Imprecise Probabilistic Beliefs

Klaus Nehring

University of California, Davis*

First Version: November 2002†

ABSTRACT

Imprecise probabilistic beliefs are modelled as incomplete comparative likelihood relations admitting a multiple-prior representation. We provide an axiomatization of such relations for the case in which the set of priors is "convex-ranged". We also show that the multiple-prior representation is unique whenever the set of priors is "almost-convex-ranged". Such uniqueness ensures the adequacy of likelihood relations as models of probabilistic beliefs.

In the second part of the paper, two conditions relating preferences and specified probabilistic beliefs are proposed. The weaker one requires simply that the decision maker prefer betting on events in line with specified likelihood comparisons. A somewhat stronger one amounts to a requirement of probabilistic sophistication relative to the specified probabilistic beliefs; it leads to an Anscombe-Aumann like structure derived in an epistemically enriched Savage framework.

*e-mail: kdnehring@ucdavis.edu ; homepage: http://www.econ.ucdavis.edu/faculty/nehring/

†Precursors of some of the ideas of this paper were presented in December 1996 at LOFT2 in Torino, Italy, July 2001 at RUD in Venice, Italy, and at the University of Heidelberg, and June 2002 at RUD in Paris.

1. INTRODUCTION

In this paper, we shall study decision makers who have precise probabilistic beliefs about some events, while their beliefs about others may be imprecise or "ambiguous". Behaviorally, such ambiguity is characterized by violations of the sure-thing principle exemplified by the Ellsberg's (1961) celebrated experiments. While this and much other evidence provide compelling reasons for abandoning the assumption that behavior can be globally explained in terms of precise probabilistic beliefs, they do not render the notion of probabilistic belief useless if it is applied "locally", that is: if applied to some events or event comparisons. Indeed, the very formulation of Ellsberg's original experiment suggests a comparison of events with probabilistic beliefs to other potentially ambiguous ones. Likewise, in many economically relevant situations in which ambiguity is plausible and interesting, the existence of partial probabilistic beliefs is plausible as well. For example, while the expected return on equity (as a whole) is notoriously hard to pin down through statistical information, this is much less true for the volatility of the return. Thus, it seems much more plausible to assume precise probabilistic beliefs on the latter than on the former. Likewise, in the context of games under incomplete information, it may well make sense to ascribe to players precise beliefs about others' types along with ambiguity about their strategy choices.

To model such "decision making in the context of probabilistic beliefs", we shall propose to represent a decision-maker in terms of two entities, a preference relation and a (not necessarily exhaustive) description of his probabilistic beliefs. This departure from the behaviorist tradition following Ramsey and Savage of *defining* beliefs in terms of preferences is motivated by the loss of a canonical one-to-one relation between beliefs and betting preferences in the presence of ambiguity, for now there are (at least) two determinants of betting preferences: beliefs –however construed– and ambiguity attitudes. Both common sense and the practice of economic modeling support an independent, non-derived role for beliefs: as real-world actors, we prefer certain acts over others *because* we have certain beliefs rather than others; as economic modellers, we attribute to economic agents particular preferences over uncertain acts *because* we have some idea what beliefs are reasonable in a particular situation. In both cases, we think directly in terms of beliefs rather than preferences. This really is the substance of the proposed dual framework.

Two basic, interrelated questions arise:

"How are preferences (rationally) constrained by probabilistic beliefs ?",

and, more fundamentally,

"How are imprecise probabilistic beliefs themselves to be represented formally?"

Imprecise Probabilistic Beliefs as Comparative Likelihood Relations

Following the lead of Keynes (1921), Koopmans (1940), and Savage (1954), we shall model "imprecise probabilistic beliefs" formally as comparative likelihood relations $\trianglerighteq$ over events, with "$A \trianglerighteq B$" denoting the judgement "A is at least as likely as B". Comparative likelihood relations constrain preferences canonically: if A is judged at least as likely as B, then betting on A must be weakly preferred to betting on B; in this case, the preference relation will be said to be *compatible* with the likelihood relation. A "decision maker in the context of probabilistic beliefs" is thus given as a pair $(\succsim, \trianglerighteq)$ such that $\succsim$ and $\trianglerighteq$ are compatible. The likelihood relation will frequently be referred to as the DM's "explicit" probabilistic beliefs; it will typically be *non-exhaustive* in the sense that the DM may have further "non-explicit" probabilistic judgments not listed in $\trianglerighteq$.

An important virtue of using likelihood relations as the epistemic primitive is their *behavioral generality,* in that our formulation does not constrain the DM's risk or ambiguity attitudes; in particular, Compatibility accomodates Allais- and Ellsberg-style choice patterns as well as their converses, and is not tied to assumptions about functional form. Behavioral generality is important since, as argued compellingly by Machina-Schmeidler (1992) and Epstein-Zhang (2001), issues about the representation of probabilistic beliefs are more fundamental than particular behavioral assumptions. The goal of the paper is a) to provide axiomatic foundations for incomplete comparative likelihood relations, and b), to demonstrate that such relations are adequately expressive in sufficiently general circumstances.

Representation by Multiple Priors

The incompleteness of the set of explicit likelihood judgments is naturally reflected in a representation in terms of a sets of admissible probability measures ("priors") according to which judging an event A as at least as likely than B is equivalent to A's probability weakly exceeding that of B, for any admissible probability in the set. A comparative likelihood relation for which such a multi-prior representation exists will be called *coherent.*

An axiomatization of coherent likelihood relations will rely on conditions of three kinds; rationality axioms that account for the logical interrelations among various judgements, conditions reflecting the real-valued character of the desired representation, and auxiliary "boosting" assumptions that are not implied by coherence as such but that are needed to make the other conditions sufficiently powerful. Savage (1954) achieved a representation of this kind under completeness. In particular, by an appropriate choice of auxiliary conditions, he was able to make do with one fundamental rationality axiom, "Additivity", according to which the judgment that A is at least as likely as B entails and is entailed by the judgment that "A or C" is at least as likely as "B or C", for any event C disjoint from A and B. At the same time, Savage had to pay a price by restricting attention to "convex-ranged" (finitely additive) probability measures. (In the countably additivite case axiomatized later by Villegas (#), convex-rangedness is equivalent to the absence of probability atoms, a very natural regularity condition).

The main result of this paper, Theorem 2, presents a counterpart (and generalization of sorts) of Savage's result to incomplete comparative likelihood relations; it is apparently the first such result in the literature.[1] If the completeness assumption is dropped, almost all of Savage's axioms need to be modified or augmented. In particular, Additivity is no longer enough to fully capture the "logical syntax of probability"; a second rationality axiom called "Splitting" is needed as well. This axiom requires that if two events A and B are split into two equally likely parts, and if A is judged at least as likely as B, then any "half" of A must be at least as likely as any "half" of B. "Splitting" is accompanied by an auxiliary "Equidivisibility" condition that assumes any event can indeed be split into two equally likely subevents. Equidivisibility leads to convex-rangedness of the *set* of priors. That is, given any non-null event and any value between 0 and 1, there exists a subevent with that value as its conditional probability with respect to any prior in the set. Besides non-atomicity, Equidivisibility thus assumes a minimal degree of completeness of the likelihood relation. It is satisfied, for example, in the presence of a continuous (subjective) random device, as assumed (in objective form) in the widely-used Anscombe-Aumann framework In an important sense, Equidivisibility is not really restrictive at all since any coherent likelihood relation on any (possibly finite) state space can be extended to a larger one incorporating a hypothetical random-device on a larger state space. Other examples are given in section 3.

Uniqueness

Next to the provision of axiomatic foundations, a second main concern of the paper is to establish that comparative likelihood relations are adequately expressive as formal representations of "im-

[1] By contrast, there is a fair number of contributions on the complete case; see, for example, ## .

precise probabilistic beliefs" (understood in an in intuitive, not yet formally committed sense). A natural formal criterion for this is the *uniqueness* of the multi-prior representation (within the class of closed, convex sets of priors). Without uniqueness, a representation of imprecise beliefs by sets of priors could be viewed as more expressive than a representation in terms of comparative likelihood relations; this would cast doubt on the adequacy of comparative likelihood relations as "canonical" primitive.[2]

Fortunately, Equidivisibility ensures not only the existence of a multi-prior representation, but also its *uniqueness.* Indeed, since any extension of an equidivisible relation is equidivisible, it even guarantees the uniqueness of the multi-prior representation of all of its extensions.[3] In section 4, we investigate to what extent is it possible to weaken Equidivibility while still preserving uniqueness. On the positive side, we show that uniqueness continues to obtain when equidivisibility respectively convex-rangedness are only satisfied "arbitrarily closely". Specifically, the second major result of the paper, Theorem 3, establishes a one-to-one relation between "almost-equidivisible" likelihood relations and "almost-convex-ranged" sets of priors. On the other hand, we show by example that uniqueness is easily lost when the likelihood relation is not even almost-convex-ranged.

The difference between almost- and strict convex-rangedness is substantial. For example, it is frequently appropriate that all admissible priors on the realization of a real-valued random-variable have a uniformly continuous distribution, as advocated forcefully in an inspired recent paper by Machina (2001) on "Almost-Objective Uncertainty". In such cases, the set of priors will be almost but not necessarily strictly convex-ranged (Proposition 2). Machina (2001) captures the imprecise probabilistic belief in "uniform continuity in distribution" by a continuity assumption on preferences; we clarify its epistemic substance by explicitly representing this belief as a comparative likelihood relation (Proposition 3). Machina's notion of almost-objective uncertainty can thus be viewed as an important special case of our model of decision-making in the context of probabilistic beliefs. [4]

[2] Alternatively, if one insisted on comparative likelihood relations as "canonical" primitive, this would undermine the usefulness of the well-established, intuitively and mathematically compelling multiple-prior representation in that only equivalence classes of convex sets/ maximal convex sets associated with a given likelihood relation would be epistemically meaningful.

[3] That comparative likelihood relations can match multi-prior representations in their expressiveness at all in non-degenerate situations seems in fact fairly remarkable a priori; we are not aware of any hint of this in the literature.

[4] We have left to future research the task of characterizing coherent likelihood relations with an almost-convex-ranged representation. While this should not be impossible, we anticipate that it leads to a substantially more complicated, but probably not much more insightful proof than that of our result for the convex-ranged case, Theorem 2.

Applications to Decision Theory

In sum, we will show that, provided that comparative likelihood relations are sufficiently rich, they allow to represent imprecise probabilistic beliefs for the modelling of economic decisions in a behaviorally and epistemically general way. They also promise to be very fruitful in decision theory itself. From the mathematical point of view, this happens because convex-ranged beliefs effectively endow the event- and hence the act-spaces with a mixture-space structure. In particular, we show that if preferences over multi-valued acts are probabilistically sophisticated relative to a convex-ranged set of priors (in a sense defined in section 4.2), they can be represented within an Anscombe-Aumann framework. The analytical power of this framework is well-known, even though it is sometimes viewed with suspicion (see, e.g. Epstein (1999)). Our derivation not only clarifies the assumptions on preferences and beliefs implicit in the Anscombe-Aumann model, it leads to an even more powerful structure since all uncertainty is treated at the same level.[5]

In a companion paper (Nehring 2001), we have used the present framework to address three basic issues in the theory of decision making under ambiguity:

1. how to infer beliefs from preferences;
2. how to characterize decision-makers that depart from subjective expected utility only for reasons of ambiguity; and
3. how to define ambiguity attitudes in terms of betting preferences only to ensure behavioral generality.

In each case, the additional structure provided by a convex-ranged belief context is crucial.

Related Literature

1. Our first main result, Theorem 2, is most closely related to, and indeed builds on, the multiple-prior representations of partial orderings due to and Bewley (1986) and Walley (1991) following Smith (1961). All of these, however, use preferences as their primitive and derive the multiple-prior representation together with expected-utility maximization with respect to those priors, and thus

[5]Note that this holds for any convex-ranged belief context; the existence of a subjective random device is not assumed; as a counterpart of sorts to results in Machina (#), we also generalize this construction to almost-convex-ranged contexts.

fail to be behaviorally general.[6] Multiple-prior representations of complete preference orderings have obtained by Gilboa-Schmeidler (1989), GMM and K; again, these are about preferences, not belief, and behaviorally quite restrictive.

2. The published literature has not addressed the issue of *compatibility* of preferences with probabilistic beliefs explicitly, as far as we know. Implicitly, however, it offers proposals for the special case of unconditional probabilistic beliefs through primitive[7] definitions of "unambiguous events" revealed by the preference relation; see Epstein-Zhang (2001), Ghirardato-Marinacci (2001a) and Nehring (1999)). As shown in the companion paper Nehring (2001), the compatibility requirements derived from the extant definitions fail to adequately capture the "syntax of probability" and/or are behaviorally restrictive. The present paper originated in fact in an attempt at to overcome these limitations by defining "revealed unambiguous beliefs" as a coherent comparative likelihood relation that respects the logical syntax of probability by construction. This is worked out in Nehring (2001), where revealed unambiguous beliefs are defined as the maximal coherent likelihood relation extending a given convex-ranged belief context with which betting preferences are compatible.[8]

3. Machina (2001, 2002) formulates a model which reproduces the power of the Anscombe-Aumann framework in an enriched Savage setting, with the different but not unrelated goal of "robustifying" the classical (SEU) analysis of risk preferences and beliefs.[9] Indeed, Machina (2001) inspired our interest in almost-convex-ranged sets of priors; otherwise, his contribution arose independently of ours. Congenial with our work, Machina imposes epistemically motivated restrictions on preferences. However, in contrast to our work, these assumptions are imposed directly in the form of a continuity condition, while we model the postulated probabilistic belief explicitly as a likelihood relation, and obtain analogous preference restrictions via compatibility. Our approach can thus be viewed as

[6]In principle, one could reinterpret these results purely epistemically as partial orders over random variables, although this does not seem to have been advocated anywhere. Such an epistemic interpretation would run into two problems. On the one hand, the meaning of a comparison of random variables in terms of their expectation seems intuitively not very transparent; moreover, without assuming expected-utility maximization (at least relative to the specified ordering), their link to decision-making is not clear.

