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 axioms stated below can then be restated simply by replacing the "objective" Anscombe-Aumann mixture operation with subjective mixtures. This yields a characterization of plausible-priors preferences in a fully subjective environment.

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). Next, I discuss the notion of hedging against ambiguity and define robust mixture neutrality (§2.2.3); the latter is then employed to formulate the key axiom for preferences that admit plausible priors: "No Local Hedging" (Axiom 7 in §2.2.4). Two final regularity conditions (Axioms 8 and 9) are introduced in §2.2.5.

2.2.1 Basic Structural Axioms

The first five axioms will be applied both to the set L of all acts, and to certain subsets of L. I state them using intentionally vague expressions such as "for all acts f, g..." to alert the reader to this fact.

Axioms 1–4 appear in Gilboa and Schmeidler [15] and Schmeidler [30], as well as in "textbook" treatments of the Anscombe-Aumann characterization result; Axiom 5, due to Gilboa and Schmeidler [15], weakens the standard Independence requirement by imposing invariance of preferences to mixtures with constant acts only.

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 $\alpha \in (0,1)$: $f \succeq g$ implies $\alpha f + (1-\alpha)y \succeq \alpha g + (1-\alpha)y$.*

2.2.2 Mixture Neutrality and Plausible Priors

Recall that, loosely speaking, a prior P is deemed plausible if (i) preferences satisfy the Anscombe-Aumann axioms for SEU on a subset $C \subset L$ of acts, and (ii) P is the unique probability that yields a SEU representation of preferences on C. The Anscombe-Aumann characterization of SEU employs Axioms 1–4, plus the standard Independence axiom in lieu of Axiom 5. An alternative characterization employs Axioms 1–5 plus an additional property, *Mixture Neutrality.* Thus, since this paper considers preferences that satisfy Axioms 1–5, condition (i) corresponds to the requirement that Mixture Neutrality hold on C.

This subsection formalizes the preceding discussion. First, the following notation and terminology is convenient. Two acts f and g are *mixture-neutral* if the decision-maker is indifferent between any mixture $\alpha f + (1-\alpha)g$ and the corresponding mixture $\alpha y + (1-\alpha)y'$ of lotteries equivalent to them, i.e. such that $y \sim f$ and $y' \sim g$. In light of Axiom 5, C-Independence, this requirement can be formalized as follows.

Definition 2.1 (Mixture-neutral acts) Two acts $f, g \in L$ are *mixture-neutral* ($f \simeq g$) if and only if, for every $y \in Y$, $g \sim y$ implies $\alpha f + (1-\alpha)g \sim \alpha f + (1-\alpha)y$ for every $\alpha \in [0,1]$.

Axiom 6 (Mixture Neutrality) *For all acts f, g, $f \simeq g$.*

A note on terminology: "mixture neutrality" refers to the property of acts in Def. 2.1; "Mixture Neutrality" (capital initials) refers to Axiom 6.

The connection between mixture neutrality and ambiguity is discussed at length in the next subsection. The following proposition confirms that Mixture Neutrality is the key axiom characterizing SEU preferences in the class of preferences that satisfy Axioms 1–5. Furthermore, this is the case whether axioms are applied to the entire set L of acts, or to an appropriate subset thereof. This validates the proposed approach.[7]

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 and 6 for acts in C.
(2) there exists a probability charge P on (S, Σ), and an affine, cardinally unique function $u : Y \to \mathbb{R}$, such that, for all acts $f, g \in C$, $f \succeq g$ if and only if $\int u(f(\cdot))\, dP \geq \int u(g(\cdot))\, dP$.

Proposition 2.2 does *not* guarantee that the probability charge P in (2) is unique. This is explicitly required in the formal definition of plausible prior, which can finally be stated.

[7]This result is standard if the set C in the statement below equals L (or the collection of simple acts); see Remark 5 in Subsection 5.1.5 of the Appendix for a sketch of the argument in the general case.

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 $\succeq$ if and only if there exists a convex subset C of L containing all constant acts and such that

(i) $\succeq$ satisfies Axiom 6 on C;

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

I emphasize that a plausible prior is *not* required to be countably additive.

2.2.3 Hedging and Robust Mixture Neutrality

Gilboa and Schmeidler [15] and Schmeidler [30] suggest that an ambiguity-averse (MEU) decision-maker may violate Mixture Neutrality when contemplating a mixture of two acts f, g that provides a *hedge* against perceived ambiguity.[8] On the other hand, if (as prescribed by the definition of plausible prior) a decision-maker exhibits SEU preferences over a set C of acts, hedging considerations must be irrelevant for mixtures of acts in C. Thus, loosely speaking, preferences for which plausible priors can be elicited may exhibit arbitrary "global" attitudes towards ambiguity, but are "locally" indistinguishable from SEU preferences.

Axiom 7 ("No Local Hedging"), to be introduced in §2.2.4, characterizes this specific aspect of preferences, building on the notion of *robust mixture neutrality*. This subsection formalizes this notion, and clarifies its interpretation by means of examples that will also prove useful in motivating the axiom.

Notation. All examples in this section employ a finite state space $S = \{s_1, \ldots, s_N\}$ and a common set of prizes $X = \{\$0, \$10\}$. The set of probability distributions on S is denoted by $\Delta(S)$. Also, a lottery $y \in Y$ can be identified with the probability of receiving the prize \$10; similarly, an act f is represented by a tuple $(f(s_1), \ldots, f(s_N)) \in [0,1]^N$, where $f(s_n)$ is the probability of receiving the prize \$10 in state s_n. Axioms 1–5 imply that the decision-maker has EU preferences over Y, so such tuples can also be interpreted as utility profiles.

A mixture of two acts f, g intuitively provides a hedge against ambiguity if it *reduces outcome variability across ambiguous events* relative to *both* f and g. Example 1 is a particularly simple instance of this phenomenon; Examples 2 and 3 are possibly more interesting.

[8] 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).

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).

The acts $f_g = (0,1,0)$ and $f_b = (0,0,1)$ exhibit considerable variability across the intuitively ambiguous events $\{g\}$ and $\{b\}$; however, they vary in opposite, hence complementary directions—$f_g(g) \succ f_g(b)$ and $f_b(g) \prec f_b(b)$. Consequently, their strict mixtures display less variability across $\{g\}$ and $\{b\}$ than both f_g and f_b: indeed, $\frac{1}{2}f_g + \frac{1}{2}f_b$ is constant on $\{g,b\}$. Consistently with this hedging intuition, $f_g \not\approx f_b$.

On the other hand, two acts f, f' are mixture-neutral if and only if $f(g) \succeq f(b)$ and $f'(g) \succeq f'(b)$, or $f(g) \preceq f(b)$ and $f'(g) \preceq f'(b)$. Mixtures of such acts will display *more* outcome variability across the intuitively ambiguous events g and b than one of the acts f, f', and less variability than the other.[9]

Example 2 (Bob) (cf. Klibanoff [21], Ex. 1). A ball is drawn from an urn containing an equal, but unspecified number of red and blue balls, and an unspecified number of green balls; thus, $S = \{r,g,b\}$. Bob has MEU preferences, with priors $\mathcal{Q}_B = \{q \in \Delta(S) : p(r) = p(b)\}$. Thus, the relative likelihood of $\{r\}$ vs. $\{b\}$ is unambiguous.

Let $f = (.2,.3,.5)$ and $f' = (.1,.4,.6)$. Then f and f' are comonotonic,[10] but the *expected* outcomes they yield on the ambiguous events $\{r,b\}$ and $\{g\}$ vary in complementary directions. Hence, their mixtures reduce or eliminate expected-outcome variability across $\{r,b\}$ and $\{g\}$. Consistently with this hedging intuition, $f \not\approx f'$.

On the other hand, two acts f, f' are mixture-neutral if and only if $\frac{1}{2}f(r) + \frac{1}{2}f(b) \succeq f(g)$ and $\frac{1}{2}f'(r) + \frac{1}{2}f'(b) \succeq f'(g)$, or $\frac{1}{2}f(r) + \frac{1}{2}f(b) \preceq f(g)$ and $\frac{1}{2}f'(r) + \frac{1}{2}f'(b) \preceq f'(g)$; mixtures of such acts do not reduce *expected* outcome variability across the ambiguous events $\{r,b\}$ and $\{g\}$ relative to *both* f and f', and hence provide no hedging opportunities.

While mixture neutrality correctly reflects absence of hedging opportunities for MEU preferences, this is *not* the case for more general preferences that satisfy Axioms 1–5. The following example demonstrates this.

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.

[9] For such acts, $|[\gamma f(g) + (1-\gamma)f'(g)] - [\gamma f(b) - (1-\gamma)f'(b)]| = \gamma|f(g) - f(b)| + (1-\gamma)|f'(g) - f'(b)|$.

[10] Two acts f, f' are comonotonic if there do not exist $s, s' \in S$ such that $f(s) \succ f(s')$ and $f'(s) \prec f'(s')$: see Schmeidler [30]. It may be shown that Axioms 1-5 and Mixture Neutrality restricted to comonotonic acts characterize Choquet-Expected Utility preferences.

Let $f = \left(1, \frac{2}{3}, \frac{1}{2}, \frac{1}{2}\right)$ and $f' = \left(\frac{1}{2}, \frac{1}{2}, 1, \frac{2}{3}\right)$. The *sets* of outcomes delivered by f and f' on the intuitively ambiguous events[11] $\{r, g\}$ and $\{b, w\}$ vary in complementary directions, so their mixtures reduce or eliminate the variability of *sets* of outcomes across these events. For instance, the mixture $\frac{1}{2}f + \frac{1}{2}f' = \left(\frac{3}{4}, \frac{7}{12}, \frac{3}{4}, \frac{7}{12}\right)$, yields an outcome in the set $\{\frac{3}{4}, \frac{7}{12}\}$ if either $\{r, g\}$ or $\{b, w\}$ obtains. Thus, by analogy with Examples 1 and 2, it seems at least conceivable that a mixture of f and g might have some hedging value; yet, $f \simeq g$.

On the other hand, the mixture neutrality of f and g is not "robust". Consider for instance a small perturbation of f, such as the act $f_\epsilon = \left(1 - \epsilon, \frac{2}{3}, \frac{1}{2}, \frac{1}{2}\right)$.[12] Mixtures of f_ϵ and g also exhibit less variability across the events $\{r, g\}$ and $\{b, w\}$ than do f_ϵ or g. Consistently with this intuition, it may be verified that $f_\epsilon \not\simeq g$.

Example 3 indicates that non-MEU preferences allow for knife-edge, non-robust instances of mixture-neutrality for acts that may conceivably provide hedging opportunities.[13] Appropriate "perturbations" of the acts under consideration may then be used to filter out such non-robust instances of mixture neutrality, and hence identify pairs of acts that genuinely provide no hedging opportunities. The definition of robust mixture neutrality reflects these considerations by identifying suitable perturbations.

For the present purposes, a *perturbation* of an act h_0 is a mixture $(1 - \gamma)h_0 + \gamma h$, where $h \in L$ is arbitrary and $\gamma \in (0, 1)$ should be thought of as "small".[14] Two acts $f, g \in L$ are deemed robustly mixture-neutral if every small but arbitrary perturbation of every *strict* mixture of f and g is mixture-neutral with both f and g:

Definition 2.4 (Robustly mixture-neutral acts) Two acts $f, g \in L$ are *robustly mixture-neutral* (written $f \approx g$) if and only if for every $h \in L$ and $h_0 \in \{\lambda f + (1 - \lambda)g : \lambda \in (0, 1)\}$, there exists $\gamma \in (0, 1)$ such that

$$\gamma h + (1 - \gamma)h_0 \simeq f, \quad \gamma h + (1 - \gamma)h_0 \simeq g.$$

To motivate Def. 2.4, note first that $f \approx g$ implies $f \simeq g$ (cf. Lemma 5.2, Part 5 in the Appendix). Moreover, for the MEU preferences in Examples 1 and 2, $f \simeq g$ implies

[11] These events are ambiguous according to the definition provided by Ghirardato and Marinacci [13]. The example can be slightly modified so as to ensure that they are also ambiguous according to the definition put forth by Epstein and Zhang [10].

[12] Formally, the restriction $\epsilon \in (0, \frac{1}{3})$ is required; informally, ϵ can be "arbitrarily small".

[13] The choice of f and g in Example 3 is "knife-edge" also in a different, but related sense. It can be shown that, for *any* α*-MEU preference with* $\alpha \neq \frac{3}{4}$, f and g are *not* mixture-neutral.

[14] The specific mixture-based notion of perturbation I adopt is sufficient for the present purposes, but it is *not* crucial to the analysis. One might consider more general perturbations, at the expense of introducing additional definitions and notation (e.g. a notion of "neighborhood" of the act h_0).

$f \approx g$: for such MEU preferences, mixture neutrality is always robust.[15] This suggests that the relation $\approx$ is the appropriate counterpart of $\simeq$ for non-MEU preferences. The main characterization result will further strengthen this interpretation.

Turn now to the specific requirements of Def. 2.4. Assume first that λ is close to 1.[16] A perturbation $\gamma h + (1-\gamma)h_0$ of $h_0 = \lambda f + (1-\lambda)g$ can then be viewed as a "two-stage" perturbation of f. The first stage entails perturbing f in the direction of g, and is necessary to ensure that no opportunity for hedging is introduced by perturbations, when none exist. For concreteness, in Example 1, $0 \simeq f_g$, and indeed $0 \approx f_g$; this is in accordance with the intuition that mixtures with constant acts do not provide any hedge against ambiguity.[17] However, $(1-\gamma)(0)+\gamma f_b \not\approx f_g$ for any positive value of γ; intuitively, perturbing the constant act 0 in the direction of f_b introduces spurious hedging opportunities relative to f_g. On the other hand, for every strict mixture $h_0 = \lambda(0) + (1-\lambda)f_g$ there exists a (sufficiently small) $\gamma \in (0,1)$ such that $[\gamma f_b + (1-\gamma)h_0](g) \succ [\gamma f_b + (1-\gamma)h_0](b)$, and hence the two-stage perturbation $[\gamma f_b + (1-\gamma)h_0](b)$ does not provide hedging opportunities relative to f_g.[18]

Perturbations of $h_0 = \lambda f + (1-\lambda)g$ for intermediate values of λ need also be considered in order to ensure that the relation $\approx$ can be viewed as extending $\simeq$ to non-MEU preferences. Observe that, by Def. 2.1, if $f \simeq g$, then for *all* strict mixtures $\lambda f + (1-\lambda)g$, it is the case that $\lambda f + (1-\lambda)g \simeq f$ and $\lambda f + (1-\lambda)g \simeq g$. Allowing for arbitrary $\lambda \in (0,1)$ ensures that robust mixture neutrality also satisfies this property.

2.2.4 The Key Axiom

As noted above, the two key axiom proposed in this paper characterizes preferences that are "locally" consistent with SEU, but are otherwise arbitrary.

Examples 1 and 2 suggest that the conditions identifying (robustly) mixture-neutral acts can be quite different for different preferences.[19] However, there is some commonality among all three examples: if, for two acts f and f', *the outcomes $f(s)$ and $f'(s)$ are nearly indifferent, i.e. "close", in every state s*, then f, f' are (robustly) mixture-neutral.

This is in accordance with the hedging intuition. For simplicity, assume that the state

[15] It can be shown that MEU preferences satisfy this property whenever the representing set of priors is the weak*-closed, convex hull of a finite or countable collection of points.

[16] Analogous arguments hold for λ close to 0.

[17] Recall that this is the motivation for Axiom 5.

[18] Note also that the act f_ϵ in Example 3 can be viewed as a "two-stage" perturbation of f: first, perturb f by considering the mixture $(1-\epsilon)f + \epsilon g \equiv h_0$, for $\epsilon > 0$ "small"; then, perturb h by considering the mixture $\frac{1}{2}h + \frac{1}{2}h_0$, where $h = \left(1 - \frac{3}{2}\epsilon, \frac{2}{3} + \frac{1}{6}\epsilon, \frac{1}{2} - \frac{1}{2}\epsilon, \frac{1}{2} + \frac{1}{2}\epsilon\right)$.

[19] Similar conditions can be provided for Example 3.

space S is finite. If two acts f, f' are sufficiently "close" to each other state-by-state, then they vary in the *same* direction across *all* disjoint events $E, F \subset S$, including potentially ambiguous events: if f yields better outcomes at states $s \in E$ than at states $s \in F$ (or, as in Examples 2 and 3 respectively, a better *expected* outcome, or better *sets* of outcomes, on E than on F) , then so does f'. Mixtures of f and f' therefore exhibit *more* variability across E and F than one of the acts f, f', and less than the other; hence, they provide no hedging opportunities. Therefore, if arbitrarily close acts are *not* (robustly) mixture-neutral, this must be a consequence of considerations other than the simple variability-reduction, or hedging intuition. The axiom I propose thus requires a degree of consistency with the latter.