[7]That is: definitions that are not based on a prior notion of compatibility.

[8]The idea of defining "revealed unambiguous beliefs" in terms of a maximal independent subrelation had first been proposed in the talk Nehring (1996) in an Anscombe-Aumann framework. From the perspective of Nehring (2001), the older formulation corresponds to the special case of "utility sophisticated preferences" in which all departures from SEU maximization are due to ambiguity. Recently, this formulation has been taken up and developed further by Ghirardato-Maccheroni-Marinacci (2001c).

[9]A very different, utility- rather than event-based subjective version of the Anscombe-Aumann model has been proposed in the interesting recent paper by Ghirardato et al. (2001d).

generalizing and grounding Machina's. [10] Among many other things, Machina (2002) defines a global comparative likelihood relation from preferences under assumptions of event-differentiability of preferences. Nonetheless, this paper does not overlap with the present one, as neither it nor Machina 2001 analyze further the internal structure of comparative likelihood relation it derives.[11]

Overview

All proofs are contained in the appendix.

2. COHERENT LIKELIHOOD RELATIONS

A decision maker's probabilistic beliefs shall be modelled in terms of a partial ordering $\trianglerighteq$ on an algebra of events Σ in a state space Ω, his "comparative likelihood relation", with the instance $A \trianglerighteq B$ denoting the DM's judgment that A is at least as likely as B. We shall denote the symmetric component of $\trianglerighteq$ ("is as likely as") by $\equiv$. The comparative likelihood relation can be viewed as representing a non-exhaustive set of probabilistic judgments attributed to the DM , his *explicit probabilistic beliefs.* These judgments, in turn, may reflect probabilistic information available to and accepted by him, for example in the form of statistical frequencies or physical propensities. The likelihood relation may be "non-exhaustive" in the sense that the DM may have further probabilistic judgments not listed in $\trianglerighteq$. For now, we shall treat the comparative likelihood relation $\trianglerighteq$ as a non-behavioral primitive; we will consider its relation to behavior below in section 4.

2.1. Savage's Probability Theorem

As a reference point, we briefly review Savage's Probability Theorem delivers a unique representation of comparative likelihood relations in terms of finitely additive probabilities.

The following axioms are canonical for comparative likelihood in any context; disjoint union is denoted by "+".

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

Axiom 2 (Nondegeneracy) $\Omega \triangleright \emptyset$.

[10] There are further differences, for example in the treatment of probabilistic sophistication. In addition, Machina assumes event-differentiability of preferences which is not unrestrictive, excluding, for example, MEU and its cousins.

[11] The relation to Nehring (2001) will be discussed in future versions of that paper.

Axiom 3 (Nonnegativity) $A \trianglerighteq \emptyset$ *for all* $A \in \Sigma$.

Axiom 4 (Additivity) $A \trianglerighteq B$ *if and only if* $A + C \trianglerighteq B + C$ *, for any* C *such that* $A \cap C = B \cap C = \emptyset$.[12]

Additivity is the hallmark of comparative *likelihood*. Normatively, it can be read as saying that in comparing two events in terms of likelihood, only states that are not common can matter.

It is well-known that, on finite state-spaces, Additivity is far from sufficient to guarantee the existence of a probability-measure representing the complete comparative likelihood relation; see Kraft-Pratt-Seidenberg (1959). Savage (1954) realized, however, that Additivity suffices for the characterization of convex-ranged probability measures;[13] the probability measure π is **convex-ranged** if, for any event A and any $\alpha \in (0, 1)$, there exists an event $B \subseteq A$ such that $\pi(B) = \alpha\pi(A)$. We state a version of his result for the sake of further comparison. It requires two more axioms; the event A is *non-null* if $A \triangleright \emptyset$.

Axiom 5 (Fineness) *For any non-null* A *there exists a finite partition of* Ω $\{C_1, ..., C_n\}$ *such that for no* $i \leq n$, $A \triangleleft C_i$.

Axiom 6 (Tightness) *For any* A, B *such that* $B \triangleright A$ *there exist non-null events* C *and* D *such that* $B \backslash D \triangleright A \cup C$.

Theorem 1 (Savage) *Let* Σ *be a* σ*-algebra. The likelihood relation* $\trianglerighteq$ *satisfies Axioms 1 through 6 if and only if there exists a (unique) finitely additive, convex-ranged probability measure* π *on* Σ *such that for all* $A, B \in \Sigma$:

$$A \trianglerighteq B \text{ if and only if } \pi(A) \geq \pi(B).$$

An important feature of Savage's result is the uniqueness of the representing probability. It justifies the view that the comparative likelihood relation represents the DM's beliefs fully. This is non-trivial, and holds only rarely in finite state-spaces.

2.2 Dropping Completeness

To allow for imprecision in specified beliefs, likelihood relations will now allowed to be incomplete.

[12] In this notation, we quantify over all C disjoint from A and B.

[13] This result was in fact a crucial first step in his famous characterization of SEU maximization, Addivity of the "revealed likelihood relation" being a consequence of the Sure-Thing Principle.

Axiom 7 (Partial Order)[14] $\trianglerighteq$ *is transitive and reflexive.*

It is not immediately obvious how an incomplete likelihood relation is to represented in order to fully capture the logical syntax of probability. A natural minimal criterion of the latter is the possibility of extending the given incomplete relation $\trianglerighteq$ to a complete one that is representable by some subjective probability.

$$\text{There exists a finitely additve } \pi \text{ such that } \pi(A) \geq \pi(B) \text{ whenever } A \trianglerighteq B. \tag{1}$$

Likelihood relations satisfying (1) will be referred to as *non-contradictory,* and the associated probability measures π as *admissible.* Note that if one were to admit "contradictory" likelihood relations as representing well-defined probabilistic beliefs, one would in effect claim that there are probabilistic beliefs over some events/ event-comparisons that preclude in themselves the existence of precise probabilistic beliefs over some other events/ event-comparisons – a fairly radical claim.

While codition (1) rules out inconstistencies among likelihood judgments, it does not entail "deductive closure". For example, while it precludes the assertion of "$A^c \triangleright B^c$" given the judgment that "$A \trianglerighteq B$", it does not allow one to infer that "$B^c \trianglerighteq A^c$". Deductive closure is achieved by requiring that any absence of a comparative likelihood judgment can be rationalized by the existence of a prior satsifying (1) implying the contrary judgment, as stated by the following condition:

$$\text{For any } A, B \text{ such that not } A \trianglerighteq B, \text{ there exists } \pi \in \Delta(\Sigma) \text{ satisfying (1) such that } \pi(B) > \pi(A). \tag{2}$$

It is easily seen that this condition is equivalent to the existence of a set of finitely additive probability measures $\Pi \subseteq \Delta(\Omega)$ of the following form. For all $A, B \in \Sigma$:

$$A \trianglerighteq B \text{ if and only if } \pi(A) \geq \pi(B) \text{ for all } \pi \in \Pi. \tag{3}$$

A comparative likelihood relation with the representation (3) will be called **coherent**, and also referred to as an *imprecise qualitative probability.* Coherence entails deductive closure in the sense that, if a set of likelihood judgments entails another judgment "$C \trianglerighteq D$" assuming completeness, i.e. if $\pi(C) \geq \pi(D)$ for any π satisfying (1), the coherent relation $\trianglerighteq$ contains in fact this judgment. Note that if $\trianglerighteq$ satisfies (3) for some set of priors Π, then it satisfies (3) also for the convex hull of Π, as it does for the closure of Π (in the product or "weak*"-topology which will be assumed throughout). Thus, it is without loss of generality to assume Π to be a closed convex set; let the

[14] Technically, the proper label would be "preorder".

class of closed (hence compact), convex subsets of $\Delta(\Sigma)$ be denoted by $\mathcal{K}(\Delta(\Sigma))$. Some additional notation will prove convenient: for any set of priors $\Pi \in \Delta(\Sigma)$, let $\trianglerighteq_\Pi$ denote the likelihood relation induced by the unanimity condition (3); conversely, for any non-contradictory relation $\trianglerighteq$ on Σ, let $\Pi_\trianglerighteq$ denote the set of admissible priors π. Also, given Π, the lower and upper probabilities of an event $\min_{\pi\in\Pi} \pi(A)$ and $\max_{\pi\in\Pi} \pi(A)$ will be denoted by $\pi^-(A)$ and $\pi^+(A)$, respectively; given $\trianglerighteq$, the lower and upper probabilities are taken to be those associated with $\Pi_\trianglerighteq$.

A main achievement of Savage's Probability Theorem is its reliance on Additivity as the sole axiom capturing the logical syntax of probability. If the completeness assumption is dropped, this seems no longer feasible. For example, while under completeness, one can use Additivity to infer that if A is at least as likely than B, B's complement ("not B") must be at least as likely than that of A, this no longer follows without completeness. Yet this implication seems essential to a proper *likelihood* interpretation of the relation.

In the present context, it turns out to be sufficient to complement Additivity by the following second rationality axiom called "Splitting".

Axiom 8 (Splitting) *If* $A_1 + A_2 \trianglerighteq B_1 + B_2$, $A_1 \equiv A_2$ *and* $B_1 \equiv B_2$, *then* $A_i \trianglerighteq B_j$.

In words: If two events are split into two equally likely parts, then any part of the more likely event must be more likely than any part of the less likely event. Note that from additivity one can only infer that not $B_j \triangleright A_i$.

Splitting is made powerful with the help of the following assumption, according to which any event can be split into two equally likely parts.

Axiom 9 (Equidivisibility) *For any* $A \in \Sigma$, *there exists* $B \subseteq A$ *such that* $B \equiv A \backslash B$.

Note also Equidivisibility is consistent with the existence of "atoms" A of Σ (such as singletons), as long as these atoms are null-events.[15] For imprecise qualitative probabilities on σ-algebras, Equidivisibility is equivalent to the following condition of "convex-rangedness" of the representing set of priors; if Σ is merely an algebra, it is equivalent to "dyadic convex-rangedness". Let $\mathbf{D}$ denote the set of dyadic numbers between 0 and 1, i.e. of numbers of the form $\alpha = \frac{\ell}{2^k}$, where k and ℓ are non-negative integers such that ℓ does not exceed 2^k.

Definition 1 *A set of priors* Π *is* **convex-ranged** *if, for any event* $A \in \Sigma$ *and any* $\alpha \in (0, 1)$, *there*

[15] That is, for any A for which $B \subset A$ implies $B = \varnothing$, Equivdisibility implies $A \equiv \emptyset$.

exists an event $B \in \Sigma$, $B \subseteq A$ *such that* $\pi(B) = \alpha\pi(A)$ *for all* $\pi \in \Pi$. *The set* Π *is* **dyadically convex-ranged** *if this holds for all* $\alpha \in \mathbf{D}$.

Note that while range convexity of Π implies range convexity of every $\pi \in \Pi$, the converse is far from true.

Fact 1 *If* Σ *is a* σ*-algebra,* Π *is convex-ranged if and only if it is dyadically convex-ranged.*

Finally, Savage's Fineness and Tightness axioms are no longer adequate. To obtain a real-valued representation, a condition express the notion of "continuity in probability" is needed. It relies on the following notion of a "small", "$\frac{1}{K} -$ "event: A is a $\frac{1}{K}-$**event** if there exist $K-1$ mutually disjoint events A_i, disjoint from A, such that $A \trianglelefteq A_i$ for all i. Clearly, for coherent $\trianglerighteq$ and any $\pi \in \Pi$ and any $\frac{1}{K}-$event A, $\pi(A) \leq \frac{1}{K}$; if Π is convex-ranged, the converse holds as well.

Axiom 10 (Continuity) *If not* $A \trianglerighteq B$, *then there exists* $K < \infty$ *such that, for any* $\frac{1}{K}-$*events* C, D, *it is not the case that* $A \cup C \trianglerighteq B \backslash D$.

Note that Continuity is entailed by coherence. In particular, Continuity is applicable to any state space, finite or infinite.[16]

Theorem 2 *A relation* $\trianglerighteq$ *on an event algebra* Σ *has a multi-prior representation with a dyadically convex-ranged set of priors* Π *if and only if it satisfies Partial Order, Additivity, Nonnegativity, Splitting, Continuity, Equidivisibility, and Nondegeneracy. The representing* Π *is unique in* $\mathcal{K}(\Delta(\Sigma))$.

We shall sketch the idea of the proof of Theorem 2 with a bit of "reverse engineering". The key is the "mixture-space" structure of the event-space resulting from the convex-rangedness of the set of priors. Specifically, one can extend everycoherent likelihood relation represented by the convex-ranged set of priors Π to the domain $B_0(\Sigma, [0,1])$ of finite-valued functions $Z : \Omega \to [0,1]$ by associating with each Z an equivalence class $[Z]$ of events $A \in \Sigma$ as follows. Let $A \in [Z]$ if, for some

[16]By contrast, neither Tightness nor Fineness is entailed by coherence, or even the existence of a representing probability measure, as both rule out finite state spaces. On the other hand, both Fineness and Tightness are implied by Continuity plus Equidivisibility; both in effect mix non-atomicity and continuity aspects.