In the simple setting of Examples 1–3, "closeness" is measured by the absolute difference between probabilities. Axiom 7 generalizes this by considering pointwise convergent sequences of acts.

Axiom 7 (No Local Hedging) *If $\{f_n\}_{n\geq 1} \subset L$ is u.b. and, for some $f \in L$, $f_n \xrightarrow{\forall s} f$, then there exists $N \geq 1$ such that $n \geq N$ implies $f_n \approx f$.*

The preceding discussion is subject to two qualifications. First, Axiom 7 allows for the possibility that two acts f, f' fail to be (robustly) mixture-neutral, even though they vary in the same direction across all disjoint ambiguous events; however, it *does* rule out the possibility that *all* acts f' that vary in the same direction as f across ambiguous events be perceived as providing hedging opportunities relative to f. Second, the discussion applies *verbatim* to infinite state spaces, if there are finitely many ambiguous events. More generally, Axiom 7 intuitively suggests that, for two acts f, f' to violate robust mixture neutrality, they must display "sufficiently complementary" patterns of variability across "sufficiently many" ambiguous events. For further discussion, see Section 4.1.

2.2.5 Additional Regularity Conditions

Axioms 1–5 and 7 should be viewed as capturing the key behavioral properties of preferences that admit plausible priors. It turns out that two additional regularity conditions are required. The first is only needed if the state space is infinite; the second is automatically satisfied in certain specific models of choice (see below), but must be explicitly stated if one wishes to identify plausible priors without imposing restrictions on ambiguity attitudes.

The first regularity condition requires a form of pointwise continuity for uniformly bounded sequences of acts; if S is finite, it is of course implied by Axiom 3.[20]

[20] Versions of this axiom appear in characterizations of SEU and probabilistically sophisticated preferences with countably additive probabilities: see e.g. Epstein and Zhang [10], as well as references therein.

Axiom 8 (Uniformly Bounded Continuity) *If $\{f_n\}_{n\geq 1} \subset L$ is u.b. and, for some act $f \in L$, $f_n \xrightarrow{\forall s} f$, then $f_n \to f$.*

Axiom 8 plays a key role in Subsection 5.1.4 of the Appendix. It also implies that plausible priors are countably additive, but this should be considered a side effect. The restriction to *uniformly bounded* sequences of acts is of course essential: even SEU preferences characterized by countably additive probabilities may fail to be continuous with respect to sequences of (bounded) acts that are not uniformly bounded.

To motivate Axiom 9, observe first that, for arbitrary MEU preferences, mixture neutrality satisfies the following "consistency" property:

Remark 1 Let $\succeq$ be a MEU preference relation. For any triple of acts f, g, h, if $f \simeq g$ and $h \simeq \gamma f + (1-\gamma)g$ for some $\gamma \in (0,1)$, then

$$\alpha f + (1-\alpha)h \simeq \beta g + (1-\beta)h.$$

for all $\alpha, \beta \in [0,1]$.[21] [22]

This property ensures that the no-hedging intuition applies to distinct pairs of mixture-neutral acts in a consistent manner, as the following example suggests.

Example 4 (Dave) $S = \{rr, rb, br, bb\}$, representing draws from two urns containing red and blue balls, in unknown proportions. Dave has MEU preferences, with priors $\mathcal{Q}_D = \Delta(S)$.

Let $f = (1,1,0,0)$, $g = (1,0,1,0)$, and $h = (0,1,1,0)$. It may be verified that $f \simeq g$. Since f and g yield the same outcomes in states rr and bb, their mixtures cannot provide a hedge against the ambiguity of $\{rr\}$ and $\{bb\}$. Such mixtures *can* reduce outcome variability across the ambiguous events $\{rb\}$ vs. $\{br\}$, but this turns out to be irrelevant for Dave.

Furthermore, $h \simeq \frac{1}{2}f + \frac{1}{2}g = \left(1, \frac{1}{2}, \frac{1}{2}, 0\right)$. In particular, mixtures of these acts do not provide a hedge against the ambiguity of $\{rr\}$ and $\{bb\}$, because h is constant on $\{rr, bb\}$.[23]

It follows that mixtures of f and h, as well as mixtures of g and h, all exhibit the same pattern of variability across $\{rr\}$ and $\{bb\}$. This suggests that $\alpha f + (1-\alpha)h \simeq \beta g + (1-\beta)h$ for all $\alpha, \beta \in [0,1]$. It may be verified that such acts are indeed mixture-neutral.

In other words, $f \simeq g$ and $\gamma f + (1-\gamma)g \simeq h$ suggest that f, g and h exhibit a common pattern of variability across at least one pair of ambiguous events, and that furthermore

[21] The cases $\alpha = \beta = 0$ and $\alpha = \beta = 1$ are uninteresting.

[22] The restriction $\gamma \in (0,1)$ is essential. In Example 1, $f_g \simeq (0, \frac{1}{4}, \frac{1}{4})$ and $(0, \frac{1}{4}, \frac{1}{4}) \simeq f_b$, but $f_g \not\simeq f_b$.

[23] That is: outcome variability is of course reduced relative to $\frac{1}{2}f + \frac{1}{2}g$, but not relative to h, so there is no hedging. Analogously, mixing any act $\bar{f}$ with a constant act y reduces variability relative to $\bar{f}$, but $\bar{f} \simeq y$.

this variability is crucial to the evaluation of these acts.[24] The mixtures $\alpha f + (1-\alpha)h$ and $\beta g + (1-\beta)h$ also display this pattern of variability, so mixtures of these composite acts should not provide hedging opportunities. Remark 1 confirms that, indeed, they do not.

As was discussed above, I suggest that robust mixture neutrality be viewed as the appropriate extension of mixture neutrality to non-MEU preferences. However, the preceding axioms are not sufficient to ensure that the former will satisfy a property analogous to Remark 1. Axiom 9 explicitly requires that such a property hold, and hence reinforces the suggested interpretation of robust mixture neutrality.

Axiom 9 (Hedging Consistency) *For all acts $f, g, h \in L$: if $f \approx g$ and, for some $\gamma \in (0,1)$, $h \approx \gamma f + (1-\gamma)g$, then $\alpha f + (1-\alpha)h \approx \beta g + (1-\beta)h$ for all $\alpha, \beta \in [0,1]$.*

2.3 The Main Result

Recall that the definition of plausible priors involves sets of acts restricted to which preferences satisfy Axiom 6, Mixture Neutrality. The utility profiles of acts in any such collection is a subset of the space $B(S,\Sigma)$ of bounded, Σ-measurable real functions on S. Theorem 2.6 indicates that, under the axioms proposed here, $B(S,\Sigma)$ can be covered by (the conic hull of) such sets of utility profiles; furthermore, the latter satisfy specific algebraic and topological properties. For expository convenience, these properties are listed in Definition 2.5.

Definition 2.5 A collection $\mathcal{C}$ of subsets of $B(S,\Sigma)$ is a *proper covering* if $\bigcup\{C : C \in \mathcal{C}\} = B(S,\Sigma)$, elements of $\mathcal{C}$ are not nested, and:

(1) every element $C \in \mathcal{C}$ is a convex cone with non-empty interior that contains the constant functions $\gamma 1_S$, for $\gamma \in \mathbb{R}$;

(2) for all distinct $C, C' \in \mathcal{C}$: if $c \in C \cap C'$ and $a, b \in C$ are such that $\alpha a + (1-\alpha)b = c$ for some $\alpha \in (0,1)$, then $a, b \in C'$;[25] and

(3) if a uniformly bounded sequence $\{a_n\}_{n\geq 1}$ in $B(S,\Sigma)$ converges pointwise to some $a \in B(S,\Sigma)$, then there exists $N \geq 1$ such that, for all $n \geq N$, there exists $C_n \in \mathcal{C}$ with $a_n, a \in C_n$.

[24] Mixtures of f,g and h may eliminate variability across other events, but this does not lead to departures from mixture neutrality. So, the residual variability must be a crucial concern for the decision-maker.

[25] That is, $C \cap C'$ is an extremal subset of C: see e.g. Holmes [18], §2.C.

Property (1) can be seen to correspond closely to Part (i) in the definition of a plausible prior; furthermore, nonemptiness of the interior ensures uniqueness of the elicited probability. Property (2) states that intersections of distinct cones are "small": in particular, it implies that they have empty interior. Finally, if $\mathcal{C}$ is finite, Property (3) can be replaced by the assumption that each set $C \in \mathcal{C}$ is closed with respect to uniformly bounded pointwise limits.

The main result of this paper can finally be stated; N denotes a generic index set.

Theorem 2.6 *Let (S, Σ) be a standard Borel space, and consider a preference relation $\succeq$ on L. The following statements are equivalent:*

(1) $\succeq$ satisfies Axioms 1–5, 7, and 8–9;

(2) There exist an affine function $u : Y \to \mathbb{R}$, a finite or countable proper covering $\{C_n : n \in N\}$ and a corresponding collection of probability measures $\{P_n : n \in N\}$ such that, for all $n, m \in N$ and $a \in C_n \cap C_m$, $\int a\, dP_n = \int a\, dP_m$, and furthermore, for all $f, g \in L$: (i) if $u \circ f \in C_n$ and $u \circ g \in C_m$ for some $n, m \in N$, then

$$f \succeq g \iff \int u \circ f\, dP_n \geq \int u \circ g\, dP_m; \tag{1}$$

(ii) if $f \approx g$, then there exists $n \in N$ such that $u \circ f, u \circ g \in C_n$; and (iii) if $\int u \circ f\, dP_n \geq \int u \circ g\, dP_n$ for all $n \in N$, then $f \succeq g$.

In (2), u is unique up to positive affine transformations, $\{C_n\}$ is unique, and for every $n \in N$, P_n is the unique probability charge such that (i) holds for acts in $\{f : u \circ f \in C_n\}$.

Henceforth, I will employ the expression *plausible-priors preference* to indicate a binary relation $\succeq$ on L for which (2) in Theorem 2.6 holds.

Preferences in Examples 1–4 all satisfy the axioms in (1). In the MEU examples (Ex. 1, 2 and 4), the probabilities $\{P_n\}$ in (2), i.e. the plausible priors, turn out to coincide with the priors indicated above. In Example 3, which employs α-MEU preferences, the set of plausible priors consists of all mixtures of degenerate probabilities of the form $\frac{3}{4}\delta_s + \frac{1}{4}\delta_{s'}$, for distinct states s, s'.[26]

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.[27]

[26] In Example 3, the $\frac{3}{4}$-MEU criterion associates the value $\frac{3}{4}\int u \circ f\, d\delta_s + \frac{1}{4}\int u \circ f\, d\delta_{s'}$ to acts f such that $f(s) \prec f(s'') \prec f(s')$ for all $s'' \notin \{s, s'\}$. This clearly corresponds to $\int u \circ f\, d[\frac{3}{4}\delta_s + \frac{1}{4}\delta_{s'}]$.

[27] I owe this interpretation to Mark Machina.

I now discuss Theorem 2.6 by listing a number of corollaries. First of all, each (countably additive) probability measure P_n appearing in (2) is a plausible prior; in particular, note that every P_n is uniquely determined by preferences over acts whose utility profile lies in C_n. Furthermore, 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 collection of plausible priors for $\succeq$ is given by $\{P_n : n \in N\}$ in (2).*

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 $n \geq 1$, 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

$$\alpha g + (1-\alpha) b_E \sim \alpha g + (1-\alpha)[\pi_E x + (1-\pi_E) x'].$$

Simple calculations then show that $\pi_E = P_n(E)$.

The "procedure" just described should be viewed merely as a thought experiment.[28] It 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.

Condition (ii) in Part (2) confirms that, under the axioms proposed here, robust mixture neutrality reflects a strong notion of absence of hedging opportunities: if $f \approx g$, then f and g belong to a set of acts over which preferences are consistent with SEU. Indeed, the converse of (ii) is implied by (i)-(iii). Thus:

Corollary 2.8 *Under the equivalent conditions of Theorem 2.6, for all $f, g \in L$, $f \approx g$ if and only if $u \circ f$, $u \circ g \in C_n$ for some $n \geq 1$.*

Part (ii) of Theorem 2.6 also implies that the elements of the proper covering $\{C_n\}$ have an important *maximality* property:

Corollary 2.9 *Suppose the equivalent conditions of Theorem 2.6 hold, and let $C \subset L$ be such that, for all $f, g \in C$, $f \approx g$. Then there exists $n \geq 1$ such that $\{u \circ f : f \in C\} \subset C_n$.*

[28] In practice, identifying points in the interior of each cone C_n seems non-trivial.

In particular, the domain of each plausible prior P_n cannot be extended beyond C_n.

Finally, Condition (iii) 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. Let $\mathcal{R} = \{(\int a\,dP_n)_{n\in N} : a \in B(S,\Sigma)\}$ be the collection of (possibly countably infinite) vectors of integrals of functions with respect to each plausible prior obtained in Theorem 2.6. If N is finite, $\mathcal{R}$ is a vector subspace of $\mathbb{R}^N$ that includes the diagonal $\{(\gamma,\ldots,\gamma) : \gamma \in \mathbb{R}\}$.[29] Then define a map $V : \mathcal{R} \to \mathbb{R}$ by

$$\forall a \in B(S,\Sigma), \quad V\left(\left(\int a\,dP_n\right)_{n\in N}\right) = \int a\,dP_{n^*}, \tag{2}$$

where n^* is such that $a \in C_{n^*}$. By (iii) in Theorem 2.6, V is well-defined and unique.[30]

The map V can be thought of as "selecting" an index n^* as a function of the vector $\varphi = (\varphi_n)_{n\in N} \in \mathcal{R}$, and associating to φ the value φ_{n^*}. 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.

The above claim can now be made precise. For all acts $f,g \in L$, $f \succeq g$ if and only if $V((\int u\circ f\,dP_n)_{n\in N}) \geq V((\int u\circ g\,dP_n)_{n\in N})$, so the ordering of f and g is *entirely determined by the vectors* $(\int u\circ f\,dP_n)_{n\in N}$ *and* $(\int u\circ g\,dP_n)_{n\in N}$. As a function of these alone, V selects indices n_f and n_g such that $f \succeq g$ if and only if $\int u\circ f\,dP_{n_f} \geq \int u\circ g\,dP_{n_g}$. Thus, plausible priors fully determine preferences.

The functional V can be interpreted as the decision-maker's "elicited" choice criterion. Since plausible priors are identified without assuming any specific functional representation of overall preferences, V arguably provides a behaviorally accurate description of the decision-maker's attitudes towards ambiguity.

2.4 Examples of Plausible-Priors Preferences

This section provides simple sufficient conditions for MEU (or 0-MEU), α-MEU and CEU preferences to permit the elicitation of plausible priors. Conceptually, these examples demonstrate that *Axiom 7 does not restrict the decision-maker's attitudes towards ambiguity*, and hence is compatible with a variety of decision models. From a practical standpoint, the

[29] If $N = \mathbb{N}$, it can be shown that $\mathcal{R}$ is a vector subspace of $B(\mathbb{N}, 2^{\mathbb{N}})$ that includes all constant functions.

[30] Additional functional properties of V can be easily established. V is *normalized*, i.e. $V(1_N) = 1$; *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 $\varphi \in \mathcal{R}$, $V(\alpha\varphi + \beta) = \alpha V(\varphi) + \beta$. Finally, if $\{\varphi^k\}$ is a sequence in $\mathcal{R}$ such that $\varphi^k_n \to \varphi_n$ for all $n \in N$, and $\alpha \leq \varphi^k_n \leq \beta$ for some $\alpha,\beta \in \mathbb{R}$ and all $n \in N$, $k \geq 1$, then $V(\varphi^k) \to V(\varphi)$.

results in this subsection may be directly invoked in applications to ensure that the intuitive interpretation of priors is behaviorally sound.

Unless otherwise noted, no cardinality restriction is imposed on the state space S. For simplicity, I focus on preferences that admit a *finite* collection of plausible priors. These are characterized by the following additional axiom (see Online Appendix, §6.2.2):

Axiom 10 *If $\{f_n\}_{n\geq 1} \subset L$ is u.b. and, for some $f \in L$, $f_n \xrightarrow{\forall s} f$, then there exists a subsequence $\{f_{n_k}\}_{k\geq 1}$ such that, for all $k, k' \geq 1$, $f_{n_k} \approx f_{n_{k'}}$.*

Recall that Axiom 7 only requires that elements of the sequence $\{f_n\}$ be robustly mixture-neutral with f, for n large. Intuitively, Axiom 10 additionally requires that, as elements of the sequence $\{f_n\}$ get closer and closer in preference at each state, a "cluster" of elements become *mutually* robustly mixture-neutral.

Given a functional representation of preferences, Condition (ii) in Theorem 2.6 is typically harder to verify than Conditions (i) or (iii). The following result (Proposition 6.2 in the Online Appendix) considerably simplifies this task, and is also of independent interest.

Proposition 2.10 *Let $\{C_n\}_{n\geq 1}$ be a proper covering and $\{P_n\}_{n\geq 1}$ the corresponding collection of probabilities. Assume that (i) and (iii) in Theorem 2.6 hold. Finally, let $f, g \in L$ be such that $f \approx g$ and, for some $\lambda, \mu \in [0,1]$ with $\lambda < \mu$, and an appropriate index n, $u(\lambda f + (1-\lambda)g), u(\mu f + (1-\mu)g) \in C_n$.*

Then, for all $\alpha \in [0,1]$, there exists n_α such that $u(\alpha f + (1-\alpha)g) \in C_{n_\alpha}$ and $P_{n_\alpha} = P_n$.

Loosely speaking, if $f \approx g$ and the "line segment" $\{\alpha f + (1-\alpha)g\}$ has a non-trivial intersection with some C_n, then every mixture of f and g is evaluated using the prior P_n—even if it does not belong to C_n. By way of contrast, if Condition (ii) also holds, $f \approx g$ implies $u(\alpha f + (1-\alpha)g) \in C_n$ for all α.

An analogous property holds for MEU preferences: if $f \simeq g$ and preferences are represented by a set $\mathcal{Q}$ of priors on (S, Σ), then there is $Q \in \mathcal{Q}$ such that $Q \in \arg\min_{Q' \in \mathcal{Q}} \int u(\alpha f + (1-\alpha)g)\, dQ'$ for all $\alpha \in [0,1]$. This reconfirms that robust mixture neutrality is the appropriate generalization of mixture neutrality to non-MEU preferences.

2.4.1 Maxmin and Maxmax EU preferences

MEU preferences serve as a model for the main axioms for plausible-priors preferences. Thus, unsurprisingly, a rich (indeed, dense) subset of such preferences satisfies the axioms proposed here. Section 4.1 provides an example of a MEU preference that does not.

Remark 2 Let $\succeq$ be a 0-MEU or 1-MEU preference, and let $\mathcal{P}$ be the corresponding set of priors. If $\mathcal{P}$ is the weak*-closed, convex hull of *finitely* many probability measures $\{P_1, \ldots, P_N\}$, then $\succeq$ satisfies the plausible-priors axioms, and the corresponding set of priors is precisely $\{P_1, \ldots, P_N\}$.

For finite state spaces, I am unaware of examples of MEU preferences employing countably many plausible priors. On the other hand, examples for infinite state spaces are easy to construct (e.g. consider $S = \mathbb{N}$ and the set of (degenerate) priors $\{\delta_{\{n\}} : n \in \mathbb{N}\}$).

2.4.2 Generic α-MEU Preferences

Consider now preferences consistent with Hurwicz' α-criterion. I provide a simple sufficient condition that relies on Proposition 2.10 and a genericity assumption.

Remark 3 Let $\succeq$ be an α-MEU preference, and let $\mathcal{Q}$ be the corresponding set of priors. Assume that $\mathcal{Q}$ is the weak*-closed, convex hull of finitely many probability measures $\{Q_1, \ldots, Q_N\}$, and let $\mathcal{M} \subset \{1, \ldots, M\}^2$ be defined by

$$\mathcal{M} = \left\{(n,m) \;:\; \{Q_n\} = \arg\min_k \int u \circ f \, dQ_k, \; \{Q_m\} = \arg\max_k \int u \circ f \, dQ_k \text{ for some } f \in L\right\}.$$

If, for all distinct pairs $(n,m), (n',m') \in \mathcal{M}$, $\alpha Q_n + (1-\alpha)Q_m \neq \alpha Q_{n'} + (1-\alpha)Q_{m'}$, then $\succeq$ satisfies the plausible-priors axioms, and the corresponding set of priors is precisely $\{\alpha Q_n + (1-\alpha)Q_m : (n,m) \in \mathcal{M}\}$.

A sketch of the argument is as follows. For all $(n,m) \in \mathcal{M}$, let $C_{n,m}$ be the set of functions in $B(S,\Sigma)$ whose expectation is minimized by Q_n and maximized by Q_m. The collection of such sets is a proper covering (see the Online Appendix), and functions in $C_{n,m}$ are evaluated using the measure $P_{n,m} = \alpha Q_n + (1-\alpha)Q_m$. Thus, Condition (i) in Theorem 2.6 holds trivially, and Condition (iii) is easily verified (cf. the argument following Lemma 5.12 in the Appendix). To verify Condition (ii), suppose that $f \approx g$ and let α and (n,m) be such that $u \circ f, u \circ [\alpha f + (1-\alpha)g] \in C_{n,m}$. If $\alpha = 0$, there is nothing to prove; otherwise, Proposition 2.10 implies that there exists $(n',m') \in \mathcal{M}$ such that $P_{n',m'} = P_{n,m}$; but this violates the genericity assumption in Remark 3. Thus, Condition (ii) holds.

2.4.3 Generic CEU preferences in finite state spaces

CEU preferences (Schmeidler [30]) can also satisfy the plausible-priors axioms, provided the state space S is finite. Let $v : 2^S \to [0,1]$ be a capacity on S: that is, $A \subset B \subset S$ imply

$v(A) \le v(B)$, and $v(\emptyset) = 0 = 1 - v(S)$. Assume that $S = \{s_1, \ldots, s_M\}$, and let Π_M be the set of all permutations $(\pi_1, \ldots, \pi_M)$ of $\{1, \ldots, M\}$.

Again, I employ Proposition 2.10 and genericity. In particular, note that the collections of maximal cones of comonotonic functions is a proper covering, and the proofs that Conditions (ii) and (iii) in Theorem 2.6 hold are analogous to the ones sketched above. Thus:

Remark 4 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

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

If, for all $\pi, \pi' \in \Pi_M$, $P_\pi \neq P_{\pi'}$, then $\succeq$ satisfies the axioms proposed here, and $\{P_\pi : \pi \in \Pi_M\}$ is the collection of plausible priors for $\succeq$.

3 Probabilistic Sophistication and Updating

3.1 Probabilistic Sophistication

As is well-known, preferences that admit a non-degenerate "multiple-priors" representation may nevertheless be *probabilistically sophisticated* in the sense of Machina and Schmeidler [23].[31] Whether or not such preferences should be treated as revealing a concern for ambiguity is a somewhat contentious issue; see e.g. e.g. Epstein and Zhang [10], §8.3, and Ghirardato and Marinacci [13], §6.3 for a discussion of alternative viewpoints.

Fortunately, this issue does not arise if the axioms proposed here hold: *a probabilistically sophisticated plausible-priors preference is consistent with SEU.*

To clarify, observe first that, within the Anscombe-Aumann decision framework, Axioms 1–5 imply that preferences over lotteries are consistent with EU maximization. The preceding statement is thus trivially true in such circumstances.[32] This section is concerned with establishing this property in a "fully subjective" setup where lotteries are not available, and hence cannot be employed to pin down the decision-maker's risk preferences.

I adopt the decision setup discussed at the end of §2.1, wherein a characterization of plausible-priors preferences can be provided. Thus, (i) acts are maps from S to a topologically

[31]The "canonical" example is a CEU preference $\succeq$ represented by a capacity ν that is a convex transformation of a probability measure (e.g. $S = [0,1]$ and $\nu(E) = [\lambda(E)]^2$ for all Borel sets E, where λ denotes Lebesgue measure). Since ν is convex, $\succeq$ also admits a MEU representation.

[32]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 [24] for details). In the Anscombe-Aumann setup, Axioms 1–5 imply that V is the EU functional.

rich, but otherwise arbitrary set of prizes X, and (ii) regularity assumptions on preferences guarantee the existence of a *convex-ranged* utility function u on X.

Probabilistic sophistication can be defined as a restriction on preferences in order to economize on notation. An act $f \in L$ is deemed *simple* if $\{x : \exists s \in S,\ f(s) = x\}$ is finite.

Definition 3.1 A preference relation $\succeq$ is *probabilistically sophisticated* (with respect to μ) if 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.[33] Furthermore, the probability μ represents her "qualitative beliefs", as revealed by preferences over binary acts.

Although Def. 3.1 does not require this, the axiomatization of probabilistic sophistication provided by Machina and Schmeidler [23] 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 $\mu(F) = \alpha\mu(E)$. On the other hand, μ need not be a measure; however, for plausible-priors preferences, it will be (due to Axiom 8; cf. also Epstein and Zhang [10]).

Proposition 3.2 *Let $\succeq$ be a plausible-priors preference in the decision setting under consideration. 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.*

Apart from resolving issues related to differences in the definition of "ambiguity", Proposition 3.2 is conceptually relevant to the interpretation of plausible priors. It is *never* the case that a decision-maker who perceives no ambiguity, but has non-EU risk preferences, is (incorrectly) deemed to consider more than one prior "plausible": if her preferences admit a plausible-priors representation as per (2) in Theorem 2.6, then either she is *not* probabilistically sophisticated, or she is a SEU decision-maker.[34]

[33] Def. 3.1 does not explicitly impose "mixture continuity" (cf. [23], pp. 754-755), because it is immaterial to the proof of Proposition 3.2. Also, under Axioms 1–8, Def. 3.1 implies that preferences among non-simple acts are also consistent with FOSD, but this is not used in the proof of Proposition 3.2.

[34] I emphasize that the assumption that μ is convex-ranged is essential for Proposition 3.2 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 [24], which utilizes objective lotteries. As noted above, the result is trivially true under Axioms 1–5 in that setup.

3.2 Updating

The theory presented so far accommodates static decision problems only. This section extends it to dynamic choice problems, adapting analogous axioms and results for MEU preferences in Siniscalchi [32].

Consider an event $E \in \Sigma$; interpret it as information the decision-maker may receive in the dynamic context under consideration.[35] Correspondingly, 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. It shows that, if the unconditional and conditional preference relations satisfy the axioms of the previous section, plus two joint consistency requirements, then:

- the (unique) set of "plausible posteriors" representing the updated preference is related to the set of "plausible priors" via Bayesian updating;
- conditional preferences are determined by a unique updating rule.

As is the case for SEU preferences, updating is well-defined only for a subclass of events. The following definition indicates the relevant restriction. The intuition is that the event E under consideration "matters".

Definition 3.3 An event $E \in \Sigma$ is *non-null* iff, for all pairs of acts $f, g \in L$ such that $f(s) \sim g(s)$ for all $s \in S \setminus E$ and $f(s) \succ g(s)$ for all $s \in E$, $f \succ g$.

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

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

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. This is a version of *consequentialism*.

Axiom 11 (Consequentialism) *For every pair of acts $f, h \in L$: $f \sim_E fEh$.*

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

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 [32]. Loosely speaking, Axiom 12 imposes consistency in situations where hedging considerations are arguably less likely to lead to preference reversals.

Axiom 12 (Dynamic c-Consistency) *For every act $f \in L$ and outcome $y \in Y$:*

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

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

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

The main result follows.

Theorem 3.4 *Consider an event $E \in \Sigma$. Suppose the preferences $\succeq$ and $\succeq_E$ satisfy Axioms 1–5, 7, and 8–9, and assume that E is non-null. Let $\succeq$ be represented by u, $\{C_n : n \in N\}$ and $\{P_n : n \in N\}$ as in Theorem 2.6; similarly, let $\succeq_E$ be represented by u^E, $\{C^E_k : k \in K\}$ and $\{P^E_k : k \in K\}$. Then the following are equivalent:*

(1) $\succeq_E$ satisfies Axiom 11, and $\succeq$, $\succeq_E$ jointly satisfy Axiom 12;

(2) u^E is a positive affine transformation of u, and for every $k \in K$, there exists $n_k \in N$ such that $P^E_k(F) = P_{n_k}(F|E)$ for all $F \in \Sigma$. Moreover, for every $k \in K$ and $a \in C^E_k$,

$$\gamma = \int a \, dP_{n_k}(\cdot|E) \quad \Longrightarrow \quad \forall m \text{ s.t. } 1_E a + 1_{E^c}\gamma \in C_m, \ \int [1_E a + 1_{E^c}\gamma] \, dP_m = \gamma. \tag{3}$$

A few remarks are in order. First, the Theorem ensures that *every* plausible-priors preference relation can be uniquely updated in a manner consistent with Axioms 11 and 12, so as to ensure that the the resulting conditional preference has an analogous "plausible-posteriors" representation. Conceptually, this is perhaps the most important part of Theorem 3.4, because it indicates that *the class of plausible-priors preferences is closed under updating.*

Second, every posterior P^E_k is obtained by updating one of the priors in $\{P_n : n \in N\}$. However, not every plausible prior generates a plausible posterior. Intuitively, this reflects the possibility that, upon acquiring new information, certain ex-ante plausible probabilistic models of the underlying uncertainty might have to be discarded.

Third, the condition in Eq. (3) *characterizes the posterior evaluation of a function* in terms of the prior evaluation of a related function. To clarify, define the"evaluation" of a function $a \in B(S, \Sigma)$ by $I(a) = \int a \, dP_m$ whenever $a \in C_m$; as a consequence of Theorem 2.6, this definition is well-posed. The "posterior evaluation" of a function can be similarly defined. As is shown in the Appendix, Part (2) in Theorem 3.4 implies that the converse implication in Eq. (3) also holds. Thus, the latter equation states that γ is the posterior evaluation of a if and only if γ solves the equation

$$\gamma = I(1_E a + 1_{E^c}\gamma).$$

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 [19], Pires [28] and references therein). On the other hand, Theorem 3.4 shows that Eq. (3) is a *result* of consequentialism and consistency axioms on prior and posterior preferences.

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. This highlights important features of plausible-priors preferences. Notation and assumptions about outcomes are as in the examples of Section 2.

Example 5 (Edith) Let $S = \{s_1, s_2, s_3\}$ and consider a MEU decision-maker, Edith, with priors $\mathcal{Q}_E = \{q \in \Delta(S) : \sum_{i=1,2,3}[q(s_i) - \frac{1}{3}]^2 \leq \varepsilon^2\}$ for $\varepsilon \in (0, \frac{1}{\sqrt{6}}]$. Graphically, $\mathcal{Q}_E$ is a circle of radius ε in the simplex in $\mathbb{R}^3$, centered at the uniform distribution on S.

Two acts f, g satisfy $f \simeq g$ if and only if they are *affinely related*, i.e. if and only if $f(s) = \alpha g(s) + \beta$ for some $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$.[36] Now let C be any maximal collection of affinely related acts. Note that there is a unique prior $q_C \in \mathcal{Q}_E$ that minimizes $\int f \, dq$ over $\mathcal{Q}_E$ 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

[36] 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 $u(y) = y$.

related. Thus, there exists a one-parameter family (hence, a continuum) of probabilities that represent preferences on C.[37] Consequently, no plausible prior can be elicited.

Since in this example $f \simeq g$ (if and) only if f and g are affinely related, it is easy to see that only constant acts are robustly mixture neutral. Hence, Axiom 7 is violated in a relatively trivial sense. However, the preferences in Example 5 also violate a variant of Axiom 7 wherein robust mixture neutrality is replaced with simple mixture neutrality.[38] Since the state space is finite, the discussion preceding Axiom 7 strongly suggests that considerations other than variability reduction, or hedging, determine Edith's violations of mixture neutrality. Moreover, Edith's preferences satisfy a strict version of Schmeidler's "Uncertainy Aversion" axiom [30, p. 582]. In this respect, Edith behaves very differently from a SEU decision-maker when choosing from a sufficiently rich (although possibly "small") set of acts: she is *not* a "local" SEU maximizer.

Note that the preferences in Example 5, as well as arbitrary MEU preferences, can be approximated arbitrarily closely by MEU preferences that do satisfy Axiom 7. This indicates that, loosely speaking, plausible-priors preferences are "dense" in the class of MEU preferences, so that the additional restrictions they are required to satisfy do not exact a significant price in terms of expressive power.