Moreover, in the presence of the other Savage axioms (including Fineness), Tightness can directly be shown to be equivalent to Continuity. Thus, in Theorem 1, one could replace Tightness by Continuity. This would have the conceptual advantage of having one condition (Continuity) entailed by the real-valuedness of the probability-representation, leaving Fineness as the condition solely responsible for the convex-rangedness of the representing measure.

appropriate partition of Ω $\{E_i\}$, $Z = \sum z_i 1_{E_i}$, and, for all $i \in I$ and $\pi \in \Pi : \pi(A \cap E_i) = z_i \pi(E_i)$. It is easily seen that for any two $A, B \in [Z] : \pi(A) = \pi(B)$ for all $\pi \in \Pi$, and thus $A \equiv B$. One therefore arrives at a well-defined partial ordering on $B_0(\Sigma, [0,1])$, denoted by $\widehat{\trianglerighteq}$, by setting

$$Y \widehat{\trianglerighteq} Z \text{ if } A \trianglerighteq B \text{ for some } A \in [Y] \text{ and } B \in [Z].$$

It is easily verified that this ordering is monotone, continuous and satisfies the following two conditions:

$$\textbf{(Additivity)} \quad Y \widehat{\trianglerighteq} Z \text{ if and only if } Y + X \widehat{\trianglerighteq} Z + X \text{ for any } X, Y, Z \;, \tag{4}$$

and

$$\textbf{(Homogeneity)} \quad Y \widehat{\trianglerighteq} Z \text{ if and only if } \alpha Y \widehat{\trianglerighteq} \alpha Z \text{ for any } Y, Z \text{ and } \alpha \in (0,1].$$

In the sequel, we shall refer to partial orderings on $B_0(\Sigma, [0,1])$ satisfying these four conditions as *coherent expectation orderings*. By well-known results due to Walley (1991) and Bewley (1986, for finite state-spaces), coherent expectation orderings admit a unique representation in terms of a closed, convex set of priors; cf. Theorem 4 in the appendix.

The actual proof of Theorem 2 proceeds by constructing this extension from the given likelihood relation and by deriving the properties of the induced relation from the axioms on the primitive relation. In particular, the Additivity and Homogeneity properties of the expectation ordering correspond to the Additivity and Splitting axioms satisfied by the underlying likelihood relation. The proof then invokes the just-quoted Theorem to obtain the desired (unique) multi-prior representation.

2.3. Examples of Equidivisibility

Equidivisibility is not a weak assumption. While it implies Fineness in the presence of Continuity, the converse is not close to being true, unless the likelihood relation is complete. While Fineness is in substance a strong non-atomicity condition, Equidivisibility makes positive assumptions about the existence of indifferences. Correspondingly, convex-rangedness assumes that the set of priors not be too large, or, more specifically, not too heterogeneous.

Besides this broad intuition motivating it, it is of interest to verify its content in the context of the following specific examples.

Example 1 (Limited Imprecision). One way to make the intuition of a limited extent of overall ambiguity precise is to assume that Σ is a σ-algebra and that Π is the convex hull of a *finite* set of *non-atomic, countably additive* priors. A well-known corollary of Lyapunov's (1940) convexity theorem due to Dubins and Spanier (1961) states, in effect, that Π is convex-ranged.

Convex-ranged contexts of this kind occur naturally in social belief aggregation, where $\trianglerighteq_I$ represents the unanimity likelihood ordering induced by the finite set of precise likelihood orderings $\trianglerighteq_i$ with representing measures μ_i. Assume that social decisions are based on a precise likelihood ordering $\trianglerighteq_I$ represented by some measure μ_I that respects unanimity in beliefs, i.e. such that $\trianglerighteq_I^* \supseteq \trianglerighteq_I$. Then Theorem 2 implies that $\Pi_{(\trianglerighteq_I)} = co\{\mu_i\}_{i \in I}$; the "social prior" μ_I must therefore be a convex combination of individual priors.[17]

Example 2 (Missing Information).

In some situations, ambiguity may only concern certain aspects of the state-space, and beliefs conditional on knowing these aspects may be precise. Formally, suppose that conditional on each event A in some finite finite partition $\mathcal{A}$ of Ω, the DM's beliefs are described by a convex-ranged probability measure μ_A; that is, for any $\pi \in \Pi$ and any $A \in \mathcal{A}$, $\pi(./A) = \mu_A$ or $\pi(A) = 0$. Then Π is convex-ranged, however imprecise the DM's beliefs about the events in $\mathcal{A}$ may be.

Example 3 (External Randomization Device)

In view of the above construction, a coherent likelihood relation can be viewed as inhabiting a mixture-space in the manner of Anscombe-Aumann without reference to an "extraneous" randomization device. On the other hand, state-spaces with a continuous randomization device furnish an important example of equidivisible likelihood relations. Specifically, consider a state space that can be written as $\Omega = \Omega_1 \times \Omega_2$, where the space Ω_1 is the space of "generic states" , and Ω_2 that of independent "random states" with associated algebras Σ_1 and σ-algebra Σ_2. The "continuity" and stochastic independence of the random device are captured by a coherent likelihood relation $\trianglerighteq_{AA}$ defined on the product algebra $\Sigma = \Sigma_1 \times \Sigma_2$ that satisfies the following two conditions, noting that any $A \in \Sigma_1 \times \Sigma_2$ can be written as $A = \sum_i S_i \times T_i$, where the $\{S_i\}$ form a finite partition of Ω_1.

AA1) The restriction of $\trianglerighteq_{AA}$ to $\{\Omega_1\} \times \Sigma_2$ satisfies all of Savage's axioms (axioms 1 through 6).

AA2) $\sum_i S_i \times T_i \trianglerighteq_{AA} \sum_i S_i \times T_i'$ if and only if, for all $i \in I$, $\qquad \Omega_1 \times T_i \trianglerighteq_{AA} \Omega_1 \times T_i'$.

While the first condition ensures the existence of a convex-ranged probability measure over random events $\overline{\pi}_2 \in \Delta(\Sigma_2)$, the second describes their stochastic independence.

[17] This is in effect a version of a result by Gilboa-Samet-Schmeidler (2001), who derive from social respect for unanimous indifferences a representation of the social prior as an affine linear combination of individual priors.

By AA1 and AA2, it is easily verified that $\trianglerighteq_{AA}$ satisfies all the assumptions of Theorem 2 including Equidivisibility. Hence there exists a unique set of priors Π_{AA} representing $\trianglerighteq_{AA}$; indeed, Π_{AA} is simply the set of all product-measures $\pi_1 \times \overline{\pi}_2$, where π_1 can be any finitely additive measure on Σ_1, reflecting the stochastic independence of the random device.[18]

The example of an external randomization device is particularly important in that it counters the potential initial impression that convex-rangedness is an empirically rather restrictive assumption, for it is possible to embed any coherent likelihood relation in any state-space in a larger likelihood relation on a larger state-space that incorporates the device.

3. WHEN IS THE MULTI-PRIOR REPRESENTATION UNIQUE?

Intuitively, different (closed and convex) set of probabilities convey different imprecise probabilistic beliefs. To flesh out this intuition in a decision-making context, consider a risk-neutral DM with linear utility over monetary outcomes, and think of random variables as monetary gambles. Then different sets of priors are associated with different partial preference orderings $\succsim_\Pi$ given by

$$X \succsim_\Pi Y \text{ iff } E_\pi X \geq E_\pi Y \text{ for all } \pi \in \Pi.$$[19]

Indeed, in general, these can easily be contradictory, in that a likelihood-relation can be represented by two disjoint sets Π, Π' such that, for some X, Y, $X \succ_\Pi Y$ and $Y \succ_{\Pi'} X$). On the other hand, since consequence-attitudes have been fixed, these difference in preference are naturally attributed to differences in beliefs. Thus, for comparative likelihood relations to comprehensively represent a DM's beliefs, their multi-prior representation must be unique.

Theorem 2 has already shown that Equidivisibility is sufficient for uniqueness. A priori, it seems indeed quite remarkable that uniqueness can be achieved at all in non-degenerate cases; and indeed, in finite settings this seems not to be possible. To see the source of the difficulty, note that any judgment $A \trianglerighteq B$ constrains any $\pi \in \Pi$ to satisfy the condition $\pi(A) \geq \pi(B)$, which may also be written as $\pi \cdot (1_A - 1_B) \geq 0$. In other words, comparative likelihood relations can separate convex sets Π only through vectors taking values in $\{-1, 0, +1\}$. It is therefore perhaps not surprising that

[18]Note that if the assumption of an external randomization device is accepted, Theorem 2 has not merely the implication of making the AA modeling kosher. Instead, its main importance is to ensure that any *further* restriction of the set of priors is captured one-to-one by a corresponding expansion of the likelihood relation.

[19]This follows from the uniqueness results due to Smith (1961), Bewley (1986), and Walley (1991); cf. Theorem 4 in the Appendix.

the uniqueness problem does not seem to have been adressed at all in the literature. Walley (1991, section 4.5), for instance, views comparative likelihood as merely one mode of assessing imprecise probabilities, if an especially simple and natural one.

Inspired by a construction of "almost-objective events" in Machina (2001), we show in this section that Equidivisibility/convex-rangedness can be weakened to "Almost Equidivisibility"/"almost convex-rangedness" (where the conditional probabilities fall within an arbitrarily small interval around the fixed number). While this weakening is fairly substantial, it continues to assume that the likelihood relation is sufficiently complete, in other words: that the imprecision is not all-pervasive. If, however, imprecision is pervasive in the sense that there are no events with (almost) precise conditional probabilities, uniqueness may fail to obtain easily, as we show by example.

3.1. Almost-Equidivisibility Implies Uniqueness

To generalize the one-to-one correspondence between equidivisible coherent likelihood relations and convex-ranged sets of priors, we formulate "approximate" generalizations.

Axiom 11 (Almost-Equidivisibility) *For all $E \in \Sigma$ and all $n \in \mathbf{N}$, there exists a partition of E into $2n-1$ sets $\{A_1, ..., A_{2n-1}\}$ such that, for any subfamily of n sets $\{A_{i_1}, ..., A_{i_n}\}$,*

$$\sum_{j=1,..,n} A_{i_j} \trianglerighteq E \backslash \left(\sum_{j=1,..,n} A_{i_j} \right).$$

Definition 2 *A set of priors Π is* **almost-convex-ranged** *if, for any event $A \in \Sigma$ and any $\alpha < \beta \in (0,1)$, there exists an event $B \in \Sigma$, $B \subseteq A$ such that $\alpha\pi(A) \leq \pi(B) \leq \beta\pi(A)$ for all $\pi \in \Pi$.*

By the following result, there is a one-to-one relation between almost-equidivisible likelihood relations and almost-convex-ranged sets of priors.

Theorem 3 *i) If $\trianglerighteq$ is coherent and almost-equidivisible, $\Pi_{\trianglerighteq}$ is almost-convex-ranged. Conversely, if Π is almost-convex-ranged, $\trianglerighteq_{\Pi}$ is almost-equidivisible.*

ii) Any coherent and almost-equidivisible comparative likelihood relation $\trianglerighteq$ has a unique multi-prior representation.

The proof of Theorem 3,ii) proceeds by showing through an approximate mixture-space construction that a coherent and almost-equidivisible likelihood relation $\trianglerighteq$ can be uniquely extended to an expectation ordering $\widehat{\trianglerighteq}$; since the multi-prior representation of such orderings is always unique, the multi-prior representation of $\trianglerighteq_{\Pi}$ is unique as well.

3.2 Example of Almost-Convex-Ranged Belief Sets: "Uniformly Continuous Densities"

In many situtations, random quantities are naturally assessed to be continuously distributed. For example, the subjective probability of any reasonable, contemporary human DM (assuming him to be Bayesian for now, for the sake of the argument) over the true temperature (idealized as real-valued magnitude) in Seoul on August 15, 2004, would be described by a continuous density function. Poincare (1912)[20] recognized that if one assumes a certain amount of additional regularity, then one can derive the existence of events with a chance of approximately one half, whatever the specific probability distribution of the DM within these constraints; in his important recent contribution by which this section was inspired, Machina (2001) refers to such events as "almost objective".

Specifically, suppose that $\Omega = [0, 1]$, and let A_k denote the event that the k-th digit in decimal expansion of ω is odd; thus A_k is the evenly spaced union of $\frac{1}{2}10^k$ intervals of length $\varepsilon_k := 10^{-k}$, and has Lebesgue measure $\frac{1}{2}$. Poincare obtained the following result:

Proposition 1 *For any differentiable density function ϕ such that $\mid \phi'(.) \mid \leq M$ on $[0, 1]$, $\mid prob(A_k) - \frac{1}{2} \mid \leq \frac{1}{2}M\varepsilon_k$.*

In the framework of the present paper, if Π denotes the set of all probability measures satisfying the assmption of Proposition 1, then $\frac{1}{2} - \frac{1}{2}M\varepsilon_k \leq \pi^-(A_k)$ and $\pi^+(A_k) \leq \frac{1}{2} + \frac{1}{2}M\varepsilon_k$. If one assumes in addition that densities are bounded below strictly above zero ($\phi(.) \geq L > 0$), then Π is in fact almost-convex-ranged, as we will show now. In fact, we will demonstrate a much more general result for metric spaces. Furthermore, we will represent explicitly the intutive notion of a "uniformly continuous density" of any reasonable prior in terms of an appropriate comparative likelihood relation.