As the approximation becomes more accurate, it might appear that the behavioral distinction between preferences that permit the elicitation of plausible priors and those that don't "becomes small"—and hence, so does the distinction between plausible priors and elements of sets such as $\mathcal{Q}_E$ in Example 5. But, intuitively and formally, the proposed notion of plausibility is binary: either a prior provides a unique SEU representation of preferences on some set of acts, or it does not. On the other hand, to formalize the above "continuity" intuition, a behaviorally founded notion of *degrees* of plausibility is required. Similar issues arise in connection with ambiguity: for instance, refer to Example 1, the Ellsberg Paradox, and consider the set of priors $\mathcal{Q}^{\epsilon}_A = \{q : q(r) = \frac{1}{3}, q(g) \in [\frac{1}{3} - \epsilon, \frac{1}{3} + \epsilon]\}$. For all $\epsilon > 0$, however "small", the events $\{g\}$ and $\{b\}$ are ambiguous, and the decision-maker is deemed ambiguity averse; but $\epsilon = 0$ corresponds to SEU preferences.

I emphasize that a preference that does not admit plausible priors need not be considered "defective". As discussed in the Introduction, the intuitive interpretation of multiple

[37] I.e., the set $Q_C = \{q \in \mathbb{R}^3_+ : \sum_i q(s_i) = 1, \sum_i f_i q(s_i) = \sum_i f_i q_C(s_i)\}$, where $f \in C$ is non-constant. Geometrically, Q_C is a line segment in the simplex in $\mathbb{R}^3$ tangent to $\mathcal{Q}_E$ at q_C. Due to the shape of $\mathcal{Q}_E$, the set Q_C is non-degenerate (i.e. not a singleton) for any maximal collection C of affinely related acts.

[38] It can be shown that such a version of Axiom 7 characterizes MEU preferences that admit plausible priors

priors may or may not apply to such preferences; however, alternative, behaviorally-based interpretations may be possible. For instance, Wang [33] provides an axiomatization of an entropy-based multiple-priors model. Another alternative is considered in Siniscalchi [31]. 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 [12, GMM henceforth] and Nehring [27] 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 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 $f, g \in L$,

$$f \succeq^* g \quad \Leftrightarrow \quad \forall Q \in \mathcal{Q}, \int u \circ f \, dQ \geq \int u \circ g \, dQ. \tag{4}$$

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. (4). 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. (4); but, just like a MEU preference, a Bewley preference admits alternative representations, characterized by different sets of priors.[39]

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

[39]Section 6.1 in the Online Appendix discusses Bewley preferences and provides examples.

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 Axioms 7 and 9 do not hold, the intuitive interpretation of the elements of $\mathcal{Q}$ as possible probabilistic models of the underlying uncertainty is problematic.

Also note that a probabilistically sophisticated preference may give rise to a non-singleton set $\mathcal{Q}$ in the GMM setup. By Proposition 3.2, this is never the case if Axioms 7 and 9 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 representation analogous to Eq. (1); 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 [25] 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." [25, 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

Overview. Lemma 5.1 provides a basic representation for the preference $\succeq$. Henceforth, the analysis can be carried out with reference to this representation. Lemma 5.2 then furnishes basic properties of the relations $\simeq$ and $\approx$. These are employed in Lemma 5.3 to translate Axioms 7–9 into properties of the functional representation of preferences. Building on these properties, Lemma 5.4 derives further key properties of $\simeq$ and $\approx$. All proofs 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 $u : Y \to \mathbb{R}$ 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.

Denote by $B_1(S,\Sigma)$ the unit ball of $B(S,\Sigma)$. With an abuse of notation, I write $a \simeq b$ for functions $a, b \in B(S,\Sigma)$ as a shorthand for "$I(\alpha a + (1-\alpha)b) = \alpha I(a) + (1-\alpha)I(b)$ for all $\alpha \in [0,1]$". Similarly, I write $a \approx b$ iff, for every $c \in B(S,\Sigma)$ and $c' \in \{\lambda a + (1-\lambda)b : \lambda \in (0,1)\}$, there exists $\gamma \in (0,1)$ such that $\gamma c + (1-\gamma)c' \simeq a$ and $\gamma c + (1-\gamma)c' \simeq b$.

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

1. *For all $f, g \in L$, $f \simeq g$ iff $u \circ f \simeq u \circ g$.*
2. *For all $a, b \in B(S,\Sigma)$, and $\alpha, \beta \in \mathbb{R}$ with $\alpha \geq 0$: $a \simeq b$ implies $a \simeq \alpha b + \beta$.*
3. *For all $a, b \in B(S,\Sigma)$: $a \approx b$ iff, for all $c \in B_1(S,\Sigma)$ and $c' \in \{\lambda a+(1-\lambda)b : \lambda \in (0,1)\}$, there exists $\gamma \in (0,1)$ such that $\gamma c + (1-\gamma)c' \simeq a$ and $\gamma c + (1-\gamma)c' \simeq b$. That is, only functions $c \in B_1(S,\Sigma)$ need be considered in the definition of $\approx$ for functions.*
4. *For all $f, g \in L$, $f \approx g$ iff $u \circ f \approx u \circ g$.*
5. *For all $a, b \in B(S,\Sigma)$: $a \approx b$ implies $a \simeq b$.*
6. *For all $a, b \in B(S,\Sigma)$, and $\alpha, \beta \in \mathbb{R}$ with $\alpha > 0$: $a \approx b$ implies $a \approx \alpha b + \beta$.*

Lemma 5.3 *Suppose $\succeq$ satisfies Axioms 1, 2, 3, 4 and 5, and let I, u be its representation.*

1. *$\succeq$ satisfies Axiom 7 iff, for any sequence $\{a_n\}_{n\geq 1} \subset B_1(S,\Sigma)$ such that $a_n(s) \to a(s)$ for all $s \in S$ and some $a \in B_1(S,\Sigma)$, there is $N \geq 1$ such that $n \geq N$ implies $a_n \approx a$.*
2. *$\succeq$ satisfies Axiom 8 iff, for any sequence $\{a_n\}_{n\geq 1} \subset B_1(S,\Sigma)$ such that $a_n(s) \to a(s)$ for all $s \in S$, $I(a_n) \to I(a)$.*
3. *$\succeq$ satisfies Axiom 9 iff, for all $a, b, c \in B_1(S,\Sigma)$, $a \approx b$ and $c \approx \gamma b + (1-\gamma)c$ for some $\gamma \in (0,1)$ imply $\alpha a + (1-\alpha)c \approx \beta b + (1-\beta)c$ for all $\alpha, \beta \in [0,1]$.*

Henceforth, the analysis will focus on the properties and representation of the functional I on $B_1(S,\Sigma)$. In order to streamline the exposition, expressions such as "by Axiom 8 and Part (2) of Lemma 5.3, $I(a_n) \to I(a)$" will be shortened to "by Axiom 8, $I(a_n) \to I(a)$." That is, references to Axioms 7, 8, or 9 should be interpreted as references to the respective equivalent conditions on I provided in Lemma 5.3.

I remark that Axiom 9 will be most often invoked with $\alpha = 0$, $\beta = 1$ or $\alpha = 1$,$\beta = 1$: that is, $c \approx \gamma a + (1-\gamma)b$ implies $c \approx a$ and $c \approx b$. However, the full force of the Axiom will be needed in the proof of Lemma 5.6, Part (3).

Lemma 5.4 *Suppose $\succeq$ satisfies Axioms 1, 2, 3, 4, 5, 8, 7, 9. Let I, u be its representation.*

1. *If $\{a_n\}, \{b_n\} \subset B_1(S,\Sigma)$, $a_n(s) \to a(s)$, $b_n(s) \to b(s)$ for all s, and $a_n \simeq b_n$ for all n, then $a \simeq b$.*
2. *For every $a \in B_1(S,\Sigma)$, $a \approx a$.*
3. *For every $a \in B_1(S,\Sigma)$ and $\beta \in \mathbb{R}$, $a \approx \beta$ (thus, $a \approx b$ implies $a \approx \alpha b + \beta$ for $\alpha \geq 0$ and $\beta \in \mathbb{R}$: cf. Lemma 5.2, Part 6).*
4. *If $a \approx b$, then for all $\{c_n\} \subset B_1(S,\Sigma)$ such that, for some $c \in \{\lambda a + (1-\lambda)b : \lambda \in (0,1)\}$, $c_n(s) \to c(s)$ for all s, there exists $N \geq 1$ such that $n \geq N$ implies $c_n \approx a, b$, hence $c_n \simeq a, b$.*
5. *If $a \approx b$ and $\alpha, \beta \in [0,1]$, then $\alpha a + (1-\alpha)b \approx \beta a + (1-\beta)b$.*
6. *If $\{a_n\}, \{b_n\}, \{a, b\} \subset B_1(S,\Sigma)$ are such that $a_n(s) \to a(s)$, $b_n(s) \to b(s)$ for all s, and $a_n \approx b_n$ for each n, then $a \approx b$.*

5.1.2 Necessity of the Axioms

It can now be shown that (2) implies (1) in Theorem 2.6. The following, preliminary result collects useful properties of proper coverings; in particular, (1) below is related to (ii) in Theorem 2.6.

Lemma 5.5 *Consider a proper covering $\mathcal{C}$. Then:*

(1) Suppose that $a \approx b$ implies $a, b \in C$ for some $C \in \mathcal{C}$. Then every convex set D having non-empty algebraic interior[40] *and such that $a, b \in D$ implies $a \approx b$ is included in some $C \in \mathcal{C}$.*

(2) If $C, C' \in \mathcal{C}$, then $C \cap C'$ is a subset of the boundary of C (and C'). Consequently, the intersection of any two $C, C' \in \mathcal{C}$ has empty interior, and a point in the interior of some

[40] That is, there exists $d \in D$ such that, for all $a \in D$, there exists $b \in D$ such that $\frac{1}{2}a + \frac{1}{2}b = d$.

$C \in \mathcal{C}$ belongs to no other $C' \in \mathcal{C}$. Thus, if $a \approx b$ implies $a, b \in C'$ for some $C' \in \mathcal{C}$, c is in the interior of $C \in \mathcal{C}$, and $d \approx c$, then $d \in C$.

Proof.

(1) Let D have the required properties and consider $d \in D$ in the algebraic interior of D. Since $\mathcal{C}$ is a covering, $d \in C$ for some $C \in \mathcal{C}$. Now consider an arbitrary $a \in D$: then there exists $b \in D$ such that $\frac{1}{2}a + \frac{1}{2}b = d$. By assumption, $a, b \in C'$ for some (possibly different) $C' \in \mathcal{C}$; since C' is convex, $d \in C'$. Thus, one has $d \in C \cap C'$ and $a, b \in C'$: then, by Part (2) in Def. 2.5, $a, b \in C$ as well. Since a was arbitrary, $D \subset C$.

(2) Suppose there exists $c \in C \cap C'$ such that c is an interior point of C. The (non-empty) topological interior of C coincides with its algebraic interior (cf. e.g. Holmes [18], §11.A), so for every $a \in C$ there exists $b \in C$ such that $\frac{1}{2}a + \frac{1}{2}b = c$. By Part (2) in Def. 2.5, $a, b \in C'$: that is, $C \subset C'$, which contradicts the requirement that no two elements of $\mathcal{C}$ be nested. Thus, c must lie in the boundary of C; similarly, it must lie in the boundary of C'. ∎

Now let u, $\mathcal{C}$ and $\{P_n\}$ be as in Theorem 2.6, (2). 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 a cone, 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, I is also unique.

Let $a \in B(S, \Sigma)$, $\beta \in \mathbb{R}_+$ and $\gamma \in \mathbb{R}$. Since each C_n is a convex cone that contains all constant functions, $\beta a + \gamma \in C_n$ 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.e. I is c-linear.

Let $a, b \in B(S, \Sigma)$ be such that $a(s) \geq b(s)$ for all s; then $\int a \, dP_n \geq \int b \, dP_n$ for all $n \geq 1$. Since $u(Y) \supset [-1, 1]$, there exist $f, g \in L$ and $\alpha > 0$ such that $u \circ f = \alpha a$ and $u \circ g = \alpha b$. Then (iii) implies that $f \succeq g$; since (I, u) represent $\succeq$ and I is positively homogeneous, this is equivalent to $I(a) = \frac{1}{\alpha} I(u \circ f) \geq \frac{1}{\alpha} I(u \circ g) = I(b)$, i.e. I is monotonic.

Now Lemma 5.1 implies that I satisfies Axioms 1, 2, 3, 4, and 5. To see that Axioms 8 and 7 hold, consider $a \in B_1(S, \Sigma)$ and a bounded sequence $\{a_k\}$ in $B_1(S, \Sigma)$ such that $a_k(s) \to a(s)$ for all s.

If a lies in the interior of some $C_m \in \mathcal{C}$, so that, by Part (2) of Lemma 5.5, it belongs to no other element of $\mathcal{C}$, then Part (3) in Def. 2.5 implies that there exists $K \geq 1$ such that, for $k \geq K$, $a_k, a \in C_m$. Then, since P_m is countably additive and the sequence $\{a_k\}$ is bounded, by the Dominated Convergence theorem, $\int a_k \, dP_m \to \int a \, dP_m = I(a)$; and since $I(a_k) = \int a_k \, dP_m$ for $k \geq K$, $I(a_k) \to I(a)$, as required.

Now assume instead that a lies on the boundary of C_m. Let $\epsilon > 0$. Then there exists c^ϵ in the interior of C_m such that $\|a - c^\epsilon\| < \frac{\epsilon}{3}$. Furthermore, the sequence $\{a_k - [a - c^\epsilon]\}_{k \geq 1}$

is bounded and converges pointwise to c^ϵ; hence, by the argument just given, there exists K such that $k \geq K$ implies $|I(c^\epsilon) - I(a_k - [a - c^\epsilon])| < \frac{\epsilon}{3}$. We then have, for $k \geq K$,

$$\begin{aligned} |I(a) - I(a_k)| &\leq |I(a) - I(c^\epsilon)| + |I(c^\epsilon) - I(a_k - [a - c^\epsilon])| + |I(a_k - [a - c^\epsilon]) - I(a_k)| < \\ &< \|a - c^\epsilon\| + \frac{\epsilon}{3} + \|a_k - a + c^\epsilon - a_k\| = 2\|a - c^\epsilon\| + \frac{\epsilon}{3} < \epsilon. \end{aligned}$$

Thus, again $I(a_k) \to I(a)$. Part (1) of Lemma 5.3 now implies that Axiom 8 holds.

Next, I claim that $a, b \in C_n \in \mathcal{C}$ implies $a \approx b$. Fix $c \in B(S, \Sigma)$ and $c' = \lambda a + (1-\lambda) b \in C_n$ for some $\lambda \in (0,1)$. Let $c_k = \frac{1}{k} c + \frac{1-k}{k} c'$, so $c_k \to c'$ uniformly. By Part (3) in Def. 2.5, for k large, there exists C_{n_k} such that $c_k, c' \in C_{n_k}$. Then $c' \in C_{n_k} \cap C_n$, which, by Part (2) in Def. 2.5, implies that also $a, b \in C_{n_k}$. Thus, for such k, and for every $\alpha \in [0,1]$, $I(\alpha a + (1 - \alpha) c_k) = \int [\alpha a + (1-\alpha) c_k] \, dP_{n_k} = \alpha \int a \, dP_{n_k} + (1-\alpha) \int c_k \, dP_{n_k} = \alpha I(a) + (1-\alpha) I(c_k)$, i.e. $\frac{1}{k} c + \frac{k-1}{k} c' = c_k \simeq a$; similarly, $\frac{1}{k} c + \frac{k-1}{k} c' \simeq b$. Thus, $a \approx b$.

Now, to see that Axiom 7 holds, consider $\{a_k\} \subset B_1(S, \Sigma)$ u.b. and such that $a_k(s) \to a(s) \in B_1(S, \Sigma)$ for all s. By Part (3) in Def. 2.5, for k large, there exists C_{n_k} such that $a_k, a \in C_{n_k}$; as was just shown, this implies that $a_k \approx a$ for such k. Part (2) of Lemma 5.3 then implies that Axiom 7 holds.

Finally, consider $a, b, c \in B_1(S, \Sigma)$ such that $a \approx b$ and $c \approx \gamma a + (1-\gamma) b$ for some $\gamma \in (0,1)$. By (ii) in Theorem 2.6, there exists $n, m \geq 1$ such that $a, b \in C_n$ and $c, \gamma a + (1-\gamma) b \in C_m$. Thus, $\gamma a + (1-\gamma) b \in C_n \cap C_m$. By Part (2) in Def. 2.5, $a, b \in C_m$; hence, for all $\alpha, \beta \in [0,1]$, $\alpha a + (1-\alpha) c, \beta b + (1-\beta) c \in C_m$ as well. As was shown above, this implies that $\alpha a + (1-\alpha) c \approx \beta b + (1-\beta) c$. Thus, by Part (3) of Lemma 5.3, Axiom 9 holds, and the proof of necessity is complete.

5.1.3 Sufficiency: Covering $C(S) \cap B_1(S, \Sigma)$.

Turn now to the other direction. Assume throughout that Axioms 1-8, 7 and 9 hold for the preference relation $\succeq$. Continue to denote by I, u its representation per Lemma 5.1.

Recall that (S, Σ) is assumed to be a standard Borel space: that is, there exists a Polish topology $\tau \subset 2^S$ such that Σ is the Borel σ-algebra generated by τ. However, by the Borel Isomorphism theorem (Kechris [20], Theorem 15.6), it is sufficient to prove the characterization result for S compact metrizable, with Σ its Borel σ-algebra [See the Online Appendix for additional details.]

This subsection and the next contain the heart of the proof that (1) $\Rightarrow$ (2) in Theorem 2.6. The objective is to construct a proper covering of $B(S, \Sigma)$; since I is c-linear, it is actually sufficient to construct a suitable covering of $B_1(S, \Sigma)$, the unit ball of $B(S, \Sigma)$.

Overview. Lemma 5.6 establishes the existence of a covering of certain convex subsets of $B_1(S,\Sigma)$ whose elements are maximal convex sets containing all constant functions, characterized by the property that, if a, b are functions belonging to the same subset, then $a \approx b$. Lemma 5.7 then shows that the covering $\mathcal{C}_0$ of the set $C(S) \cap B_1(S,\Sigma)$ of *continuous* functions with norm at most 1 delivered by Lemma 5.6 has a special structure (and in fact extends to a proper covering of $C(S)$, although this is not essential to the argument). Finally, Lemmata 5.8 and 5.9 in the next subsection exploit the structure of $\mathcal{C}_0$ to construct a covering of $B_1(S,\Sigma)$ that extends (by c-linearity of I) to a proper covering of $B(S,\Sigma)$.

Lemma 5.6 *Let M be a convex subset of $B_1(S,\Sigma)$ that contains all constant functions $\gamma 1_S$, $\gamma \in [-1,1]$. There exists a unique, non-empty collection $\mathcal{C}$ of convex subsets of M such that:*

(1) For all $C \in \mathcal{C}$, $\gamma 1_S \in C$ for all $\gamma \in [-1,1]$;

(2) For all $C \in \mathcal{C}$ and $a, b \in C$, $a \approx b$;

(3) For all $C \in \mathcal{C}$, if $a \in M$ satisfies $a \approx c$ for all $c \in C$, then $a \in C$, i.e. C is $\subset$-maximal with respect to the relation $\approx$. In particular, for every $C, D \in \mathcal{D}$ and $a \in C \backslash D$ there exists $b \in D$ such that $a \not\approx b$, so the elements of $\mathcal{C}$ are not nested.

(4) For all $a \in M$, there exists $C \in \mathcal{C}$ such that $a \in C$. More generally, every convex subset D of M that satisfies (1) and (2) is contained in some $C \in \mathcal{C}$.

Finally, if M is norm-closed, then so are the elements of $\mathcal{C}$.

Note that, in general, elements of the covering constructed in this Lemma may have empty interior. Note also that, although attention is restricted to elements of a given subset of $B_1(S,\Sigma)$, in order to determine whether $a \approx b$, one still needs to consider *all* functions $c \in B_1(S,\Sigma)$ and verify that, for all $c' \in \{\lambda a + (1-\lambda)b : \lambda \in (0,1)\}$, there is $\gamma \in (0,1)$ such that $\gamma c + (1-\gamma)c' \simeq a, b$. That is, the "test set" of perturbations is all of $B_1(S,\Sigma)$.

Proof. Let $\mathcal{C}'$ be the collection of all convex subsets of M satisfying properties (1) and (2) above. In particular, for every $a \in M$, $\{\alpha a + (1-\alpha)\gamma 1_S \ : \ \alpha \in [0,1], \gamma \in [-1,1]\} \in \mathcal{C}'$, because M is a convex subset of $B_1(S,\Sigma)$ that includes $\gamma 1_S$ for all $\gamma \in [-1,1]$, and $\alpha a + (1-\alpha)\gamma \approx \alpha' a + (1-\alpha')\gamma'$ for all appropriate $\alpha, \alpha', \gamma, \gamma'$, by Lemma 5.4, Parts 2 and 3. Partially order $\mathcal{C}'$ by set inclusion ($\subset$). If $\mathcal{C}'' \subset \mathcal{C}'$ is a chain, consider $C = \bigcup_{C' \in \mathcal{C}''} C'$; then C satisfies (1), and, furthermore, if $a, b \in C$, then $a, b \in C'$ for some $C' \in \mathcal{C}''$, which ensures that C is a convex set that satisfies (2) as well. Hence, $C \in \mathcal{C}'$.

Now let $\mathcal{C}$ be the set of all maximal elements of $(\mathcal{C}', \subset)$; Zorn's Lemma ensures that $\mathcal{C}$ is non-empty. Every element $C \in \mathcal{C}$ is a convex set that satisfies (1) and (2). The norm-closure $\bar{C}$ of any $C \in \mathcal{C}$ is also convex and satisfies (1); furthermore, it satisfies (2), because, for any pair of sequences $\{a^n\}, \{b^n\} \subset C$ such that $a^n \to a$, $b^n \to b$ for $a, b \in \bar{C}$, $a_n \approx b_n$ for all n

implies that $a \approx b$ by Lemma 5.4, Part 6. Therefore, if M is norm-closed, every maximal element C of $(\mathcal{C}', \subset)$ must be norm-closed.

For (3), consider $D \in \mathcal{C}$, so D satisfies (1) and (2). Suppose that, for some $c \in M \setminus D$, $c \approx d$ for all $d \in D$. The set $D^{+c} = \{\alpha c + (1-\alpha)d \ : \ d \in D, \ \alpha \in [0,1]\}$ is a convex subset of M that properly contains D, and hence satisfies (1). To see that it also satisfies (2), consider $\alpha c + (1-\alpha)d, \alpha' c + (1-\alpha')d' \in D^{+c}$, so $d, d' \in D$ and $\alpha, \alpha' \in [0,1]$. Since $D \in \mathcal{C}$, $d \approx d'$ and $\frac{1}{2}d + \frac{1}{2}d' \in D$; furthermore, by assumption, $c \approx \frac{1}{2}d + \frac{1}{2}d'$. Thus, Axiom 9 implies that also $\alpha c + (1-\alpha)d \approx \alpha' c + (1-\alpha')d'$. Thus, $D \subsetneq D^{+c} \in \mathcal{C}'$, a contradiction. In particular, if $c \in C \setminus D$, then $c \in M$, so there must be $d \in D$ such that $c \not\approx d$.

For (4), consider any convex set $D \in \mathcal{C}'$ and let $\mathcal{C}'_D \subset \mathcal{C}'$ be the collection of convex sets $C' \in \mathcal{C}'$ such that $D \subset C'$. Order $\mathcal{C}'_D$ by set inclusion, and argue as above to conclude that $\mathcal{C}'_D$ has at least one maximal element, say $\bar{D}$. Then $\bar{D}$ must also be a maximal element of $\mathcal{C}'$: suppose that $\bar{D} \subset C'$ for some $C' \in \mathcal{C}'$. Since $\bar{D} \in \mathcal{C}'_D$, $D \subset \bar{D} \subset C'$, so $C' \in \mathcal{C}'_D$; but since $\bar{D}$ is maximal in $\mathcal{C}'_D$, $C' \subset \bar{D}$ must hold, i.e. $\bar{D}$ is also maximal in $\mathcal{C}'$. Hence, every $D \in \mathcal{C}'$ is contained in some $C \in \mathcal{C}$; in particular, by the argument given above, every $a \in M$ is contained in some $C \in \mathcal{C}$.

Finally, $\mathcal{C}$ is the only collection of (closed), convex sets for which (1)-(4) hold. To see this, consider another collection $\mathcal{D}$ having the same properties. Fix $D \in \mathcal{D}$; then D is a convex set that satisfies (1) and (2), and therefore it is contained in some $C \in \mathcal{C}$. Moreover, if $D \neq C$, then there exists $a \in C \setminus D$ such that $a \approx b$ for all $b \in D \subset C$; by (4), $a \in D' \in \mathcal{D}$, and indeed $a \in D' \setminus D$; hence, (3) is violated. Thus, $D = C$. Therefore, $\mathcal{D} \subset \mathcal{C}$; the same argument shows that $\mathcal{C} \subset \mathcal{D}$, so $\mathcal{D} = \mathcal{C}$. ■

Now let $\mathcal{C}_0$ be the covering provided by Lemma 5.6 for $M = C(S) \cap B_1(S, \Sigma)$, the unit ball of the set of continuous functions on S.

Lemma 5.7 *For every $C \in \mathcal{C}_0$, there exists $c \in C$ such that $a \in B(S, \Sigma)$ and $a \approx c$ imply $a \approx b$ for all $b \in C$. In particular, $a \in C(S) \cap B_1(S, \Sigma)$ and $a \approx b$ imply $a \in C$.*

A function $c \in C$ with the properties mentioned in the above statement will be henceforth referred to as a *critical point*.

Proof. Recall that $C(S)$, endowed with its relative norm topology, is a separable (cf e.g. [2], Theorem 7.47) metric space. Thus, every $C \in \mathcal{C}_0$ is also separable ([2], Corollary 3.2). Furthermore, since $C(S) \cap B_1(S, \Sigma)$ is norm-closed, Lemma 5.6 also ensures that every $C \in \mathcal{C}_0$ is closed as well. This implies that, if $\{b_n\}$ is a sequence in $C \in \mathcal{C}_0$ and $\{\beta_n\}_{n \geq 1}$ is a sequence in $(0,1)$ such that $\sum_n \beta_n \leq 1$, the series $\sum_n \beta_n b_n$ converges in C.

Now fix $C \in \mathcal{C}_0$ and let $\{c_1, c_2, ...,\}$ be a countable dense subset of C. Fix a collection $\{\alpha_n\}_{n\geq 1} \in (0,1)^{\mathbb{N}}$ such that $\sum_n \alpha_n = 1$, and define $c = \sum_n \alpha_n c_n$.

I claim that, if $a \in B(S,\Sigma)$ satisfies $a \approx c$, then $a \approx c_n$ for all $n \geq 1$. To see this, note first that, for all $n \geq 1$,

$$c = \sum_{k\geq 1} \alpha_k c_k = \alpha_n c_n + (1-\alpha_n) \sum_{k\in\mathbb{N}\setminus\{n\}} \frac{\alpha_k}{1-\alpha_n} c_k.$$

Since $\sum_{k\in\mathbb{N}\setminus\{n\}} \frac{\alpha_k}{1-\alpha_n} c_k \in C$ and $\alpha_n \in (0,1)$, Axiom 9 implies the claim. Now fix $b \in C$; I claim that $a \approx c$ implies $a \approx b$. To see this, note that there is a sequence $\{c_{n_k}\}_{k\geq 1} \subset \{c_n\}_{n\geq 1}$ such that $c_{n_k} \to b$ in norm. Since $a \approx c_{n_k}$ for all k, Lemma 5.4, Part 6 implies that $a \approx b$, as needed. In particular, by Lemma 5.6, Part (3), $a \approx c$ for some $a \in C(S) \cap B_1(S,\Sigma)$ implies $a \in C$, as required. ■

5.1.4 Sufficiency: Proper Covering of $B(S,\Sigma)$

In order to extend $\mathcal{C}_0$ to a covering of $B_1(S,\Sigma)$, consider the Baire hierarchy of functions from S to $[-1,1]$ (Kechris [20], §24).

Let $\mathcal{B}_1$ contain all pointwise limits of sequences of continuous functions from S to $[-1,1]$: that is, $b \in \mathcal{B}_1$ iff there exists a sequence $\{a_n\}_{n\geq 1}$ in $C(S) \cap B_1(S,\Sigma)$ such that $a_n(s) \to b(s)$ for all $s \in S$. Recursively, for every ordinal ξ such that $1 < \xi < \omega_1$ (the latter symbol denotes the first uncountable ordinal), consider the set $\mathcal{B}_\xi$ of functions $b : S \to [-1,1]$ for which one can find a sequence $\{a_n\}$ of functions such that (i) for every $n \geq 1$, $a_n \in \mathcal{B}_{\xi_n}$ for some $\xi_n < \xi$, and (ii) $a_n(s) \to b(s)$ for all $s \in S$.

Note that, trivially, all limits above involve uniformly bounded sequences of functions. Furthermore, $C(S) \cap B_1(S,\Sigma) \subset \mathcal{B}_1 \subset \ldots \subset \mathcal{B}_\xi \subset \ldots \subset \mathcal{B}_\eta \subset ...$, for any $\xi \leq \eta < \omega_1$.

By Theorem 24.10 and Exercise 24.13 in Kechris [20], $\mathcal{B}_1$ is the set of functions of Baire class 1. Consequently, by Theorem 24.3 in [20], $\bigcup_\xi \mathcal{B}_\xi$ is the class of Borel-measurable functions from S to $[-1,1]$, i.e. $B_1(S,\Sigma)$. Furthermore, each member $\mathcal{B}_\xi$ of the Baire hierarchy is easily seen to be a convex[41] subset of $B_1(S,\Sigma)$ that contains all constant functions $\gamma 1_S$ for $\gamma \in [0,1]$.

For notational convenience, let $\mathcal{B}_0 = C(S) \cap B_1(S,\Sigma)$; also let $\varphi_0 : \mathcal{C}_0 \to \mathcal{C}_0$ be the identity function. Let $\mathcal{C}_\xi$ be a covering of $\mathcal{B}_\xi$ by convex sets, as per Lemma 5.6.

Lemma 5.8 *For every ordinal $\xi < \omega_1$, there exists a one-to-one map $\varphi_\xi : \mathcal{C}_0 \to \mathcal{C}_\xi$ such that, for every $C \in \mathcal{C}_0$:*

[41] See the argument for the convexity of $\varphi_\xi(C)$ in the proof of Lemma 5.8.

(i) if $\eta < \xi$, then $\varphi_\eta(C) \subset \varphi_\xi(C)$;

(ii) if $\xi > 0$, then $a \in \varphi_\xi(C)$ iff there exists a sequence b_n that converges pointwise to a and such that, for every $n \geq 1$, there is an ordinal $\xi_n < \xi$ such that $b_n \in \varphi_{\xi_n}(C)$;

(iii) if c is a critical point of C, then $a \in B(S, \Sigma)$ and $a \approx c$ imply $a \approx b$ for all $b \in \varphi_\xi(C)$—so, in particular $a \in \varphi_\xi(C)$ if $a \in \mathcal{B}_\xi$;

(iv) the collection $\{\varphi_\xi(C) : C \in \mathcal{C}_0\}$ covers $\mathcal{B}_\xi$.

Proof. The claim is true for $\xi = 0$: (i) and (ii) hold vacuously, (iii) follows directly from Lemma 5.7, and (iv) holds because φ_0 is the identity function on $\mathcal{C}_0$. For $\xi > 0$, consider $C \in \mathcal{C}_0$ and let $\varphi_\xi(C)$ be the set of functions $a \in \mathcal{B}_\xi$ that are pointwise limits of sequences $\{b_n\}$ in $B_1(S, \Sigma)$ such that, for every $n \geq 1$, $b_n \in \varphi_{\xi_n}(C)$ for some $\xi_n < \xi$.