Let Ω be a compact metric space endowed with a metric δ and a non-atomic, countably additive "reference-measure" λ on the Borel-σ-algebra Σ. If $\pi \in \Delta(\Sigma)$ is absolutely continuous with respect to λ, let ϕ denote its density (Radon-Nikodym derivative), and let $\Delta(\Sigma)_{cont}$ denote the set of all measures with continuous density. Let $\Psi : \Omega \times \Omega \to \mathbf{R}$ denote a function such that $\Psi(a, b)$ constrains how different the densities at a and b can reasonably be. Given λ and Ψ, define the set of compatible probability measures $\Pi_{\lambda,\Psi}$ as follows,

$$\Pi_{\lambda,\Psi} := \{\pi \in \Delta(\Sigma)_{cont} \mid \phi(a) \leq \phi(b)\Psi(a, b) \text{ for all } a, b \in \Omega\}. \quad (5)$$

[20] Quoted from Machina (2001).

In an interpretation alternative to one in terms of continuous quantities, the metric δ may capture a notion of similarity among states; "uniform continuity" of densities captures the intuition that similar states are assessed similarly under all admissible priors, possibly adjusted for differences in the size of the state. We will maintain the following assumptions on Ψ.

Assumption 1 *i)* $\Psi(a,b) \geq 1$ *for all* a, b.
ii) Ψ *is continuous, and* $\Psi(a,a) = 1$ *for all* a.
iii) For all $a, b, c \in \Omega : \Psi(a,c) \leq \Psi(a,b)\Psi(b,c)$.

Assumption i) ensures that the reference measure itself is compatible with the evidence, i.e. $\lambda \in \Pi_{\lambda,\Psi}$, while ii) entails that densities are continuous; since Ω is assumed compact, this implies also that Ψ is bounded above, and hence that densities are uniformly bounded above and below. Finally, iii) is w.l.o.g. in view of (5). Note that if Ψ is symmetric, then in view of i) and iii), $\log \Psi$ is a (quasi-)metric. Note also that setting $\Psi(a,b) := e^{\delta(a,b)M}$ is equivalent to requiring $\log \phi$ to be Lipschitz continuous of rank M.

Proposition 2 $\Pi_{\lambda,\Psi}$ *is almost-convex-ranged.*

We have chosen this somewhat more specific framework (rather than merely assuming uniform continuity and boundedness of the densities), in order to describe the associated likelihood relation more precisely. To do this, let $\Psi(A,B) := \sup_{a \in A, b \in B} \Psi(a,b)$, and define a comparative likelihood relation $\trianglerighteq_{\lambda,\Psi}$ as follows. For any A, B :

$$A \trianglerighteq_{\lambda,\Psi} B :\Leftrightarrow \quad \lambda(A) \geq \Psi(A,B)\lambda(B).$$

While $\trianglerighteq_{\lambda,\Psi}$ is not coherent, it is non-contradictory as λ is evidently compatible with $\trianglerighteq_{\lambda,\Psi}$. Let $\trianglerighteq^*_{\lambda,\Psi}$ denote the "coherent hull", i.e. the smallest coherent superrelation of $\trianglerighteq_{\lambda,\Psi}$; $\trianglerighteq^*_{\lambda,\Psi}$ adds to the "basic judgments" $\trianglerighteq_{\lambda,\Psi}$ further judgements inferred through the logic of probability. We will show that $\trianglerighteq_{\lambda,\Psi}$ respectively $\trianglerighteq^*_{\lambda,\Psi}$ fully encode the probabilistic beliefs described by $\Pi_{\lambda,\Psi}$ in a primitive way. Formally, one has

Proposition 3 $\Pi_{\lambda,\Psi}$ *is the unique multi-prior representation of* $\trianglerighteq^*_{\lambda,\Psi}$.

To show this, it is easy to verify $\trianglerighteq_{\lambda,\Psi} \subseteq \trianglerighteq_{(\Pi_{\lambda,\Psi})}$ and thus $\trianglerighteq^*_{\lambda,\Psi} \subseteq \trianglerighteq_{(\Pi_{\lambda,\Psi})}$, i.e. that $\trianglerighteq^*_{\lambda,\Psi}$ contains only comparative likelihood judgements induced by $\Pi_{\lambda,\Psi}$. It is not self-evident, and substantially more difficult prove that, conversely, any probability measure admissible with respect to $\trianglerighteq_{\lambda,\Psi}$ is in fact contained in $\Pi_{\lambda,\Psi}$. Together, this shows that $\Pi_{\lambda,\Psi} = \Pi_{(\trianglerighteq^*_{\lambda,\Psi})}$, hence that $\Pi_{\lambda,\Psi}$ is *a* multi-prior representation of $\trianglerighteq^*_{\lambda,\Psi}$. Uniqueness follows from the almost-convex-rangedness of $\Pi_{\lambda,\Psi}$ asserted in Proposition 2 in view of Theorem 3.

Proposition 2 goes beyond Machina (2001) by modelling the postulated "uniform continuity" probabilistic belief explicitly as a likelihood relation. It entails restrictions on preferences analogous his through the notion of Compatibility introduced in section 4. By contrast, Machina's formulation restricts preferences directly, and is based on the not unrestrictive assumption of event-differentiability. [21]

3.3 Example of Non-Uniqueness

We conclude by providing an example that shows that Equidivisibility/convex-rangedness cannot be weakened greatly without losing uniqueness. Let Σ denote the Borel-σ-algebra on the unit interval with Lebesgue-measure λ, and fix $K > 1$, and define a coherent likelihood relation $\trianglerighteq^K$as follows:

$$A \trianglerighteq^K B \text{ if and only if } \lambda(A \backslash B) \geq K\lambda(B \backslash A). \tag{6}$$

Note that this specification can be viewed as a special case of specification (5), with δ given as the discrete metric[22]; obviously, the unit interval fails to be compact under this metric.

It is easily verified that the associated set of admissible priors $\Pi^K := \Pi_{(\trianglerighteq^K)}$ consists of all probability measures π with Lebesgue-density ϕ such that $\operatorname{ess\,sup}_{\omega \in [0,1]} \phi(\omega) \leq K \operatorname{ess\,inf}_{\omega \in [0,1]} \phi(\omega)$; in particular, the extreme points of Π^K consist of all probability measures π_D with density ϕ_D, where D ranges over Σ with $0 < \lambda(D) < 1$, and ϕ_D is given by

$$\phi_D(\omega) = \begin{cases} \frac{K}{1+(K-1)\lambda(D)} & \text{if } \omega \in D, \\ \frac{1}{1+(K-1)\lambda(D)} & \text{if } \omega \notin D. \end{cases}$$

Let $\widetilde{\Pi}^K$ denote the closed, convex hull of $\{\pi_D | \lambda(D) = \frac{1}{K+1}\}$; in the Appendix, we show that $\widetilde{\Pi}^K$ induces the same likelihood relation $\trianglerighteq^K$. On the other hand, Π^K and $\widetilde{\Pi}^K$ induce different lower probabilities denoted by π_K^- and $\widetilde{\pi}_K^-$, respectively, and must therefore be different sets. These assertions are summarized by the following Fact.

Fact 2 *i)* $\trianglerighteq_{(\widetilde{\Pi}^K)} = \trianglerighteq^K$;

[21] This assumption excludes, for example, MEU and its cousins.

[22] With $\delta(x, y) = 1$ whenever $x \neq y$, and =0 otherwise.

ii) For all $A \in \Sigma : \pi^{-}_{K}(A) = \frac{\lambda(A)}{1+(1-\lambda(A))(K-1)}$;

iii) For all $A \in \Sigma : \widetilde{\pi}^{-}_{K}(A) = \begin{cases} \frac{K+1}{2K}\lambda(A) & \text{if } \lambda(A) \leq \frac{K}{K+1}, \\ 1 - \frac{K+1}{2}(1-\lambda(A)) & \text{if } \lambda(A) \geq \frac{K}{K+1}. \end{cases}$

Note that if K is close to $1,$ all admissible probabilities are uniformly close to the Lebesgue measure; nonetheless, uniqueness is lost.

Heuristically, this example suggests that "almost convex-rangedness" may be to close to being necessary for obtaining uniqueness through minimal completeness assumptions on the likelihood relation, but a formal result to this effect seems to be difficult to obtain. In any cae, almost-convex-rangedness is not strictly necessary for uniqueness. For example, in finite state spaces $\Omega = \{1, .., n\}$, the likelihood relation that judges all states equally likely has a unique represention by the singleton probability measures that assigns probability $\frac{1}{n}$ to each state. We leave a more exhaustive treatment of the uniqueness issue to future research.

4. PREFERENCES CONSTRAINED BY PROBABILISTIC BELIEFS

4.1. Compatibility of Betting Preferences with Probabilistic Beliefs

Consider now a DM described by a preference ordering over acts and salient beliefs over events. Let X be a set of *consequences*. An *act* is a finite-valued mapping from states to consequences, $f : \Omega \rightarrow X$, that is measurable with respect to a σ-algebra of events Σ; the set of all acts is denoted by $\mathcal{F}$. A *preference ordering* $\succsim$ is a weak order (complete and transitive relation) on $\mathcal{F}$. We shall write $[x_1, A_1; x_2, A_2; ...]$ for the act with consequence x_i in event A_i; for the act $[x, A; y, A^c]$ we will also use the shorthand $x_A y$. More generally, the act h that agrees with f on A and with g on A^c will be denoted by $f_A g$. As usual, constant acts $[x, \Omega]$ are typically referred to by their constant consequence x.

As part of the ground rules, we will assume throughout that preferences are state-independent.[23]

Axiom 12 (State Independence). *For all* $x, y \in X$, *all* $h \in \mathcal{F}$ *and all* $\succsim$*-non-null* $A \in \Sigma$, $x_A h \succsim y_A h$ *if and only if* $x \succsim y$.

The DM also has probabilistic beliefs described *non-exhaustively* by a coherent comparative likelihood relation $\trianglerighteq^0$ on Σ. This relation represents *some* of the DM's probabilistic beliefs; he may have

[23] State-Independence can be justified along Savagean lines by stipulating that consequences must be individuated so as to fully specify everthing evaluation-relevant.

others not included in it. The relation $\trianglerighteq^0$ will be referred to as the "*probabilistic context*" of the decision situation. Thus, decision-making in a probabilistic context is described by the pair $(\succsim, \trianglerighteq^0)$. In the companion paper Nehring (2001), we define a richer relation $\trianglerighteq^*$ derived from his preferences that captures *all* probabilistic beliefs that can be meaningfully described to the DM.

Probabilistic beliefs determine primarily preferences over bets. A *bet* is a pair of acts with the same two outcomes, i.e. a pair of the form $([x, A; y, A^c], [x, B; y, B^c])$. Fundamental is the following rationality requirement on the relation between preferences and probabilistic beliefs.

Axiom 13 (Compatibility) *For all $A, B \in \Sigma$ and $x, y \in X$ such that $x \succ y$:*

$$\begin{aligned} [x, A; y, A^c] &\succsim [x, B; y, B^c] \text{ if } A \trianglerighteq^0 B, \text{ and} \\ [x, A; y, A^c] &\succ [x, B; y, B^c] \text{ if } A \triangleright^0 B. \end{aligned}$$

Remark 1. If the relation $\trianglerighteq^0$ captures "objective probabilities", then Compatibility says in effect that behavioral probabilities (as embodied in betting preferences) must agree with objective ones whenever the latter are available. The philosopher David Lewis has termed this requirement the "Principal Principle".

Remark 2. Note that Compatibility (like State-Independence) needs to be justified by full individuation of consequences; it does not "follow" from the State Indepence axiom itself. For example, if consequences are described in terms of wealth, and if the marginal utility of wealth is uncertain but always strictly positive, State Independence will be satisfied but Compatibility may be rationally violated.[24]

Remark 3. Philosophically, Compatibility can be interpreted in two ways. On one view, Compatibility may be viewed as a normative requirement on the rational justification of preferences; such a requirement could be derived, for example, from the idea that likelihood comparisons, when available, should be *decisive* in ranking bets; other conceivable factors such as familiarity or "competence" should not matter rationally.

Alternatively, from a behavioral point of view broadly conceived, Compatibility is naturally taken to be part and parcel of the very meaning of likelihood judgments. If someone says that she believes event A more likely than event B but insists on choosing to bet on B over betting on A, it is unclear what to make of her claim that nonetheless she believes A to be more likely than B. On this "*semantic*" interpretation, we would have to infer that the agent uses "more likely than" in a

[24] In effect, Compatibility presupposes that the utility difference among any pair of acts is constant across states, while State Independence asserts only constancy of the ordinal ranking of the two acts.

non-standard way.

The semantic interpretation of Compatibility renders belief attributions *behaviorally falsifiable*: if the DM prefers betting on B over betting on A, one cannot attribute to her the belief that A is at least as likely as B. On the other hand, as shown by the Ellsberg paradox, it may not be legitimate to attribute to her the converse belief that B is more likely than A, since his betting preferences may be influenced by his ambiguity attitudes. Thus, this semantic interpretation does not commit to the "behaviorist" claim that beliefs must be *definable* from preferences.

It is not the purpose of this paper to settle these philosophical issues, even though they are undoubtedly important. Suffice it say that the semantic interpretation should appeal on reflection to most readers raised in the Ramsey-Savage tradition; note in particular that if the underlying likelihood relation is assumed complete (as is arguably implicit in Savage's axiom P4), Compatibility ensures the unique recoverability of likelihood judgments from preferences, and provides thus an alternative way of reading Savage's behavioral "definition" of comparative likelihood.