I claim that $\varphi_\xi(C)$ is a convex subset of $\mathcal{B}_\xi$ that satisfies conditions (1) and (2) of Lemma 5.6. To see this, note first that $\varphi_\xi(C)$ contains C (consider constant sequences from $C = \varphi_0(C)$). Hence, in particular it contains all constant functions $\gamma 1_S$, for $\gamma \in [-1, 1]$. Next, let $a, b \in \varphi_\xi(C)$, so there exist sequences $\{a_n\}$ and $\{b_n\}$ converging pointwise to a and b respectively, and such that, for every $n \geq 1$, $a_n \in \varphi_{\xi_n^a}(C)$ and $b_n \in \varphi_{\xi_n^b}(C)$ with $\xi_n^a, \xi_n^b < \xi$. By Part (i) of the induction hypothesis, for every $n \geq 1$, if $\xi_n^a < \xi_n^b$, then $\varphi_{\xi_n^a}(C) \subset \varphi_{\xi_n^b}(C)$; otherwise, the reverse inclusion holds. Moreover, again by the induction hypothesis, for every $n \geq 1$, $\varphi_{\max(\xi_n^a, \xi_n^b)}(C)$ is an element of $\mathcal{C}_{\max(\xi_n^a, \xi_n^b)}$, hence a (maximal) convex set that satisfies (1) and (2) of Lemma 5.6. Thus, for every $\alpha \in [0, 1]$, and for every $n \geq 1$, $\alpha a_n + (1-\alpha) b_n \in \varphi_{\max(\xi_n^a, \xi_n^b)}(C)$, with $\max(\xi_n^a, \xi_n^b) < \xi$; furthermore, $a_n \approx b_n$. Finally, the sequence $\{\alpha a_n + (1-\alpha) b_n\}$ converges to $\alpha a + (1-\alpha) b$ pointwise. Hence, the latter is a member of $\varphi_\xi(C)$, and by Lemma 5.4, Part 6, $a \approx b$. Thus, $\varphi_\xi(C)$ satisfies conditions (1) and (2) of Lemma 5.6 for the subset $\mathcal{B}_\xi$ of $B_1(S, \Sigma)$. Therefore, it is included in at least one element of $\mathcal{C}_\xi$. Let C' be one such element. It will now be shown that $C' \subset \varphi_\xi(C)$, so in fact $C' = \varphi_\xi(C)$.

Fix $a \in C'$ and let $c \in C$ be a critical point of C. Then $c \in C'$, $a \approx c$ and, for all $\alpha \in (0, 1)$, $\alpha c + (1-\alpha) a \in C'$. Since $a \in \mathcal{B}_\xi$, there exists a sequence b_n converging pointwise to a such that $b_n \in \mathcal{B}_{\xi_n}$ and $\xi_n < \xi$ for each n. Thus, for every $\alpha \in (0, 1)$, $\alpha c + (1-\alpha) b_n$ converges pointwise to $\alpha c + (1-\alpha) a$, and since $c \in \mathcal{B}_0$ and Baire classes are increasing and convex, $\alpha c + (1-\alpha) b_n \in \mathcal{B}_{\xi_n}$.

Now fix one such $\alpha \in (0, 1)$. By Axiom 7, there exists $N(\alpha) \geq 1$ such that $\alpha c + (1-\alpha) a \approx \alpha c + (1-\alpha) b_n$ for all $n \geq N(\alpha)$; thus, by Axiom 9, $\alpha c + (1-\alpha) b_n \approx c$ for such n. Since $\alpha c + (1-\alpha) b_n \in \mathcal{B}_{\xi_n}$, Part (iii) of the the induction hypothesis implies that, for such n, $\alpha c + (1-\alpha) b_n \in \varphi_{\xi_n}(C)$.

To summarize: for any $\alpha \in (0, 1)$ there exists $N(\alpha)$ such that, if $n \geq N(\alpha)$, then $\alpha c + (1-\alpha) b_n \in \varphi_{\xi_n}(C)$. Now define a sequence $\{\bar{b}_k\}$ as follows. Let $n_1 = 0$ and $\bar{b}_1 = c$; then, for $k \geq 2$, let $n_k = \max(N(\frac{1}{k}), n_{k-1}) + 1$ and$\bar{b}_k \equiv \frac{1}{k} c + \frac{k-1}{k} b_{n_k} \in \varphi_{\xi_{n_k}}(C)$. Now $\bar{b}_k(s) \to a(s)$

for all s:

$$\begin{aligned}
|\bar{b}_n(s) - a(s)| &= \left|\frac{1}{k}c(s) + \frac{k-1}{k}b_{n_k}(s) - \frac{1}{k}a(s) - \frac{k-1}{k}a(s)\right| \leq \\
&\leq \frac{1}{k}\left|c(s) - a(s)\right| + \frac{k-1}{k}\left|b_{n_k}(s) - a(s)\right| \to 0;
\end{aligned}$$

hence, by construction, $a \in \varphi_\xi(C)$.

Since a was arbitrarily chosen in C', the proof that $C' \subset \varphi_\xi(C)$ is complete. Since by assumption $\varphi_\xi(C) \subset C'$, actually $C' = \varphi_\xi(C)$; in particular, there is exactly one element of $\mathcal{C}_\xi$ that contains $\varphi_\xi(C)$. Thus, $\varphi_\xi : \mathcal{C}_0 \to \mathcal{C}_\xi$ is well-defined and, by construction, it satisfies (i) and (ii); furthermore, φ_ξ must be one-to-one, because $\varphi_\xi(C) = \varphi_\xi(D)$ for distinct $C, D \in \mathcal{C}_0$ violates the maximality of C and D.

To see that (iii) must also hold, suppose $a \in B(S, \Sigma)$ satisfies $a \approx c$, and consider an arbitrary $b \in \varphi_\xi(C)$; by construction, b is the (u.b.) pointwise limit of a sequence $\{b_n\}$ such that, for every $n \geq 1$, $b_n \in \varphi_{\xi_n}(C)$ for some $\xi_n < \xi$. Since $c \in \varphi_{\xi_n}(C)$ for all n as well, $a \approx c$ and Part (iii) of the induction hypothesis imply that $a \approx b_n$ for all n; therefore, by Lemma 5.4, Part 6, $a \approx b$, as required. If $a \in \mathcal{B}_\xi$, then, by Part (3) of Lemma 5.6, $a \approx b$ for all $b \in \varphi_\xi(C) \in \mathcal{C}_\xi$ implies $a \in \varphi_\xi(C)$.

Turn now to (iv). Fix $a \in \mathcal{B}_\xi$, and let $\{b_n\}$ be a sequence that converges pointwise to a such that, for all $n \geq 1$, $b_n \in \mathcal{B}_{\xi_n}$ for some $\xi_n < \xi$. By the induction hypothesis, Part (iv), $b_n \in \varphi_{\xi_n}(C_n)$ for some $C_n \in \mathcal{C}_0$. Let c_n be a critical point of C_n. Recall that $c_n \in \varphi_{\xi_n}(C_n)$, so clearly $b_n \approx c_n$. For $n \geq 1$, let $\bar{b}_n = \frac{1}{n}c_n + \frac{n-1}{n}b_n$; then clearly $\bar{b}_n \in \varphi_{\xi_n}(C_n)$, and for every $s \in S$,

$$\begin{aligned}
|\bar{b}_n(s) - a(s)| &= \left|\frac{1}{n}c_n(s) + \frac{n-1}{n}b_n(s) - \frac{1}{n}a(s) - \frac{n-1}{n}a(s)\right| \leq \\
&\leq \frac{1}{n}\left|c_n(s) - a(s)\right| + \frac{n-1}{n}\left|b_n(s) - a(s)\right| \leq \\
&\leq \frac{1}{n} \cdot (\|c_n\| + \|a\|) + \frac{n-1}{n}\left|b_n(s) - a(s)\right| \leq \\
&\leq \frac{2}{n} + \frac{n-1}{n}\left|b_n(s) - a(s)\right| \to 0,
\end{aligned}$$

i.e. $\{\bar{b}_n\}$ also converges pointwise to a. By Axiom 7, $a \approx \bar{b}_n$ for n sufficiently large; by Axiom 9, $a \approx c_n$ for such n; finally, from Part (iii) of the result, since $c_n \in \varphi_\xi(C_n)$ and $a \in \mathcal{B}_\xi$, $a \in \varphi_\xi(C_n)$. ■

The unique proper covering of $B(S, \Sigma)$ in Theorem 2.6 can finally be constructed:

Lemma 5.9 *There exists a unique proper covering $\mathcal{C}$ of $B(S,\Sigma)$ such that (i) for all $C \in \mathcal{C}$, $a, b \in C$ implies $a \approx b$, and (ii) if $a, b \in B(S,\Sigma)$ are such that $a \approx b$, then $a, b \in C$ for some $C \in \mathcal{C}$. Furthermore, $\mathcal{C}$ is finite or countable.*

Proof. I first construct a covering of $B_1(S,\Sigma)$. In the notation of Lemma 5.8, let

$$\mathcal{C}^* = \left\{ \bigcup_{\xi<\omega_1} \varphi_\xi(C) \ : \ C \in \mathcal{C}_0 \right\}. \tag{5}$$

$\mathcal{C}^*$ has the following properties. Clearly, *each $C^* \in \mathcal{C}^*$ contains all constants $\gamma 1_S$, $\gamma \in [-1,1]$.*

Each $C^* \in \mathcal{C}^*$ is a convex subset of $B_1(S,\Sigma)$, and $a, b \in C^*$ implies $a \approx b$: if $a, b \in C^* \in \mathcal{C}^*$, then $a \in \varphi_\eta(C)$, $b \in \varphi_\zeta(C)$ for some pair of ordinals η, ζ; hence, letting $\xi = \max(\eta,\zeta)$, $a, b \in \varphi_\xi(C)$, so $\alpha a + (1-\alpha)b \in \varphi_\xi(C) \subset C^*$ and $a \approx b$, as required.

$\mathcal{C}^*$ covers $B_1(S,\Sigma)$, because every $a \in B_1(S,\Sigma)$ belongs to some $\mathcal{B}_\xi$, $\xi < \omega_1$, and therefore, by Part (iv) of Lemma 5.8, to some $\varphi_\xi(C)$, $C \in \mathcal{C}_0$.

Every $C^* \in \mathcal{C}^*$ contains a function c with $\|c\| < 1$ and such that, for any $a \in B_1(S,\Sigma)$, $a \approx c$ implies $a \in C^*$—i.e. c is a critical point of C^*. Consider $C \in \mathcal{C}_0$ and let c be a critical point of C. Note that $\|c\| < 1$.[42] Let $C^* = \bigcup_{\xi<\omega_1} \varphi_\xi(C)$ and consider $a \in B_1(S,\Sigma)$. Then $a \in \mathcal{B}_\xi$ for some $\xi < \omega_1$: thus, by Part (iii) of Lemma 5.8, if $a \approx c$, then $a \in \varphi_\xi(C)$, which implies that $a \in C^*$.

Distinct elements of $\mathcal{C}^*$ are not nested. If $C^*, D^* \in \mathcal{C}^*$ and $C^* \subset D^*$, then D^* contains a critical point of C^*, and therefore, by the preceding property, $D^* \subset C^*$; thus, distinct elements of $\mathcal{C}^*$ cannot be nested.

Each element of $\mathcal{C}^*$ has non-empty interior. Consider $C^* \in \mathcal{C}^*$ and let c be a critical point of C^* with $\|c\| < 1$. Suppose that, for every $n > 0$ such that the $\frac{1}{n}$-ball around c lies in $B_1(S,\Sigma)$, there exists $b_n \in B_1(S,\Sigma)$ such that $\|b_n - c\| < \frac{1}{n}$ and $b_n \notin C^*$. Thus, $b_n \to c$ in the sup-norm topology, hence pointwise. But then Axiom 7 implies that $b_n \approx c$ for sufficiently large n. Since c is a critical point of C^*, $b_n \in C^*$ for such n, which contradicts the construction of the sequence $\{b_n\}$. Therefore, for some $n > 0$, C contains an open $\frac{1}{n}$-ball in $B_1(S,\Sigma)$ around c.

$C^* \cap D^*$ is an extremal subset of C^* (and D^*) for all distinct $C^*, D^* \in \mathcal{C}^*$. To see this, let $\alpha a + (1-\alpha)b \in C^* \cap D^*$ for $a, b \in C^*$ and $\alpha \in (0,1)$; also let d be a critical point of D^*.

[42] In the notation of Lemma 5.7, suppose $\|c_n\| = 1$ for all n. Then, for $k > 1$, no $\frac{1}{k}$-ball in $C(S)$ around $0 \in C$ contains an element of $\{c_n\}$. Thus, $\|c_n\| < 1$ for some n, so $\|c\| < 1$.

Then $a \approx b$ and $d \approx \alpha a + (1-\alpha)b$, so by Axiom 9, $d \approx a$ and $d \approx b$. Since d is a critical point of D^*, it follows that $a, b \in D^*$, so $a, b \in C^* \cap D^*$.

For all $a, b \in B_1(S, \Sigma)$: if $a \approx b$, then there exists $C^* \in \mathcal{C}^*$ such that $a, b \in C^*$. Let a, b be as stated, and define $d = \frac{1}{2}a + \frac{1}{2}b$. Then $d \in C^*$ for some $C^* \in \mathcal{C}^*$. If $a = b$, there is nothing to prove. Otherwise, note that in particular $c \approx d$ for a critical point c of C^*. By Axiom 9, this implies $c \approx a$, $c \approx b$; since c is a critical point of C^*, the assertion follows.

If a sequence $\{a_n\}$ in $B_1(S, \Sigma)$ converges pointwise to a, then there exists $N \geq 1$ such that, for every $n \geq N$, there exists $C^* \in \mathcal{C}^*$ such that $a_n, a \in C^*$. Note that, by Axiom 7, there exists $N \geq 1$ such that, for $n \geq N$, $a_n \approx a$. The previous property now implies the claim.

The collection $\mathcal{C}_0$ (hence, the collection $\mathcal{C}^*$) is at most countable. To see this, recall that $C(S) \cap B_1(S, \Sigma)$ is separable, and let $\{g_n\}_{n \geq 1}$ be an enumeration of a countable dense subset. Also, for every $C \in \mathcal{C}_0$, let c_C denote a critical point of C; since c_C lies in the non-empty interior of C, there exists $\epsilon_C > 0$ such that $\|a - c\| < \epsilon_C$ implies $a \in C$. Thus, for distinct $C, D \in \mathcal{C}_0$, it must be the case that $\|c_C - c_D\| \geq \max(\epsilon_C, \epsilon_D)$.[43] Since $\{g_n\}_{n \geq 1}$ is dense in $C(S) \cap B_1(S, \Sigma)$, for every $C \in \mathcal{C}_0$ one can choose $n \geq 1$ such that $\|g_n - c_C\| < \frac{\epsilon_C}{2}$. This defines a map $N : \mathcal{C}_0 \to \mathbb{N}$. Now suppose that, for distinct $C, D \in \mathcal{C}_0$, $N(C) = N(D) = n$. Then $\|c_C - c_D\| \leq \|c_C - g_n\| + \|g_n - c_D\| < \frac{1}{2}(\epsilon_C + \epsilon_D) \leq \max(\epsilon_C, \epsilon_D)$, a contradiction. Hence, $N(\cdot)$ is an injection, so $\mathcal{C}_0$ is finite or countable.

To complete the proof, let $\{C_1^*, C_2^*, ...\}$ be an enumeration of $\mathcal{C}^*$, and define

$$\forall n \geq 1, \;\; C_n = \{\gamma a : \gamma \in \mathbb{R}_{++}, a \in C_n^*\}; \qquad \mathcal{C} = \{C_n\}_{n \geq 1}. \tag{6}$$

That is, each C_n is the cone generated by C_n^*. It may be verified that $\{C_n\}$ is a proper covering that satisfies (i) and (ii) in the Lemma [details in the Online Appendix].

To establish uniqueness, suppose $\mathcal{D}$ is another proper covering such that (i) and (ii) hold. Fix $D \in \mathcal{D}$. Then D has non-empty topological, hence algebraic interior; furthermore, $a, b \in D$ implies $a \approx b$. By Part (1) of Lemma 5.5, there exists n such that $D \subset C_n$. On the other hand, since C_n also has non-empty topological, hence algebraic interior, and $a, b \in C_n$ implies $a \approx b$, by the same argument it is contained in some $D' \in \mathcal{D}$. Hence, $D \subset C_n \subset D'$; but since elements of $\mathcal{D}$ are non-nested, $D = D' = C_n$. Conclude that, for every $D \in \mathcal{D}$, there exists n such that $D = C_n$, so $\mathcal{D} \subset \mathcal{C}$. But the same argument implies that $\mathcal{C} \subset \mathcal{D}$, so $\mathcal{C} = \mathcal{D}$, and the proof is complete. ■

[43] If not, then $c_C \in D$, say, so for all $d \in D$, $c \approx d$, which implies $d \in C$, i.e. $D \subset C$. In particular, $c_D \in C$, and by a similar argument $C \subset D$: thus, $C = D$.

The following Corollary now establishes that $\{C_n\}_{n\geq 1}$ is in fact the collection of (conic hull of) sets delivered by Lemma 5.6: that is, it comprises all $\subset$-maximal convex cones containing the constants and consisting of robustly mixture-neutral functions. Furthermore, it shows that each C_n is norm-closed.