4.2 Constraints on Preferences over Multi-Valued Acts

It is natural to ask whether probabilistic beliefs impose further constraints on preferences. Now, it is clear that any constraints on preferences also exploit information about the valuation of consequences; Compatibilty, for example, derives betting preferences from beliefs and preferences over constant acts. In this section, we shall consider natural constraints over multi-act preferences that can be formulated on the basis of such information. Much stronger normative restrictions can be obtained if one exploits information about comparisons of utility differences. For example, if the utility gain from y over z is judged to exceed that from x over y, and if A is equally likely to its complement, it can be claimed that the constant act y should be chosen over the act $x_A z$ which results in x or z with 50% probability. Such a theory of "Expected Utility in the Presence of Ambiguity" is developed in Nehring (2001, section 4). Whether or not such implications are indeed normatively compelling, it is clear that they are substantive normative requirements; their violation does not undermine the very meaning ascribing the relevant probabilistic beliefs (here: the judgment that A is equally likely to its complement). For example, while someone displaying the Allais pattern of preferences might be criticized for not appreciating the normative force of the sure-thing principle properly, few would claim that one must conclude that the agent's beliefs differ from the specified probabilities or, indeed, that he does not have any well-defined beliefs at all.

If one is restricted to exploiting only information about the ordinal ranking of consequences, a natural further rationality requirement is respect for stochastic dominance, to the extent that this criterion can be applied on the basis of the DM's specified probabilistic beliefs. Say that f **stochastically dominates** g if, for all $x \in X$, $\{\theta \mid f(\theta) \succsim x\} \trianglerighteq^0 \{\theta \mid g(\theta) \succsim x\}$.

Axiom 14 (Generalized Stochastic Dominance, GSD) *$f \succsim g$ whenever f stochastically dominates g.*

The following Fact follows immediately from the well-known utility characterization of ordinary stochastic dominance. It suggests that Generalize Stochastic Dominance is the strongest rationality requirement that relies on ordinal information about the valuation of consequences only.

Fact 3 *f stochastically dominates g if and only if $E_\pi u \circ f \geq E_\pi u \circ g$ for all $\pi \in \Pi^0$ and all $u : X \to \mathbf{R}$ such that $u(x) \geq u(y)$ whenever $x \succsim y$.*

A somewhat weaker condition that generalizes the notion of "probabilistic sophistication" due to Machina-Schmeidler (1992) to imprecise probabilities can be defined similarly. Say that f is **stochastically equivalent** to g if, for all $x \in X$, $\{\theta \mid f(\theta) \succsim x\} \equiv^0 \{\theta \mid g(\theta) \succsim x\}$. Clearly, f is stochastically equivalent to g if, for all $\pi \in \Pi^0$, the distribution of consequences under f is the same as that under g.

Axiom 15 (Generalized Probabilistic Sophistaction, GPS) *$f \succsim g$ whenever f is stochastically equivalent to g.*

Evidently, if $\trianglerighteq^0$ is complete / Π^0 is a singleton, GPS amounts to "probabilistic sophistication" a la Machina-Schmeidler (1992). While in this case GPS and GSD are equivalent given State-Independence, they are not when $\trianglerighteq^0$ is merely convex-ranged. [25]

4.3 A Subjective Interpretation of the Anscombe-Aumann Framework

The mixture-space construction of sections 2 and 3 will now be combined with GPS to obtain a subjective interpretation of the Anscombe-Aumann (1963) framework. While this framework is often used in the analysis of decision making under ambiguity, it is generally viewed as theoretically less fundamental and transparent than the Savage framework; sometimes it is even viewed with

[25] Grant (1995) characterizes probabilistic sophistication when State-Independence, and hence monotonicity with respect to stochastic dominance, fails.

outright suspicion (see, e.g., Epstein (1999)). We will first consider the simpler convex-ranged case, and then offer a generalization to the almost-convex-ranged case inspired by Machina (2001).

4.3.1 The Convex-Ranged Case.—

The Anscombe-Aumann (1963) framework is distinguished by taking acts to be mappings from states to probability distributions of consequences, rather than directly to consequences as in the Savage (1954) framework. These probability distributions are interpreted as objective probabilities of the outcome of a random device ("roulette lotteries") that is part of the explicitly modeled state space. We will show now Compatibility with a convex-ranged context of probabilistic beliefs and GPS deliver an extension of preferences over Savage acts to to Anscombe-Aumann acts *internally*, that is: without the (further) addition of an external random device.

Formally, an AA-act F is a finite-valued Σ-measurable mapping from the state space Ω to the set of probability distributions on X with finite support $\Delta(X)$. Let $\mathcal{F}^{AA}$ denote their set. Denoting elements of $\Delta(X)$ by $q = (q^x)_{x \in X}$, one can write $F = [q_1, F_1; q_2, F_2; ...]$ in analogy to the notation for Savage acts. Given any convex-ranged probabilistic context $\trianglerighteq$, any AA-act F can be identified with a class $[F]$ of Savage acts by the following stipulation: $f \in [F]$ if, for any $x \in X$ and any i,

$$\pi^-(\{\omega \in F_i | f(\omega) = x\}/F_i) = \pi^+(\{\omega \in F_i | f(\omega) = x\}/F_i) = q_i^x.$$

Thus $[F]$ consists of all Savage acts that yield the consequence probabilities specified by F as unambiguous conditional context-probabilities. Convex-rangedness of the context ensures that $[F]$ is non-empty. On the other hand, any two acts in $[F]$ are easily seen to be stochastically equivalent with respect to the context $\trianglerighteq$. Hence GSP ensures that any two acts in $[F]$ are indifferent. Thus one obtains a well-defined weak order on $\mathcal{F}^{AA}$ by setting

$$F \succsim^{AA} G :\Leftrightarrow f \succsim g \text{ for any } f \in [F] \text{ and } g \in [G].$$

Moreover, since Savage acts embed in $\mathcal{F}^{AA}$ as deterministic AA-acts, $\succsim^{AA}$ contains exactly the same information as the original preference ordering $\succsim$, except that acts are is now replicated in multiple copies.

This achieves a justification of the AA framework in a subjective, epistemically enriched setting analogous to the (purely behavioral) justification of the von Neumann-Morgenstern framework by Machina-Schmeidler (1992). It implies that any assumption on AA preferences can be translated in principle into an assumption on the underlying Savage preferences; it does not imply, however, that the Savage counterpart has an obvious interpretation. For example, State Independence of

AA preferences is much stronger an assumption than State Independence of Savage preferences. Remarkably, our justification shows that the AA framework can be interpreted without reference to an external randomization device.[26]

Earlier representations of the Anscombe-Aumann framework in a Savage setting have been presented in Pratt-Raiffa-Schleifer (1964) and Klibanoff (2001a); in contrast to ours, the former entails expected utility maximization, the latter utility sophistication. An altogether different route to mimicking the Anscombe-Aumann framework in a subjective setting based on a rich set of consequences rather than states is proposed by Ghirardato et al. (2001d); since the mixture operation in their proposal is defined in utility terms, the subjectivized interpretation of conditions on AA preferences may be very different from that in the other, probability based approaches.

4.3.2 Generalization to Almost-Convex-Ranged Contexts.—

If the context $\trianglerighteq$ is merely almost-convex-ranged, the above construction of AA acts and preferences fails because AA acts may have no Savage counterpart, that is: $[F]$ will be empty for some F. The natural remedy is to interpret AA acts $F = [q_i, F_i]_{i \in I}$ as appropriate limits of sequences of Savage acts $\{f_n\}$ with the property that, for each i, the (imprecise) conditional distribution of consequences conditional on the event F_i induced by f_n converges to that specified by the AA act, q_i. To yield a well-defined preference ordering on AA acts, the underlying preferences over Savage acts must be continuous in an appropiate sense. We will now make this construction formally precise.

Let d denote the sup-metric on $\Delta(X)$, $d(q, q') := \sup_{x \in X} \mid q(x) - q(x') \mid$. To define "convergence in distribution" of an sequence of Savage acts to an AA act, it is helpful to define the following distance measure on $\mathcal{F} \times \mathcal{F}^{AA}$ δ'.

$$\delta'(f, F) := \sup_{i \in I, \pi \in \Pi} d\left(\pi(./F_i) \circ f^{-1}, q_i\right),$$

where $\pi(./F_i) \circ f^{-1} \in \Delta(X)$ is given by $(\pi(./F_i) \circ f^{-1})(x) = \pi\left(\{\omega | f(\omega) = x\}/F_i\right)$ for any x.

An AA act F can now formally be defined as the set of all sequences of Savage acts $\{f_n\}$ converging to F, i.e.

$$\{f_n\} \in [F] \text{ if } \lim_{n \to \infty} \delta'(f_n, F) = 0.$$

It is easily verified that Almost-Convex-Rangedness ensures the non-emptiness of $[F]$ for all $F \in \mathcal{F}^{AA}$.

[26] If the probabilistic context is given by $\trianglerighteq_{AA}$ as defined in section 3 capturing an external randomization device, Anscombe-Aumann's "horse lotteries" are represented by the Σ_1-measurable acts in $\mathcal{F}^{AA}$.

Preferences over AA acts construed in this manner is are defined naturally by continuous extension.

$$F \succsim^{AA} G \text{ if there exist } \{f_n\} \in [F] \text{ and } \{g_n\} \in [G] \text{ such that } f_n \succsim g_n \text{ for all } n.$$

For the associated strict preference relation $\succ^{AA}$ defined as the asymmetric component of $\succsim^{AA}$, this amounts to eventual preference for any pair of approaching sequences.

$$F \succ^{AA} G \text{ if, for all } \{f_n\} \in [F] \text{ and } \{g_n\} \in [G],\ f_n \succ g_n \text{ for sufficiently large } n.$$

To make this ordering well-behaved, the underlying preference relation $\succsim$ must be continuous in an appropriate sense. To define such a notion of continuity, define the following quasi-metric on Savage acts δ.

$$\delta(f,g) := \sup_{\pi \in \Pi} d(\pi \circ f^{-1}, \pi \circ g^{-1}).$$

The quasi-metric δ ($=\delta_\Pi$) defines an upper bound on how far the probability distributions over consequences may be apart. Note that $\delta(f,g) = 0$ if and only if $\pi \circ f^{-1} = \pi \circ g^{-1}$ for all $\pi \in \Pi$, i.e. f is stochastically equvialent to g. Thus, intuitively, if $\delta(f,g)$ is small, then f and g are "almost" stochastically equivalent. The following continuity condition requires that almost stochastically equivalent are evaluated similarly by the agent.

Axiom 16 (Continuity in Distribution)

The weak order $\succsim$ has a utility-representation $V : \mathcal{F} \to \mathbf{R}$ that is uniformly continuous with respect to δ_Π.

Proposition 4 *If $\succsim$ satisfies Continuity in Distribution and Generalized Stochastic Domimance with respect to an almost-convex-ranged context $\trianglerighteq$, $\succsim^{AA}$ is a continuous[27] weak order extending $\succsim$.*

[27] with respect to the standard metric $\delta''(F,G) := \sup_{\omega \in \Omega} d\left(F(\omega), G(\omega)\right)$.

APPENDIX: PROOFS.

Proof of Fact 1.

Take any real number $\alpha \in (0,1)$ and any $A \in \Sigma$. Write α as the supremum of an increasing sequence of dyadic numbers $\{\frac{\ell_k}{2^k}\}_{k=1,..,\infty}$ such that

$$\frac{\ell_k+1}{2^k} \geq \alpha. \tag{7}$$

By dyadic convex-rangedness, there exists a sequence of partitions $\{\mathcal{A}_k\}$ such that $\mathcal{A}_k$ is a refinement of $\mathcal{A}_{k'}$ whenever $k \geq k'$ (i.e. $\{\mathcal{A}_k\}$ is a filtration), and such that $\pi(A') = \frac{1}{2^k}\pi(A)$ for all $\pi \in \Pi$ and all $A' \in \mathcal{A}_k$.

Thus there exists an increasing sequence $\{B_k\}$, where each B_k is the union of ℓ_k members of $\mathcal{A}_k$ and, taking account of (7), a decreasing sequence $\{D_k\}$ of members of $\mathcal{A}_k$ such that

$$B_{k'} \subseteq B_k \cup D_k \tag{8}$$

whenever $k' \geq k$.

Since Σ is a $\sigma-$algebra, $B := \cup B_k \in \Sigma$. We claim that B is the desired event. Indeed, from the construction of the sequence B_k, it follows immediately that $\pi(B) \geq \alpha\pi(A)$ for any $\pi \in \Pi$. Conversely, by (8), $B \subseteq B_k \cup D_k$ for all k, and thus, for any $\pi \in \Pi$, $\pi(B) \leq \pi(B_k) + \pi(D_k) \leq \alpha + \frac{1}{2^k}$ for all k, whence $\pi(B) \leq \alpha$. $\square$

Proof of Theorem 2.

Let E be any non-null event in Σ, and $\alpha = \frac{\ell}{2^k}$ be any dyadic number. We begin by defining, from likelihood judgments, a family αE of events A that will, in the multi-prior representation have the property that, for all $\pi \in \Pi$, $\pi(A) = \alpha\pi(E)$. Specifically, let αE be the set of all A such that there exists a partition of E into 2^k subsets A_i such that $A_i \equiv A_j$ for all i, j and $A = \sum_{i \leq \ell} A_i$.

We have the following lemmas.

Lemma 1 *(Strong Additivity) $A \trianglerighteq B$ and $A' \trianglerighteq B'$ implies $A + A' \trianglerighteq B + B'$.*

This Lemma is standard in derivations of Savage's Theorem; see, e.g. Fishburn (1970, p. 196). Its proof is therefore omitted.[28]

Lemma 2 *$A \in \frac{1}{2^k}E$ if and only if there exists $E' \in \frac{1}{2^{k-1}}E$ such that $A \in \frac{1}{2}E'$.*

[28] Fishburn's proof is for $\triangleright$ and $\equiv$, but applies equally to $\trianglerighteq$; it is applicable since it does not make use of completeness.

The "only-if" part follows directly from Strong Additivity.