Corollary 5.10 *Let $\{C_n\}_{n\geq 1}$ be the proper covering provided by Lemma 5.9. If $C \subset B(S,\Sigma)$ is a convex cone that contains all constant functions $\gamma 1_S$, $\gamma \in \mathbb{R}$, and $a, b \in C$ implies $a \approx b$, then $C \subset C_n$ for some $C_n \in \mathcal{C}$. In particular, every C_n is norm-closed.*

Proof. Suppose not: then, for every $n \geq 1$, there exists $a_n \in C$ such that $a_n \notin C_n$. Since both C and C_n are cones, assume wlog that $\|a_n\| = 1$. As in the proof of Lemma 5.7, let $\{\alpha_n\}_{n\geq 1}$ be such that $\alpha_n \in (0,1)$ for all $n \geq 1$ and $\sum_n \alpha_n = 1$; then $\sum_n \alpha_n a_n$ converges in norm, say to $a \in B(S,\Sigma)$. Moreover, since C is a convex cone and $a_\ell \in C$ for all $\ell \geq 1$, $\sum_{\ell=1}^m \alpha_\ell a_\ell \in C$ for all $m \geq 1$; it follows that $a_n \approx \sum_{n=1}^m \alpha_n a_n$ for all $n, m \geq 1$, and taking the limit of the r.h.s. as $m \to \infty$, by Lemma 5.4, Part 6, $a_n \approx a$ for all $n \geq 1$. Now, since $\{C_n\}_{n\geq 1}$ covers $B(S,\Sigma)$, there exists $n^* \geq 1$ such that $a \in C_{n^*}$. As in Lemma 5.7, $a = \alpha_{n^*} a_{n^*} + (1-\alpha_{n^*}) \sum_{n\neq n^*} \frac{\alpha_n}{1-\alpha_{n^*}} a_n$, where, arguing as above, $\sum_{n\neq n^*} \frac{\alpha_n}{1-\alpha_{n^*}} a_n$ converges in $B(S,\Sigma)$ and satisfies $\sum_{n\neq n^*} \frac{\alpha_n}{1-\alpha_{n^*}} a_n \approx a_{n^*}$.

Thus, by (ii) in Lemma 5.9, $\sum_{n\neq n^*} \frac{\alpha_n}{1-\alpha_{n^*}} a_n, a_{n^*} \in C_{n^{**}}$ for some $n^{**} \geq 1$. Since $C_{n^{**}}$ is convex, $a \in C_{n^{**}}$ as well, so $a \in C_{n^*} \cap C_{n^{**}}$; and since $C_{n^*} \cap C_{n^{**}}$ is an extremal subset of $C_{n^{**}}$, $\sum_{n\neq n^*} \frac{\alpha_n}{1-\alpha_{n^*}} a_n, a_{n^*} \in C_{n^*} \cap C_{n^{**}} \subset C_{n^*}$. This contradicts the choice of a_{n^*}.

For the last implication, let $\bar{C}_n$ denote the norm-closure of some C_n, $n \geq 1$, and consider $a, b \in \bar{C}_n$; then $\|a_n - a\| \to 0$ and $\|b_n - b\| \to 0$ for some $\{a_n\}, \{b_n\} \subset C_n$. Now Lemma 5.9, Part (i) ensures that $a_n \approx b_n$ for each n, and Lemma 5.4, Part 6 implies that $a \approx b$. Thus, by the result just established, $\bar{C}_n \subset C_m$ for some $m \geq 1$; in particular, $n = m$, for otherwise $C_n \subset C_m$, which contradicts the fact that $\{C_n\}$ is a proper covering. Thus, C_n is norm-closed. ■

5.1.5 Sufficiency: completing the argument

Throughout the remainder of this section, $\mathcal{C}$ will denote the collection of convex cones constructed in Lemma 5.9.

First, the probabilities $\{P_n\}_{n\geq 1}$ representing I on each element of $\mathcal{C}$ will be constructed.

Lemma 5.11 *For every $n \geq 1$, there exists a unique, countably additive probability measure P_n on (S,Σ) such that, for all $a \in C_n$, $I(a) = \int a \, dP_n$.*

Proof. Observe first that $a, b \in C_n \in \mathcal{C}$ implies $a \approx b$, and hence, by Lemma 5.2, Part 5, $a \simeq b$: that is, the restriction of I to every $C_n \in \mathcal{C}$ is affine.

Now fix $n \geq 1$. Let c be a point in the interior of C_n: then, for every $a \in B(S, \Sigma)$, there exists $\alpha \in (0, 1]$ such that $a_c \equiv \alpha a + (1 - \alpha)c \in C_n$. Thus, $a = \frac{1}{\alpha} a_c - \frac{1-\alpha}{\alpha} c$. Since a was arbitrary and C_n is a cone, it follows that $\{a - b : a, b \in C_n\} = B(S, \Sigma)$.

It may be verified that, for every $n \geq 1$, it is possible to define a functional $J_n : B(S, \Sigma) \to \mathbb{R}$ by letting $J_n(a - b) = I(a) - I(b)$ for all $a, b \in C_n$. Furthermore, each J_n is positive, additive, and homogeneous.

Thus, J_n is linear and norm-continuous (cf. [2], Theorem 7.6) on $B(S, \Sigma)$; furthermore, $\|J_n\| = 1$. Therefore, there exists a unique, finitely additive probability P_n on (S, Σ) such that $J_n(a) = \int a \, dP_n$ for all $a \in B(S, \Sigma)$ (cf. e.g. [2], Theorem 11.32). In particular, $I(a) = J_n(a - 0) = \int a \, dP_n$ for all $a \in C_n$.

Now consider a sequence of events $\{A_k\}_{\geq 1}$ such that $A_k \supset A_{k+1}$, $k \geq 1$, and $\bigcap_{k \geq 1} A_k = \emptyset$. Let c be an interior point of C_n such that, for some $\epsilon > 0$, $a \in B(S, \Sigma)$ and $\|a - c\| < \epsilon$ implies $a \in C_n$. Consider the sequence $\{c_k\}_{k \geq 1}$ of functions defined by $c_k(s) = c(s) + \frac{\epsilon}{2} 1_{A_k}(s)$ for all $s \in S$ and $k \geq 1$. Then $\|c_k - c\| = \frac{\epsilon}{2} < \epsilon$, so $c_k \in C_n$ for all $k \geq 1$. Since I is monotonic, $I(c_k) \geq I(c_{k+1})$ for all $k \geq 1$; furthermore, for every $s \in S$ there exists $K \geq 1$ such that, for all $k \geq K$, $s \notin A_k$, and therefore $c_k(s) = c(s)$. Thus, $c_k(s) \to c(s)$ for all $s \in S$, and the sequence $\{c_k\}$ is clearly bounded (e.g. by $\|c\| + \epsilon$). Axiom 8 then implies that $I(c_k) \downarrow I(c)$, so that

$$P_n(A_k) = J_n(1_{A_k}) = \frac{2}{\epsilon} J_n(c_k - c) = \frac{2}{\epsilon}[I(c_k) - I(c)] \downarrow 0;$$

thus (cf. e.g. [2], Lemma 8.32), P_n is countably additive. ∎

Remark 5 The preceding Lemma provides the key step in the proof of Proposition 2.2 for the case $C \subsetneq L$. Specifically, suppose that $\succeq$ satisfies Axioms 1–5 and Mixture Neutrality on such a set C of acts. Then, by Lemma 5.1 and arguments in the proof of Lemma 5.3 Part 1, $\succeq$ is represented by a cardinally unique u and a monotonic, c-linear, normalized functional I that is affine on the set $\bar{C} = \{a \in B(S, \Sigma) : a = \gamma u \circ f, \gamma \in \mathbb{R}_+, f \in C\}$. Proceeding as in the proof of Lemma 5.11, define a positive linear functional J on $\bar{C} - \bar{C} \subset B(S, \Sigma)$ by letting $J(a - b) = I(a) - I(b)$; notice that, in general, $\bar{C} - \bar{C} \neq B(S, \Sigma)$. Now let J^* denote a positive Hahn-Banach extension of J to $B(S, \Sigma)$, and apply the standard representation theorem to obtain a probability charge P on (S, Σ) such that $J^*(a) = \int a \, dP$ for all $a \in B(S, \Sigma)$, and in particular $I(u \circ f) = J(u \circ f) = J^*(u \circ f) = \int u \circ f \, dP$ for all $f \in C$. Note that, since the Hahn-Banach extension J^* is *not* unique in general, P is *not* the only probability charge that represents $\succeq$ on C.

Jointly with the cardinally unique utility function u from Lemma 5.1 and the unique proper covering $\mathcal{C}$ from Lemma 5.9, the collection $\{P_n\}_{n \geq 1}$ of unique countably additive probability measures from Lemma 5.11 satisfies (i) of Theorem 2.6, Part (2). By Lemma 5.9, it is also the case that, if $f \approx g$, hence $u \circ f \approx u \circ g$, then $u \circ f, u \circ g \in C_n$ for some $n \geq 1$: that is, (ii) also holds. Hence, to complete the proof of the Theorem, it must be shown that (iii) holds as well. The following, preliminary Lemma provides the key step.

Lemma 5.12 *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 $n_k \geq 1$ such that $\{\alpha \in [0,1] : \alpha a + (1-\alpha) b \in C_{n_k}\} = [\alpha_k, \alpha_{k+1}]$.*

Proof. Let $\alpha_0 = 0$; for $k \geq 1$, let

$$\alpha_k = \sup\{\alpha \in [\alpha_{k-1}, 1] : \alpha a + (1-\alpha) b \approx \alpha_{k-1} a + (1-\alpha_{k-1}) b\}.$$

By Lemma 5.4, Part 6, $\alpha_k a + (1-\alpha_k) b \approx \alpha_{k-1} a + (1-\alpha_{k-1}) b$ (i.e. the supremum is achieved).

I claim that, if $\alpha_{k-1} < 1$, then $\alpha_k > \alpha_{k-1}$. To see this, consider the sequence $\{b^m\}$ of functions defined by $b^m = \frac{1}{m} a + \frac{m-1}{m}[\alpha_{k-1} a + (1-\alpha_{k-1}) b]$; thus, $b^m \to \alpha_{k-1} a + (1-\alpha_{k-1}) b$ in norm. Axiom 7 then implies that there exists $M \geq 1$ such that $b^m \approx \alpha_{k-1} a + (1-\alpha_{k-1}) b$ for $m \geq M$; thus, $\alpha_k \geq \frac{1}{M} + \frac{M-1}{M}\alpha_{k-1} > \alpha_{k-1}$.

Next, I claim that there exists $K \geq 1$ such that $\alpha_K = 1$ (hence, $\alpha_k = 1$ for all $k > K$) and $\alpha_k < 1$ for all $k < K$. Suppose not; then $\alpha_k \uparrow \bar{\alpha} \in [0,1]$. Consider the sequence $\{b_k\}$ of functions defined by $b^k = \alpha_k a + (1-\alpha_k) b$; clearly, $b^k \to \bar{\alpha} a + (1-\bar{\alpha}) b$, so Axiom 7 implies that $b^k \approx \bar{\alpha} a + (1-\bar{\alpha}) b$ for large k. In other words, for large k, $\alpha_k a + (1-\alpha_k) b \approx \bar{\alpha} a + (1-\bar{\alpha}) b$; but this contradicts the fact that $\sup\{\alpha \in [\alpha_k, 1] : \alpha a + (1-\alpha) b \approx \alpha_k a + (1-\alpha_k) b\} = \alpha_{k+1} < \bar{\alpha}$.

Now Part (ii) of Lemma 5.9 and convexity of each C_n, $n \geq 1$, imply that, for every $k = 0, ..., K-1$, there exists $n_k \geq 1$ such that $\alpha a + (1-\alpha) b \in C_{n_k}$ for all $\alpha \in [\alpha_k, \alpha_{k+1}]$. Moreover, for $k > 0$, consider $\alpha \in [\alpha_h, \alpha_{h+1})$ for some $h < k$; then $\alpha a + (1-\alpha) b \not\approx \alpha_{k+1} a + (1-\alpha_{k+1}) b \in C_{n_k}$, for otherwise $\alpha_h a + (1-\alpha_h) b \approx \alpha_{k+1} a + (1-\alpha_{k+1}) b$ (either because $\alpha = \alpha_h$, or by Axiom 9), which contradicts the fact that $\alpha_{h+1} < \alpha_{k+1}$. Similarly, consider $\alpha \in (\alpha_h, \alpha_{h+1}]$ for some $h > k$; then $\alpha a + (1-\alpha) b \not\approx \alpha_k a + (1-\alpha_k) b \in C_{n_k}$, for otherwise $\alpha_{h+1} a + (1-\alpha_{h+1}) b \approx \alpha_k a + (1-\alpha_k) b$, which contradicts the fact that $\alpha_{k+1} < \alpha_{h+1}$. Thus, if $\alpha \in [0,1] \setminus [\alpha_k, \alpha_{k+1}]$, then either $\alpha a + (1-\alpha) b \not\approx \alpha_k a + (1-\alpha_k) b$, or $\alpha a + (1-\alpha) b \not\approx \alpha_{k+1} a + (1-\alpha_{k+1}) b$; hence, by Part (i) of Lemma 5.9, $\alpha a + (1-\alpha) b \notin C_{n_k}$. ■

Part (iii) of Theorem 2.6 can now be established. Assume that, for $a, b \in B(S, \Sigma)$, it is the case that $\int a \, dP_n \geq \int b \, dP_n$ for all $n \geq 1$. Let K, α_k and n_k be as in Lemma 5.12. For every $k = 0, \ldots, K-1$,

$$\int [\alpha_k a + (1-\alpha_k) b] \, dP_{n_k} \leq \int [\alpha_{k+1} a + (1-\alpha)_{k+1} b] \, dP_{n_k} = \int [\alpha_{k+1} a + (1-\alpha)_{k+1} b] \, dP_{n_{k+1}} :$$

the inequality follows from $\int a\,dP_{n_k} \geq \int b\,dP_{n_k}$ 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, $I(\alpha_k a+(1-\alpha_k)b) \leq I(\alpha_{k+1}a+(1-\alpha_{k+1})b)$ for all $k = 0, \ldots, K-1$; since $\alpha_0 = 0$ and $\alpha_K = 1$, $I(b) \leq I(a)$.

Finally, Lemma 5.13 implies that $\{P_n\}_{n\geq 1}$ comprises *all* possible priors for $\succeq$. To see this, assume that the set D below is the utility image of a collection C of acts as in Definition 2.3, and P is the unique prior identified by C. Then necessarily $P = P_n$ for some $n \geq 1$.

Lemma 5.13 *Let $D \subset B(S,\Sigma)$ be a convex set such that $a, b \in Y$ implies $a \simeq b$. Then there exists $n \geq 1$ such that $I(a) = \int a\,dP_n$ for all $a \in D$.*

Proof. Since I is norm-continuous, it is wlog to assume that D is norm-closed. Thus, D, endowed with its relative metric topology, is complete. For every $n \geq 1$, C_n is norm-closed by Corollary 5.10, so $D \cap C_n$ is relatively closed, and $D = \bigcup_{n\geq 1} D \cap C_n$. Thus, by the Baire Category Theorem ([2], Corollary 3.28), there exists $n \geq 1$ such that $D \cap C_n$ has non-empty relative interior. In particular, there exists $c \in D \cap C_n$ and $\epsilon > 0$ such that $\|a - c\| < \epsilon$ and $a \in D$ imply $a \in D \cap C_n$.

Now let $J_n : B(S,\Sigma) \to \mathbb{R}$ be a linear functional such that $J_n(a) = I(a)$ for all $a \in D \cap C_n$. I claim that then $J_n(a) = I(a)$ for all $a \in D$. To see this, consider $a \in Y$; by the preceding argument, since D is convex, there exists $\gamma \in (0,1]$ such that $\gamma a + (1-\gamma)c \in D \cap C_n$. Thus,

$$\begin{aligned}\gamma J_n(a) &= J_n(\gamma a + (1-\gamma)c) - (1-\gamma)J_n(c) = \\ &= I(\gamma a + (1-\gamma)c) - (1-\gamma)I(c) = \\ &= \gamma I(a) + (1-\gamma)I(c) - (1-\gamma)I(c) = \gamma I(a).\end{aligned}$$

Therefore, in particular, $\int a\,dP_n = I(a)$ for all $a \in D$. ■

5.2 Proof of Proposition 3.2

Denote by μ the convex-ranged probability charge in Def. 3.1. As in Subsection 5.1.2, let $I(a) = \int a\,dP_n$ for all $a \in C_n$ and $n \geq 1$. I sometimes write $P_n(a)$ in lieu of $\int a\,dP_n$. Also assume that $u(X) \supset [-1,1]$ (recall that $u(X)$ is convex by assumption).