The "if-part" holds trivially for $k = 1$. For $k > 1$, it is verified by induction. Suppose it to hold for $k' = k - 1$. Assume that there exists $E' \in \frac{1}{2^{k-1}}E$ such that $A \in \frac{1}{2}E'$. Then by the definition of $\frac{1}{2^{k-1}}E$, there exists a partition of E into events $\{E_1, ..., E_{2^{k-1}}\}$ such that $E_i \equiv E_j$ for all i, j and $E_1 = E'$. By Equidivisibility, for each $i \geq 1$, there exist events $E_{i,1}$ and $E_{i,2}$ such that $E_{i,1} \equiv E_{i,2}$, $E_{i,1} + E_{i,2} = E_i$ and $E_{1,1} = A$. By Splitting, $E_{i,m} \equiv E_{j,m'}$, and thus $A \in \frac{1}{2^k}E$.

Lemma 3 *$\alpha E \neq \emptyset$ for all $\alpha \in \mathbf{D}$ and all non-null E.*

By Equidivisibility and induction on k, the claim follows for $\alpha = \frac{1}{2^k}$ from Lemma 2, hence indeed for all $\alpha = \frac{\ell}{2^k}$ by the definition of αE.

Lemma 4 *$A \in \alpha C$, $B \in \beta D$, $\alpha \geq \beta$ and $C \trianglerighteq D$ imply $A \trianglerighteq B$.*

- seems that $\alpha = \beta$ may be enough.

From an argument as in Lemma 2, it is clear that, writing $\alpha = \frac{\ell}{2^k}$ and $\beta = \frac{\ell'}{2^k}$ with $\ell \geq \ell'$, there exist partititions of E into 2^k elements $E = \sum_{i \leq 2^k} A_i$ and $E = \sum_{i \leq 2^k} B_i$ such that $A = \sum_{i \leq \ell} A_i$ and $B = \sum_{i \leq \ell'} B_i$. First, consider the case $\ell = \ell' = 1$. Then the claim follows from Splitting and induction on k. In the general case with $\ell \geq \ell'$, this implies $A_i \equiv B_i$ for all $i \leq 2^k$, whence $A \trianglerighteq B$ by repeated application of Strong Additivity.

We are now in a position to construct the mixture- space extension $\widehat{\trianglerighteq}$ of $\trianglerighteq$. Let $\mathcal{D}$ denote the set of dyadic-valued random-variables, $\mathcal{D} := \{Z : \Omega \to \mathbf{D},\ Z \text{ is } \Sigma\text{-measurable}\}$. Any finite-valued Z can be canonically written as $\sum_i z_i 1_{E_i}$, where $E_i = Z^{-1}(\{z_i\})$. For any $Z = \sum z_i 1_{E_i} \in \mathcal{D}$, define

$$[Z] := \{A : \text{ there exist } A_i \in z_i E_i \text{ such that } A = \sum_i A_i\},$$

and define the relation $\widehat{\trianglerighteq}$ on $\mathcal{D}$ as follows,

$$X \widehat{\trianglerighteq} Y \text{ iff, for some } A \in [X] \text{ and } B \in [Y],\ A \trianglerighteq B.$$

To establish various properties of $\widehat{\trianglerighteq}$, some further auxiliary results are needed.

Lemma 5 *For all $A, B \in [Z] : A \equiv B$.*

By definition, $A = \sum_i A_i$ and $B = \sum_i B_i$ such that $A_i, B_i \in z_i E_i$. By Lemma 4, $A_i \equiv B_i$. Hence $A \equiv B$ by Strong Additivity.

Lemma 6 *If $A_i \in \alpha E_i$ for all $i \in I$, $\sum_{i\in I} A_i \in \alpha\left(\sum_{i\in I} E_i\right)$.*

Writing $\alpha = \frac{\ell}{2^k}$, by assumption there exist sets B_{ij} for $i \in I$ and $j \leq 2^k$ such that $B_{ij} \equiv B_{ij'}$ for all i, j, j', $\sum_{j\leq 2^k} B_{ij} = E_i$ for all i, and $\sum_{j\leq \ell} B_{ij} = A_i$. For $j \leq 2^k$, let $B_j := \sum_{i\in I} B_{ij}$. By construction, $\sum_{i\in I} E_i = \sum_{i\in I}\sum_{j\leq 2^k} B_{ij} = \sum_{j\leq 2^k} B_j$. By Strong Additivity, $B_j \equiv B_{j'}$ for all j, j'. Since $\sum_{i\in I} A_i = \sum_{i\in I}\sum_{j\leq \ell} B_{ij} = \sum_{j\leq \ell} B_j$, therefore $\sum_{i\in I} A_i \in \frac{\ell}{2^k}\left(\sum_{i\in I} E_i\right)$.

Lemma 7 *i) For all $X, Y, Z \in \mathcal{D}$ such that $X + Z \in \mathcal{D}$ and $Y + Z \in \mathcal{D}$, there exist $A \in [X]$, $B \in [Y]$ and $C \in [Z]$ such that $A + C \in [X + Z]$ and $B + C \in [Y + Z]$.*

ii) For all $X, Y \in \mathcal{D}$ such that $X + Y \in \mathcal{D}$ and such that Y is measurable w.r.t. the partition generated by X, and for all $A \in [X]$, there exists $B \in [Y]$ such that $A + B \in [X + Y]$.

iii) For all $X, Y \in \mathcal{D}$ such that $X + Y \in \mathcal{D}$ and such that Y is measurable w.r.t. the partition generated by $X + Y$, and for all $C \in [X + Y]$, there exists $B \in [Y]$ such that $B \subseteq C$ and $C\backslash B \in [X]$.

To verify part i), write X, Y and Z (non-canonically) as $X = \sum_i x_i 1_{D_i}$,$Y = \sum_i y_i 1_{D_i}$ and $Z = \sum_i z_i 1_{D_i}$ for an appropriate partition $\{D_i\}$ of Ω, and write $x_i = \frac{\ell_i}{2^{k_i}}, y_i = \frac{\ell_i'}{2^{k_i}}$, and $z_i = \frac{\ell_i''}{2^{k_i}}$. Split D_i into 2^{k_i} equally likely events $\{D_{i1},..., D_{i2^{k_i}}\}$, and set $C_i := \sum_{j\leq \ell_i} D_{ij} \in z_i D_i$, $A_i = \sum_{j=\ell_i+1}^{\ell_i+\ell_i'} D_{ij} \in x_i D_i$, and $B_i = \sum_{j=\ell_i+1}^{\ell_i+\ell_i''} D_{ij} \in y_i D_i$. Using Lemma 6, one infers that $\sum_i A_i \in [X]$, $\sum_i B_i \in [Y]$, $\sum_i C_i \in [Z]$, $\sum_i A_i + \sum_i C_i = \sum_i (A_i + C_i) \in [X + Z]$, and $\sum_i B_i + \sum_i C_i = \sum_i (B_i + C_i) \in [Y + Z]$ as desired.

Similar proofs verify parts ii) and iii). As to the former, write $X = \sum_i x_i 1_{E_i}$ in canonical decomposition. By assumption, Y can be written (non-canonically) as $\sum_i y_i 1_{E_i}$. Take any $A = \sum_i A_i \in [X]$. Since $x_i + y_i \leq 1$ for all i, one can find $B_i \in y_i E_i$ such that $A_i + B_i \in (x_i + y_i) E_i$. Using Lemma 6, one infers that $\sum_i B_i \in [Y]$, as well as $A + \sum_i B_i = \sum_i (A_i + B_i) \in [X + Y]$, as desired.

Finally, to verify part iii), write $X + Y = \sum_i z_i 1_{E_i}$ in canonical decomposition. By assumption, Y can be written (non-canonically) as $\sum_i y_i 1_{E_i}$. Take any $C = \sum_i C_i \in [X + Y]$. Since $y_i \leq z_i$ for all i, one can find $B_i \in y_i E_i$ such that $C_i \backslash B_i \in (z_i - y_i) E_i$. Using Lemma 6, one infers that $\sum_i B_i \in [Y]$, as well as $C\backslash\left(\sum_i B_i\right) = \sum_i (C_i \backslash B_i) \in [X]$, as desired. □

Lemma 8 *The relation $\widehat{\trianglerighteq}$ on $\mathcal{D}$ is transitive, reflexive and satisfies the following conditions*

1. *(Extension)* $1_A \widehat{\trianglerighteq} 1_B$ *if and only if* $A \trianglerighteq B$.
2. *(Positivity)* $X \widehat{\trianglerighteq} \mathbf{0}$ *for all* X.

3. *(Non-degeneracy)* $\mathbf{1}\widehat{\rhd}\mathbf{0}$.

4. *(Weak Homogeneity)* $X\widehat{\trianglerighteq}Y$ *implies* $\alpha X\widehat{\trianglerighteq}\alpha Y$ *for all* $\alpha \in \mathbf{D}$.

5. *(Additivity)* $X\widehat{\trianglerighteq}Y$ *if and only if* $X+Z\widehat{\trianglerighteq}Y+Z$.

6. *(Strong Additivity)* $X\widehat{\trianglerighteq}Y$ *and* $X'\widehat{\trianglerighteq}Y'$ *imply* $X+X'\widehat{\trianglerighteq}Y+Y'$.

7. *(Continuity)* $\{(X,Y) : X\widehat{\trianglerighteq}Y\}$ *is closed (in* $\mathcal{D}\times\mathcal{D}$*).*

Proof. Reflexivity, Extension, Positivity, and Non-degeneracy are immediate.

To verify Transitivity, consider any X, Y, Z such that $X\widehat{\trianglerighteq}Y$ and $Y\widehat{\trianglerighteq}Z$. By definition, there exist $A \in [X], B, B' \in [Y], C \in [Z]$ such that $A \trianglerighteq B$ and $B' \trianglerighteq C$. By Lemma 5, $B \equiv B'$. Hence by the transitivity of $\trianglerighteq$, $A \trianglerighteq C$, and therefore $X\widehat{\trianglerighteq}Z$ as desired.

Weak Homogeneity is an immediate consequence of Lemmas 3 and 4.

To verify Additivity, consider any X, Y, Z such that $X+Z, Y+Z \in \mathcal{D}$. According Lemma 7i), there exist $A \in [X]$, $B \in [Y]$ and $C \in [Z]$ such that $A+C \in [X+Z]$ and $B+C \in [Y+Z]$. If $X\widehat{\trianglerighteq}Y$,then $A \trianglerighteq B$ by Lemma 5, thus $A+C \trianglerighteq B+C$ by Additivity of $\trianglerighteq$, and thus $X+Z\widehat{\trianglerighteq}Y+Z$. Analogously, one obtains $X\widehat{\trianglerighteq}Y$ from $X+Z\widehat{\trianglerighteq}Y+Z$.

Strong Additivity, in turn, follows straightforwardly from 7i) and the Strong Additivity of $\trianglerighteq$.

It remains to verify Continuity. We shall show that $\{(X,Y) : \text{not } X\widehat{\trianglerighteq}Y\}$ is open in $\mathcal{D}$. Consider any X, Y such that not $X\widehat{\trianglerighteq}Y$. Take any $A \in [X], B \in [Y]$; clearly not $A \trianglerighteq B$. By the Continuity of $\trianglerighteq$, there exists $K < \infty$ such that, for any $\frac{1}{K}-$events C, D, it is not the case that $A \cup C \trianglerighteq B \backslash D$. It suffices to show that, for any X', Y' such that $\parallel X'-X \parallel \leq \frac{1}{K}$ and $\parallel Y'-Y \parallel \leq \frac{1}{K}$, it is not the case that $X'\widehat{\trianglerighteq}Y'$.

To verify this claim, take any X', Y' such that $\parallel X'-X \parallel \leq \frac{1}{K}$ and $\parallel Y'-Y \parallel \leq \frac{1}{K}$. By the Positivity and Strong Additivity of $\trianglerighteq$, it is without loss of generality to assume that X' (respectively Y') is measurable with respect to the partition generated by X (respectively Y), and that $X' \geq X$ and $Y' \leq Y$. Then there exist by Lemma 7ii) $A' \in [X'-X]$ such that $A+A' \in [X']$; likewise, by Lemma 7iii), there exist and $B' \in [Y-Y']$ and $B'' \in [Y']$ such that $B'+B'' = B$. Clearly, A' and B' are $\frac{1}{K}-$events, and therefore it is not the case that $A+A' \trianglerighteq B\backslash B' = B''$. Therefore, in view of Lemma 5, it is not the case that $X'\widehat{\trianglerighteq}Y'$, as needed to be shown. □

Now embed $\widehat{\trianglerighteq}$ (viewed as a subset of $\mathcal{D}\times\mathcal{D}$) in $\mathcal{B}\times\mathcal{B}$, with $\mathcal{B} := B(\Sigma, [0,1])$, the set of $[0,1]-$valued Σ-measurable functions, endowed with the sup-norm. Since $\mathcal{B}$ is the completion of $\mathcal{D}$, and thus $\mathcal{B}\times\mathcal{B}$

of $\mathcal{D} \times \mathcal{D}$, the closure $cl\widehat{\trianglerighteq}$ of $\widehat{\trianglerighteq}$ in $\mathcal{B} \times \mathcal{B}$ restricted to $\mathcal{D} \times \mathcal{D}$ is simply $\widehat{\trianglerighteq}$, since $\widehat{\trianglerighteq}$ is closed in $\mathcal{D} \times \mathcal{D}$. Thus, $cl\widehat{\trianglerighteq}$ is an extension of $\widehat{\trianglerighteq}$, and will be referred to as "$\widehat{\trianglerighteq}$ on $\mathcal{B}$", or simply also as "$\widehat{\trianglerighteq}$" if no misunderstanding is possible. Clearly $X\widehat{\trianglerighteq}Y$ if and only if there exist sequences $\{X_n\}$ and $\{Y_n\}$ in $\mathcal{D}$ converging to X and Y, respectively, such that $X_n\widehat{\trianglerighteq}Y_n$ for all n.

Say that $\widehat{\trianglerighteq}$ on $\mathcal{B}$ satisfies *Homogeneity* if, for all $X, Y \in \mathcal{B}$ and $\lambda \in \mathbf{R}_{++}$ such that $\lambda X, \lambda Y \in \mathcal{B}$: $X\widehat{\trianglerighteq}Y$ if and only if $\lambda X\widehat{\trianglerighteq}\lambda Y$.

Lemma 9 *The relation $\widehat{\trianglerighteq}$ on $\mathcal{B}$ is transitive, reflexive and satisfies Extension, Positivity, Non-degeneracy, Homogeneity, Strong Additivity, Additivity, and Continuity.*

Proof. Extension and Non-degeneracy are immediate. Continuity holds by construction. Positivity and reflexivity follows therefore from the corresponding properties of $\widehat{\trianglerighteq}$ on $\mathcal{D}$.

To verify Homogeneity, take $X, Y \in \mathcal{B}$ and $\lambda \in \mathbf{R}_{++}$ such that $\lambda X, \lambda Y \in \mathcal{B}$ and $X\widehat{\trianglerighteq}Y$. By definitioin, there exist sequences $\{X_n\}$ and $\{Y_n\}$ in $\mathcal{D}$ converging to X and Y, respectively.Write $\lambda = \ell\alpha$, with $\ell \in \mathbf{N}$ and $\alpha \in (0, 1]$. Choose some sequence $\{\alpha_n\}$ in $\mathbf{D}$ converging to α such that $\alpha_n \leq \min\left(\frac{\|X\|}{\|X_n\|}, \frac{\|Y\|}{\|Y_n\|}\right)$. This ensures that, for all $n, \ell\alpha_n X_n \in \mathcal{D}$ and $\ell\alpha_n Y_n \in \mathcal{D}$. By Weak Homogeneity of $\widehat{\trianglerighteq}$ on $\mathcal{D}$, $\alpha_n X_n\widehat{\trianglerighteq}\alpha_n Y_n$ for all n. Hence by $(\ell - 1)$-fold application of Strong Additivity of $\widehat{\trianglerighteq}$ on $\mathcal{D}$, also $\ell\alpha_n X_n\widehat{\trianglerighteq}\ell\alpha_n Y_n$ for all n. By Continuity on $\mathcal{B}$, $\ell\alpha X\widehat{\trianglerighteq}\ell\alpha Y$, as desired.

To verify Strong Additivity on $\mathcal{B}$, consider any $X, X', Y, Y' \in \mathcal{B}$ such that $X\widehat{\trianglerighteq}Y$ and $X'\widehat{\trianglerighteq}Y'$, and take sequences $\{X_n\}, \{X'_n\}, \{Y_n\}$ and $\{Y'_n\}$ in $\mathcal{D}$ converging to X, X', Y and Y', respectively, such that $X_n\widehat{\trianglerighteq}Y_n$ and $X'_n\widehat{\trianglerighteq}Y'_n$ for all n. By Homogeneity on $\mathcal{B}$ (just shown), $\frac{1}{2}X_n\widehat{\trianglerighteq}\frac{1}{2}Y_n$ and $\frac{1}{2}X'_n\widehat{\trianglerighteq}\frac{1}{2}Y'_n$ for all n. Disregarding an initial subsequence if necessary, $\frac{1}{2}X_n + \frac{1}{2}X'_n \in \mathcal{D}$ as well as $\frac{1}{2}Y_n + \frac{1}{2}Y'_n \in \mathcal{D}$ for all n. Hence by Strong Additivity on $\mathcal{D}$, $\frac{1}{2}X_n + \frac{1}{2}X'_n\widehat{\trianglerighteq}\frac{1}{2}Y_n + \frac{1}{2}Y'_n$. By Continuity on $\mathcal{B}$, $\frac{1}{2}X + \frac{1}{2}X'\widehat{\trianglerighteq}\frac{1}{2}Y + \frac{1}{2}Y'$, whence by Homogeneity on $\mathcal{B}$ again $X + X'\widehat{\trianglerighteq}Y + Y'$ as desired.

One direction of Additivity "$X+Z\widehat{\trianglerighteq}Y+Z$ whenever $X\widehat{\trianglerighteq}Y$" follows directly from Strong Additivity and reflexivity. For the converse, consider X, Y, Z such that $X\widehat{\trianglerighteq}Y$ and $X - Z, Y - Z \in \mathcal{B}$. Take sequences $\{X_n\}$, and $\{Y_n\}$ in $\mathcal{D}$ converging to X and Y, respectively, such that $X_n\widehat{\trianglerighteq}Y_n$ for all n. Let $\{Z_n\}$ be any sequence in $\mathcal{D}$ satisfying

$$Z - \max\left(\| X - X_n \|, \| Y - Y_n \|\right) \mathbf{1} - \frac{1}{n}\mathbf{1} \leq Z_n \leq Z - \max\left(\| X - X_n \|, \| Y - Y_n \|\right) \mathbf{1}.$$

By construction, $\{Z_n\}$ converges to Z; moreover, $X_n - Z_n \geq X - \| X - X_n \| \mathbf{1} - Z_n \geq X - Z \geq 0$, and likewise $Y_n - Z_n \geq 0$. Thus $X_n - Z_n \in \mathcal{D}$ and $Y_n - Z_n \in \mathcal{D}$ for all n. By Additivity on $\mathcal{D}$, $X_n - Z_n\widehat{\trianglerighteq}Y_n - Z_n$ for all n, whence $X - Z\widehat{\trianglerighteq}Y - Z$ as desired.

Finally, to verify Transitivity on $\mathcal{B}$, consider any $X, Y, Z \in \mathcal{B}$ such that $X \widehat{\trianglerighteq} Y$ and $Y \widehat{\trianglerighteq} Z$. By Homogeneity on $\mathcal{B}$ $\frac{1}{2}X \widehat{\trianglerighteq} \frac{1}{2}Y$ as well as $\frac{1}{2}Y \widehat{\trianglerighteq} \frac{1}{2}Z$. By Strong Additivity on $\mathcal{B}$, $\frac{1}{2}X + \frac{1}{2}Y \widehat{\trianglerighteq} \frac{1}{2}Y + \frac{1}{2}Z$. Hence by Additivity on $\mathcal{B}$, $\frac{1}{2}X \widehat{\trianglerighteq} \frac{1}{2}Z$, from which one obtains $X \widehat{\trianglerighteq} Z$ again by Homogeneity on $\mathcal{B}$. $\square$

In a final step, extend $\widehat{\trianglerighteq}$ on $\mathcal{B}$ to the set of all bounded random-variables $\mathcal{R} := B(\Sigma, \mathbf{R})$ by defining $\widehat{\trianglerighteq}$ on $B(\Sigma, \mathbf{R})$ as the unique relation $\widetilde{\trianglerighteq}$ on $B(\Sigma, \mathbf{R})$ that coincides on $\mathcal{B}$ with $\widehat{\trianglerighteq}$ on $\mathcal{B}$ and that satisfies Additivity and Homogeneity. (The uniqueness of this extension is immediate; existence follows easily form the Additivity and Homogeneity properties of $\widehat{\trianglerighteq}$ on $\mathcal{B}$). As in section 2.2, say that a relation $\widehat{\trianglerighteq}$ on $\mathcal{R}$ is a *coherent expectation ordering* if it satisfies Transitivity, Reflexivity, Positivity, Non-degeneracy, Homogeneity, Additivity, and Continuity. The following Lemma summarizes the construction, and follows immediately from Lemma 9.

Lemma 10 *The relation $\widehat{\trianglerighteq}$ on $\mathcal{R}$ is a coherent expectation ordering satisfying Extension.*

The following result establishes the existence of a multi-prior representation for coherent expectation orderings. Its proof is omitted, as it follows from combining Theorem 3.61 and 3.76 in Walley (1991); for finite state spaces, a similar result has also been obtained by Bewley (1986).

Theorem 4 *A relation $\widetilde{\trianglerighteq}$ on $\mathcal{R}$ is a coherent expectation ordering if and only if there exists a closed convex set of priors Π such that, for all $X, Y \in \mathcal{R}$,*

$$X \widetilde{\trianglerighteq} Y \text{ if and only if, for all } \pi \in \Pi,\ E_\pi X \geq E_\pi Y.$$

The representing Π is unique in $\mathcal{K}(\Delta(\Sigma))$.

To complete the proof, apply Theorem 4 to the relation $\widehat{\trianglerighteq}$ on $\mathcal{R}$ obtained in Lemma 10. By Extension, for all $A, B \in \Sigma$,

$$A \trianglerighteq B \text{ iff } 1_A \widehat{\trianglerighteq} 1_B \text{ iff, for all } \pi \in \Pi,\ E_\pi 1_A \geq E_\pi 1_B.$$

Thus Π is indeed a multi-prior representation of $\trianglerighteq$. That it is dyadically convex-ranged is an immediate consequence of Equidivisibility.

To demonstrate uniqueness, consider any $\Pi' \in \mathcal{K}(\Delta(\Sigma))$ different from Π with induced expectation ordering $\widehat{\trianglerighteq}_{\Pi'}$. From the uniqueness part of Theorem 4, there exist $X, Y \in \mathcal{R}$ such that $X \widehat{\trianglerighteq} Y$ and not $X \widehat{\trianglerighteq}_{\Pi'} Y$, or such that $X \widehat{\trianglerighteq}_{\Pi'} Y$ and not $X \widehat{\trianglerighteq} Y$. Consider the former case; the latter is dealt with symmetrically. Moreover, by Additivity and Homogeneity, it can be assumed that $X, Y \in \mathcal{B}$. By continuity, monotonicity, and the density of $\mathbf{D}$ in $[0,1]$ it can in fact be assumed that $X, Y \in \mathcal{D}$. Take any $A \in [X]$ and $B \in [Y]$. By Extension, $1_A \widehat{\equiv} X$ and $1_B \widehat{\equiv} Y$, hence $A \widehat{\trianglerighteq} B$. By assumption, for some $\pi \in \Pi'$, $E_\pi X < E_\pi Y$; in view of Lemma 11 just below, $\pi(A) < \pi(B)$, contradicting the assumption that Π' represents $\trianglerighteq$.

Lemma 11 *For any $\pi \in \Pi'$ such that $\widehat{\trianglerighteq}_{\Pi'} = \widehat{\trianglerighteq}$, and any $X \in \mathcal{D}$ and $A \in [X] : E_\pi X = \pi(A)$.*

Write $X = \sum_i \frac{\ell_i}{2^{k_i}} 1_{E_i}$ and $A = \sum_i A_i$ such that $A_i \in \frac{\ell_i}{2^{k_i}} E_i$. By assumption, one can split each E_i into 2^{k_i} equally likely events $\{E_{i1}, ..., E_{i2^{k_i}}\}$ such that $A_i = \sum_{j \leq \ell_i} E_{ij}$. For any $\pi \in \Pi'$ such that $\widehat{\trianglerighteq}_{\Pi'} = \widehat{\trianglerighteq}$, $\pi(E_{ij}) = \pi(E_{ij'})$ for all i, j, j', hence $\pi(A_i) = \frac{\ell_i}{2^{k_i}} \pi(E_i)$ by additivity of π. Hence $\pi(A) = \sum_i \frac{\ell_i}{2^{k_i}} \pi(E_i) = E_\pi X$. $\square$

Proof of Theorem 3.

To verify part i) of the Theorem, take any coherent and almost-equidivisble $\trianglerighteq$. It suffices to show that, for any E and any $n \in \mathbf{N}$, there exists a set $A \subseteq E$ such that $\frac{1}{2} \leq \pi^-(A/E)$ and $\pi^+(A/E) \leq \frac{1}{2} + \frac{1}{n}$.

Fix E and n. By Almost Equidivisibility, there exists a partition of E into $2n-1$ sets $\{A_1, ..., A_{2n-1}\}$ such that, for any subfamily of n sets $\{A_{i_1}, ..., A_{i_n}\}$, $\sum_{j=1,..,n} A_{i_j} \trianglerighteq E \backslash \left(\sum_{j=1,..,n} A_{i_j} \right)$. We claim that, for any $j \in \{1, ..., 2n-1\}$, $\pi^+(A_j/E) \leq \frac{1}{n}$. This suffices, since $\pi^+(\sum_{i=1,..,n} A_i/E) \leq \pi^+(\sum_{i=1,..,n-1} A_i/E) + \pi^+(A_n/E) \leq \frac{1}{2} + \frac{1}{n}$ as well as $\pi^-(\sum_{i=1,..,n} A_i/E) \leq \frac{1}{2}$ by construction.

To verify this claim, suppose that, by contradiction, for some $\pi \in \Pi_{\trianglerighteq}$ and $j \in \{1, ..., 2n-1\}, \pi(A_j/E) > \frac{1}{n}$. W.l.o.g., assume that $\pi(A_i/E)$ is increasing in i. Then $\pi(\sum_{i=1,..,n} A_i/E) < \frac{n}{2n-2}(1 - \frac{1}{n}) = \frac{1}{2}$, which contradicts the assumption that $\sum_{i=1,..,n} A_i \trianglerighteq E \backslash \left(\sum_{i=1,..,n} A_i \right)$.