I claim that μ is countably additive. To see this, 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 8, 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. 3.1, 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, $\mu(A_k) \downarrow 0$, so μ is countably additive.

Let $\mathcal{C}_0 = \{C_n\}_{n\geq 1}$, $\mathcal{C}^* = \{C^*_n\}_{n\geq 1}$ and $\{P_n\}_{n\geq 1}$ be as in the proof of Theorem 2.6. It is wlog to assume that S is compact metrizable; since μ is convex-valued, S must be uncountable

It is w.l.o.g. to assume that μ has full support. To elaborate, since μ is countably additive and $\mu(\{s\}) = 0$ for all $s \in S$ because μ is convex-ranged, the Borel Isomorphism theorem for measures ([20], Theorem 17.41) yields a Borel isomorphism $\varphi : S \to [0, 1]$ such that $\mu \circ \varphi^{-1}$ is Lebesgue measure. Arguing as in Subsection 6.5.5 of the Online Appendix, if the plausible priors determined by the functional $I \circ \varphi^{-1}$ agree with Lebesgue measure, the plausible priors for I coincide with μ (for an alternative, direct proof, see the Online Appendix, §6.5.7).

It will be shown that $P_n = \mu$ for all $n \geq 1$. Fix an arbitrary $n \geq 1$ and let c be a critical point of $C_n \subset C(S) \cap B_1(S, \Sigma)$. If c is constant, then $C_n = C(S) \cap B_1(S, \Sigma)$, i.e. $\succeq$ is EU; thus, assume c is nonconstant. Recall that c is in the interior of C^*_n, so for some $\epsilon > 0$, $\|a - c\| < 2\epsilon$ implies $a \in C^*_n$. Define $c_{\min} = \min_s c(s)$, $c_{\max} = \max_s c(s)$: then $-1 + 2\epsilon \leq c_{\min} < c_{\max} \leq 1 - 2\epsilon$; finally, let $R = c_{\max} - c_{\min} > 0$.

Also, c is the uniform limit of the sequence of step functions $\{a_M\}_{M\geq 1}$ defined by $a_M(s) = c_{\min} + \frac{R}{M}(m - 1)$ whenever $c(s) \in [c_{\min} + \frac{R}{M}(m - 1), c_{\min} + \frac{R}{M}m)$ for $m = 1, ..., M - 1$, and $a_M(s) = c_{\min} + \frac{R}{M}(M - 1)$ whenever $c(s) \in [c_{\min} + \frac{R}{M}(M - 1), c_{\max}]$. For $M > \frac{R}{\epsilon}$, $\|a_M - c\| = \frac{R}{M} < \epsilon$ (hence, $a_M \in C^*_n$) and furthermore $\min\{a_M(s) - a_M(t) : a_M(s) > a_M(t)\} = \frac{R}{M} < \epsilon$. Fix such a value of M; for simplicity, denote the corresponding step function a_M by a, and let $f \in L$ be a simple act such that $u \circ f = a$; write $f = (x_1, E_1; \ldots x_M, E_M)$, where $u(x_m) = c_{\min} + \frac{R}{M}(m - 1)$. Since μ has full support, $\mu(E_m) > 0$ for all m.

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: Let $x \in X$ be such that $u(x) = u(x_m) + \frac{R}{M}$; note that $x = x_{m+1}$ if $m < M$. 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 $u \circ f' \in B_1(S, \Sigma)$, and $\|u \circ f' - a\| \leq \|u \circ f' - u \circ f\| + \|u \circ f - a\| < 2\epsilon$, so $u \circ f' \in C^*_n$.

Then Def. 3.1 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.

By range convexity of μ, for every $K \geq 1$ there exists a partition $\{E^1_m, ..., E^K_m\}$ of E_m such that $\mu(E^k_m) = \frac{1}{K}\mu(E_m)$ for all $k = 1, \ldots, K$. For each such k, construct acts f^k such that $f^k(s) = f(s)$ for all $s \in S \setminus E^k_m$, and $f^k(s) = x$ for $s \in E^k_m$, where $u(x) = u(x_m) + \frac{R}{M}$ as above. Then $u \circ f^k \in C^*_n$; furthermore, Def. 3.1 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^k_m and E^h_m, a simple calculation shows that $P_n(E^k_m) = P_n(E^h_m)$, so $P_n(E^k_m) = \frac{1}{K} P_n(E_m)$. 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 $\frac{\mu(F)+\mu(F_r)}{\mu(E_m)} = r$,[44] 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, $u(x_m) - u(x_{m-1}) = \frac{R}{M} < \epsilon$, so $\|f' - a\| \leq \|f' - f\| + \|f - a\| < 2\epsilon$, 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

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

[44] 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.

5.3 Proof of Theorem 3.4

5.3.1 Notation and Preliminary results

Let u, $\mathcal{C} = \{C_n\}_{n\geq 1}$ and $\{P_n\}_{n\geq 1}$ represent $\succeq$; similarly, let u^E, $\mathcal{C}^E = \{C_k^E\}_{k\geq 1}$ and $\{P_k^E\}_{n\geq 1}$ represent $\succeq_E$. As in Subsection 5.1.2, let $I(a) = \int a\, dP_n$ for all $a \in C_n$ and $n \geq 1$, and similarly let $I^E(a) = \int a\, dP_k^E$ for all $a \in C_k^E$ and $k \geq 1$. Recall that I and I^E are monotonic, normalized, c-linear functionals. 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 if and only if, 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 $I(a) > I(b)$.

Overview. Lemma 5.14 characterizes non-null events in terms of the plausible priors $\{P_n\}$. Lemma 5.15 examines the "fixpoint condition" discussed after Theorem 3.4.

Lemma 5.14 *An event $E \in \Sigma$ is non-null for $\succeq$ if and only if, for all $n \geq 1$, $P_n(E) > 0$.*

Proof. Suppose E is non-null and let $c \in B(S, \Sigma)$ lie in the interior of C_n. Then there exists $\epsilon > 0$ such that the function c' defined by $c'(s) = c(s)$ for $s \in S \setminus E$ and $c'(s) = c(s) + \epsilon$ for $s \in E$ satisfies $c' \in C_n$. Thus, $P_n(E) = \frac{1}{\epsilon}[\int c'\, dP_n - \int c\, dP_n] = \frac{1}{\epsilon}[I(c') - I(c)] > 0$.

Conversely, assume $P_n(E) > 0$ for all $n \geq 1$, and let $a, b \in B(S,\Sigma)$ be such that $a(s) = b(s)$ for $s \in S \setminus E$, and $a(s) > b(s)$ for $s \in E$. By Part (3) in Def. 2.5, there exists $\gamma > 0$ such that $\gamma a + (1-\gamma)b,\ b \in C_n$ for some $n \geq 1$. Since $\gamma a(s) + (1-\gamma)b(s) = b(s)$ for $s \in S \setminus E$, $\gamma a(s) + (1-\gamma)b(s) > b(s)$ for $s \in E$, $P_n(E) > 0$ and P_n is countably additive, $I(\gamma a + (1-\gamma)b) = \int [\gamma a + (1-\gamma) b]\, dP_n > \int b\, dP_n = I(b)$. Furthermore, since $a(s) \geq \gamma a(s) + (1-\gamma)b(s)$ for all $s \in S$, $\int a\, dP_m \geq \int [\gamma a + (1-\gamma)b]\, dP_m$ for all $m \geq 1$; hence, by c-linearity of I [considering acts f, g such that $u \circ f = \alpha a$, $u \circ g = \alpha b$ for appropriate $\alpha > 0$] and Part (iii) of Theorem 2.6, $I(a) \geq I(\gamma a + (1-\gamma)b) > I(b)$. ■

Lemma 5.15 *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*

$$x = I(aEx). \tag{7}$$

The map $J : B(S,\Sigma) \to \mathbb{R}$ associating to each $a \in B(S,\Sigma)$ the unique solution to Eq. (7) is monotonic, c-linear and normalized.

Proof. 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 $s \in E$.

Verifying the other properties of J is straightforward, so the proof is omitted. ■

5.3.2 Necessity of the Axioms

Now turn to the proof of Theorem 3.4. To show that (2) implies (1), consider a non-null $E \in \Sigma$ and assume that $u^E = u$ (clearly w.l.o.g.) and, for all k, $P^E_k = P_{n_k}(\cdot|E)$ for some $n_k \geq 1$ such that Eq. (3) holds; conditional probabilities are well-defined by Lemma 5.14. Since $P_{n_k}(S \setminus E|E) = 0$ for all $k \geq 1$, it is clear that $\succeq_E$ satisfies Axiom 11. It remains to be shown that $\succeq, \succeq_E$ jointly satisfy Axiom 12, Dynamic c-Consistency.

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 $m \geq 1$; then, by Eq. (3), $\int u \circ [fEy] \, dP_m = \int u \circ f \, E \, u(y) \, dP_m = u(y)$, 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 5.15. 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 5.15, the solution to Eq. (7) 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.

5.3.3 Sufficiency of the Axioms

Turn now to the proof that (1) implies (2). Begin with two preliminary claims.

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 11, $fEy \sim_E f \succeq_E y$. Clearly, $fEy(s) \sim y$ for all

$s \in E^c$. Thus, by Axiom 12, $fEy \succeq y$. 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', $y \succeq_E y' \Leftrightarrow y \succeq y$.

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, it is wlog to assume $u^E = u$. 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 5.15 represents $\succeq^E$. Hence (cf. Lemma 5.1), for all $a \in B(S, \Sigma)$, $J(a) = I^E(a) = I(a\, E\, I^E(a))$.

By assumption, $I^E(a) = \int a\, dP^E_k$ whenever $a \in C^E_k$. It must now be verified that, for every $k \geq 1$, Eq. (3) holds, and $P^E_k = P_{n_k}$ for some $n_k \geq 1$. Fix $k \geq 1$ and consider the set

$$D_k = \{a\, E\, I^E(a)\ : a \in C^E_k\}.$$

D_k is convex: if $a\, E\, I^E(a),\ b\, E\, I^E(b) \in D_k$, then $a, b, \alpha a + (1-\alpha)b \in C^E_k$; also, $\alpha[a\, E\, I^E(a)] + (1-\alpha)[b\, E\, I^E(b)] = [\alpha a + (1-\alpha)b]\, E\, [\alpha I^E(a) + (1-\alpha)I^E(b)] = [\alpha a + (1-\alpha)b]\, E\, I^E(\alpha a + (1-\alpha)b) \in D_k$, because I^E is affine on C^E_k [cf. Lemma 5.9, Part (i) and Lemma 5.4, Part 5.]

Furthermore, consider $a', b' \in D_k$, so $a' = aEI^E(a)$, $b' = bEI^E(b)$ for some $a, b \in C^E_k$. Then, for all $\alpha \in [0,1]$, $I(\alpha a' + (1-\alpha)b') = I([\alpha a + (1-\alpha)b]\, E\, [\alpha I^E(a) + (1-\alpha)I^E(b)]) = I([\alpha a + (1-\alpha)b]\, E\, I^E(\alpha a + (1-\alpha)b)) = I^E(\alpha a + (1-\alpha)b) = \alpha I^E(a) + (1-\alpha)I^E(b) = \alpha I(aEI^E(a)) + (1-\alpha)I(bEI^E(b)) = \alpha I(a') + (1-\alpha)I(b')$. That is, $a' \simeq b'$ for all $a', b' \in D_k$.

Now, by Lemma 5.13, there exists $n_k \geq 1$ such that $I(a') = \int a'\, dP_{n_k}$ for every $a' \in D_k$; thus, for every $a \in C^E_k$,

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

so $I^E(a) = \int a\, dP_{n_k}(\cdot|E)$. Also, for all $a \in C^E_k$, if $x = I^E(a) = \int a\, dP_{n_k}(\cdot|E)$ and $aEx \in C_m$,

$$\int aEx\, dP_m = I(aEx) = \int aEx\, dP_{n_k} = \int aEx\, dP_{n_k}(\cdot|E) = x,$$

i.e Eq. (3) holds; finally, since P^E_k is the unique measure representing $\succeq^E$ on C^E_k, $P^E_k = P_{n_k}(\cdot|E)$, and the proof of Theorem 3.4 is complete.

References

[1] Anscombe, F. and R. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34, 199-205.

[2] Aliprantis, C., and K. Border (1994): *Infinite Dimensional Analysis.* Springer Verlag, Berlin.

[3] Bewley, T. (1986): "Knightian Decision Theory: Part I," Technical Report 807, Cowles Foundation, Yale University.

[4] Billot, A., Chateauneuf, A., Gilboa, I., and Tallon, J.M. (2000): "Sharing Beliefs: Between Agreeing and Disagreeing," *Econometrica*, 68,685-694.

[5] Camerer, C., and M. Weber (1992): "Recent Developments in Modelling Preferences: Uncertainty and Ambiguity," *Journal of Risk and Uncertainty*, 5, 325-370.

[6] Castagnoli, E. and Maccheroni, F. (2000): "Restricting Independence to Convex Cones," *Journal of Mathematical Economics* 34, 215-223.

[7] Castagnoli, E., Maccheroni, F., and Marinacci, M. (2003): "Expected Utility with Multiple Priors," mimeo, Università Bocconi.

[8] Ellsberg, D. (1961): "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics* 75, 643-669.

[9] Epstein, L. and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, 283-322.

[10] Epstein, L. and J. Zhang (2001): "Subjective Probabilities on Subjectively Unambiguous Events," *Econometrica*, 69, 265-306.

[11] Ghirardato, P., Klibanoff, P., and M. Marinacci (1998): "Additivity with Multiple Priors," *Journal of Mathematical Economics*, 30, 405-420.

[12] Ghirardato, P., Maccheroni, F., and M. Marinacci (2002): "Ambiguity from the Differential Viewpoint," Caltech Social Science Working Paper 1130.

[13] Ghirardato, P. and M. Marinacci (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102, 251-289.

[14] Ghirardato, P., Maccheroni, F., Marinacci, M., and M. Siniscalchi (2003): “A Subjective Spin on Roulette Wheels,” *Econometrica*, forthcoming.

[15] Gilboa, I. and D. Schmeidler (1989): “Maxmin Expected Utility with Non-Unique Prior,” *Journal of Mathematical Economics*, 18, 141-153.

[16] Gilboa, I. and D. Schmeidler (1993): “Updating Ambiguous Beliefs,” *Journal of Economic Theory*, 59, 33-49.

[17] Hansen, L.P., T. Sargent, and T. Tallarini (1999): “Robust Permanent Income and Pricing,” *Review of Economic Studies*, 66(4), 873-908.

[18] Holmes, R.B. (1975), *Geometric Functional Analysis and its Applications*: New York, Springer-Verlag.

[19] Jaffray, J. Y. (1994): “Dynamic Decision Making with Belief Functions,” in *Advances in the Dempster-Shafer Theory of Evidence*, Yaeger et al., eds.; Wiley.

[20] Kechris, A. S. (1994), *Classical Descriptive Set Theory*: New York, Springer-Verlag.

[21] Klibanoff, P. (2001): “Characterizing Uncertainty Aversion Through Preferences for Mixtures,” *Social Choice and Welfare*, 18, 289-301.

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

[23] Machina, M. and D. Schmeidler (1992): “A More Robust Definition of Subjective Probability,” *Econometrica* 60, 4, 745-780.

[24] Machina, M. and D. Schmeidler (1995): “Bayes Without Bernoulli: Simple Conditions for Probabilistically Sophisticated Choice.” *Journal of Economic Theory*, 67, 106-128.

[25] Machina, M. (2002): “Robustifying the Classical Model of Risk Preferences and Beliefs,” UCSD Economics Discussion Paper 2002-06.

[26] Mukerji, S. (1998): “Ambiguity Aversion and Incompleteness of Contractual Form.” *American Economic Review*, 88 n.5, 1207-1231.

[27] Nehring, K. (2001): “Ambiguity in the Context of Probabilistic Beliefs,” mimeo, UC Davis.

[28] Pires, C. P. (2001): “A Rule for Updating Ambiguous Beliefs,” mimeo, Universidade de Évora.

[29] Savage, L.J. (1954), *The Foundations of Statistics*. Wiley, New York.

[30] Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571-587.

[31] Siniscalchi, M. (2000): "Vector-Adjusted Expected Utility," mimeo, Princeton.

[32] Siniscalchi, M. (2001): "Bayesian Updating for General Maxmin-Expected Utility Preferences," mimeo, Princeton.

[33] Wang, T. (2003): "A Class of Multi-Prior Preferences," mimeo, University of British Columbia.