Conversely, take an almost-convex-ranged set Π, and consider any event E such that $\pi^+(E) > 0$ and any $n > 0$. By a straightforward inductive argument, there exists a partition of E into $2n-1$ subevents $\{A_1, ..., A_{2n-1}\}$ such that, for all i, $\frac{1}{2n} < \pi^-(A_i/E) \leq \pi^+(A_i/E) < \frac{1}{2(n-1)}$, from which the Almost Equidivisibility of $\trianglerighteq_\Pi$ is immediate.

To verify part ii) of the Theorem, we shall show that $\trianglerighteq_\Pi$, viewed as a relation on indicator-functions, has a unique extension to an expectation ordering $\widehat{\trianglerighteq_\Pi}$ on $\mathcal{F}(\Sigma, [0,1])$, and thus also $\mathcal{F}(\Sigma, \mathbf{R})$. Since by Theorem 4 of the Appendix, for any Π' different from Π, $\widehat{\trianglerighteq}_{\Pi'} \neq \widehat{\trianglerighteq}_\Pi = \widehat{\trianglerighteq_\Pi}$, this implies that in fact $\trianglerighteq_{\Pi'} \neq \trianglerighteq_\Pi$.

Consider an almost-convex-ranged set of priors Π and any extension to a coherent expectation ordering on $\mathcal{F}(\Sigma, [0,1])$ $\widehat{\trianglerighteq}$. The following Lemma ensures the possibility of an approximate mixture-space construction.

Lemma 12 *i) For any* $A \subseteq E$ *such that* $\pi^+(A/E) < \frac{1}{m}$, $1_A \widehat{\trianglelefteq} \frac{1}{m} 1_E$.

ii) For any $A \subseteq E$ *such that* $\pi^-(A/E) > \frac{1}{m}$, $1_A \widehat{\trianglerighteq} \frac{1}{m} 1_E$.

iii) For any $\alpha < \beta \in [0,1]$, *and any* $E \in \Sigma$, *there exists* $A \subseteq E$ *such that* $\alpha 1_E \widehat{\trianglelefteq} 1_A \widehat{\trianglelefteq} \beta 1_E$.

iv) For any $Y, Z \in \mathcal{F}(\Sigma, [0,1])$ *such that* $Y \geq Z$ *and* $Y(\omega) > Z(\omega)$ *whenever* $Y(\omega) > 0$ *and* $Z(\omega) < 1$, *there exists* $A \in \Sigma$ *such that* $Y \widehat{\trianglerighteq} 1_A \widehat{\trianglerighteq} Z$.

Proof of Lemma.

i) Take any $A \subseteq E$ such that $\pi^+(A/E) < \frac{1}{m}$. It is easily verified that by almost-convex-rangedness there exist $m-1$ disjoint sets B_i such that $A + \sum_i B_i = E$ and $A \widehat{\trianglelefteq} B_i$ for all i. By Strong Additivity, $m 1_A \widehat{\trianglelefteq} 1_A + \sum_i 1_{B_i} = 1_E$. By Homogeneity, one infers that $1_A \widehat{\trianglelefteq} \frac{1}{m} 1_E$ as desired.

ii) is verified analogously.

To show iii), take any m and n such that $\alpha < \frac{m}{n+1} < \frac{m}{n} < \beta$. One can easily establish from almost-convex-rangedness that there exist m disjoint subsets A_i of E such that $\frac{1}{n+1} < \pi^-(A_i/E)$ and $\pi^+(A_i/E) < \frac{1}{n}$. By parts i) and ii), $\frac{1}{n+1} 1_E \widehat{\trianglelefteq} 1_{A_i} \widehat{\trianglelefteq} \frac{1}{n} 1_E$ for all i. Setting $A = \sum_i A_i$, it follows by Strong Additivity that

$$\alpha 1_E \leq \frac{m}{n+1} 1_E \widehat{\trianglelefteq} 1_A \widehat{\trianglelefteq} \frac{m}{n} 1_E \leq \beta 1_E,$$

which suffices in view of the monotonicity of $\widehat{\trianglerighteq}$.

Finally, to verify iv), take any $Y, Z \in \mathcal{F}(\Sigma, [0,1])$ such that $Y \geq Z$ and $Y(\omega) > Z(\omega)$ whenever $Y(\omega) > 0$ and $Z(\omega) < 1$. Write $Y = \sum_i y_i 1_{E_i}$ and $Z = \sum_i z_i 1_{E_i}$. By part iii), for each i, there exists $A_i \subseteq E_i$ such that

$$y_i 1_{E_i} \widehat{\trianglelefteq} 1_{A_i} \widehat{\trianglelefteq} z_i 1_{E_i}.$$

Note that if $y_i = 0$, then also $z_i = 0$, and one can set $A_i = \emptyset$; similarly, if $z_i = 1$, then also $y_i = 1$, and one can set $A_i = E_i$. Setting $A = A_i$, the desired conclusion follows from Strong Additivity. $\square$

To conclued the proof of part ii) of the Theorem, fix any $Y, Z \in \mathcal{F}(\Sigma, [0,1])$. Take a decreasing sequence $\{Y_n\}$ in $\mathcal{F}(\Sigma, [0,1])$ converging to Y such that $Y_n(\omega) > Y(\omega)$ whenever $Y(\omega) < 1$, as well as an increasing sequence $\{Z_n\}$ in $\mathcal{F}(\Sigma, [0,1])$ converging to Z such that $Z_n(\omega) > Z(\omega)$ whenever $Z(\omega) > 0$. By part iv) of Lemma 12, there exist sequences of events $\{A_n\}$ and $\{B_n\}$ in Σ such that, for all n,

$$Y_n \widehat{\trianglerighteq} 1_{A_n} \widehat{\trianglerighteq} Y \text{ and } Z \widehat{\trianglerighteq} 1_{B_n} \widehat{\trianglerighteq} Z_n.$$

We claim that $Y \widehat{\trianglerighteq} Z$ if and only if $1_{A_n} \widehat{\trianglerighteq} 1_{B_n}$ for all n. Indeed, the only-if part is immediate from Transitivity, while the if-part follows directly from Continuity. This clearly suffices to establish uniqueness of the extension $\widehat{\trianglerighteq}$, which suffices as argued above. $\square$

Proof of Proposition 2.

Consider any (non-null) event $A \in \Sigma$ and any $\gamma \in (0, 1)$ and $\eta > 0$; we will show that there exists an event $B \in \Sigma$, $B \subseteq A$ such that $\frac{1}{1+\eta}\gamma\pi(A) \leq \pi(B) \leq (1+\eta)\pi(A)$ for all $\pi \in \Pi$.

By the compactness of Ω and assumption ii) on Ψ, there exists $\varepsilon > 0$ such that

$$\Psi(a, b) \leq 1 + \eta \text{ whenever } \delta(a, b) \leq \varepsilon. \tag{9}$$

By compactness, Ω can be covered by a finite number of ε-balls; let $\mathcal{A} = \{D_i\} \subseteq \Sigma$ denote the finite partition generated by these balls.

As a non-atomic, countably-additive measure, λ is convex-ranged (see, for example, Aliprantis-Border (1999, p. 357). Hence, for each i, there exists a set $B_i \subseteq A \cap D_i$ such that $\lambda(B_i) = \gamma\lambda(A \cap D_i)$. We claim that $B = \sum_i B_i$ is the desired set.

By the construction of the D_i, (**??**) and (9), for all $\pi \in \Pi_{\lambda,\Psi}$ and all i,

$$\frac{1}{1+\eta}\frac{\lambda(B_i)}{\lambda(A \cap D_i)} \leq \frac{\pi(B_i)}{\pi(A \cap D_i)} \leq (1+\eta)\frac{\lambda(B_i)}{\lambda(A \cap D_i)},$$

and therefore

$$\frac{1}{1+\eta}\gamma\pi(A \cap D_i) \leq \pi(B_i) \leq (1+\eta)\gamma\pi(A \cap D_i)$$

Summing over i, one obtains

$$\frac{1}{1+\eta}\gamma\pi(A) \leq \pi(B) \leq (1+\eta)\gamma\pi(A),$$

as desired. □

Proof of Proposition 3.

We will first show that $\trianglerighteq_{\lambda,\Psi} \subseteq \trianglerighteq_{(\Pi_{\lambda,\Psi})}$. To see this, take any A, B and such that $\lambda(A) \geq \Psi(B, A)\lambda(B)$ and any $\pi \in \Pi_{\lambda,\Psi}$; we need to verify that $\pi(A) \geq \pi(B)$. Indeed, $\pi(A) = \int_A \phi(\omega)\, d\lambda \geq \lambda(A) \inf_{\omega \in A} \phi(\omega) \geq \lambda(B)\, \Psi(A, B) \frac{\sup_{\omega \in B} \phi(\omega)}{\Psi(B,A)} \geq \int_B \phi(\omega)\, d\lambda = \pi(B)$.

To complete the proof, we need to show conversely that any probability measure π compatible with $\trianglerighteq_{\lambda,\Psi}$ is in fact contained in $\Pi_{\lambda,\Psi}$. We will prove this in a sequence of steps.

> **Step 1.** π is absolutely continuous with respect to λ; hence, in particular, π is countably additive.

Take any natural numbers K and L such that $L \geq \Psi(\Omega, \Omega)$, and any A such that $\lambda(A) \leq \frac{1}{KL}$. We will show that $\pi(A) \leq \frac{1}{K}$.

By the convex-rangedness of λ, there exists a partition $\{A_i\}_{i\leq KL}$ of Ω such that $A_1 \supseteq A$ and $\lambda(A_i) = \frac{1}{KL}$. By combining L members $\{A_i\}$ each, one obtains a partition of Ω $\{B_j\}_{j\leq K}$ that is coarser than $\{A_i\}$ and satisfies $\lambda(B_j) = \frac{1}{K}$ for all j. Since L was chosen to exceed $\Psi(\Omega,\Omega)$, and thus a fortiori $\Psi(A_i, B_j)$, one has for all i, j that $B_j \trianglerighteq_{\lambda,\Psi} A_i$, which implies $\pi(A) \leq \pi(A_1) \leq \min_{j\leq K} \pi(B_j) \leq \frac{1}{K}$.

By step 1, π has a Radon-Nikodym derivative ϕ with respect to λ. Let $B^{\varepsilon}(a)$ denote the closed ε-ball around a. Recall also that since (Ω,δ) is a compact metric space and thus second countable, there is a smallest closed set of full measure, the support $\operatorname{supp}\lambda$.

Step 2. For all $a, b \in \operatorname{supp}\lambda$ and all $\varepsilon > 0$, $\operatorname{ess\,sup}_{\omega\in B^{\varepsilon}(a)} \phi(\omega) \leq \Psi(B^{\varepsilon}(a), B^{\varepsilon}(b)) \operatorname{ess\,inf}_{\omega\in B^{\varepsilon}(b)} \phi(\omega)$.

We verify the claim by contradiction. That is, assume that there exist λ-non-null sets $A \subseteq B^{\varepsilon}(a)$, $B \subseteq B^{\varepsilon}(b)$ such that

$$\inf_{\omega\in A} \phi(\omega) > \sup_{\omega\in B} \phi(\omega) \Psi(B^{\varepsilon}(a), B^{\varepsilon}(b)). \tag{10}$$

By the convex-rangedness of λ, it can moreover be assumed that

$$\lambda(B) = \Psi(B^{\varepsilon}(a), B^{\varepsilon}(b)) \lambda(A). \tag{11}$$

By (10) and (11),

$$\pi(A) \geq \lambda(A) \inf_{\omega\in A} \phi(\omega) > \lambda(A) \Psi(B^{\varepsilon}(a), B^{\varepsilon}(b)) \sup_{\omega\in B} \phi(\omega) = \lambda(B) \sup_{\omega\in B} \phi(\omega) \geq \pi(B).$$

However, by (11), and the fact that $\pi \in \Pi_{\lambda,\Psi}$, by the first part of the proof, $\pi(B) \geq \pi(A)$, the desired contradiction.

In the sequel, we will make repeated use of the following implication of the continuity of Ψ.

Fact 4 *For all* $a \in \Omega : \lim_{\varepsilon\to 0} \Psi(B^{\varepsilon}(a), B^{\varepsilon}(a)) = 1$.

Step 3. π has a continuous density $\widetilde{\phi}$.

For $n \in N$, define functions ϕ_-^n and ϕ_+^n as follows:

$$\begin{aligned} \phi_-^n(a) &:= \operatorname{ess\,inf}_{\omega\in B^{\frac{1}{n}}(a)} \phi(\omega) \text{ for } a \in \operatorname{supp}\lambda, \text{ and } := \phi(a) \text{ otherwise;} \\ \phi_+^n(a) &:= \operatorname{ess\,sup}_{\omega\in B^{\frac{1}{n}}(a)} \phi(\omega) \text{ for } a \in \operatorname{supp}\lambda, \text{ and } := \phi(a) \text{ otherwise.} \end{aligned}$$

By step 2,

$$\phi_-^n(a) \geq \phi_+^n(a) \Psi(B^{\frac{1}{n}}(a), B^{\frac{1}{n}}(a)). \tag{12}$$

In view of Fact 4, the increasing and decreasing sequences $\{\phi_-^n\}$ and $\{\phi_+^n\}$ converge to the same function $\widetilde{\phi}$.

Now, for any $A \in \Sigma$,

$$\int_A \phi_-^n(\omega)\, d\lambda \leq \int_A \phi(\omega)\, d\lambda = \pi(A) \leq \int_A \phi_+^n(\omega)\, d\lambda.$$

Since by the Monotone Convergence Theorem both $\int_A \phi_-^n(\omega)\, d\lambda$ and $\int_A \phi_+^n(\omega)\, d\lambda$ converge to $\int_A \widetilde{\phi}(\omega)\, d\lambda$, one obtains $\int_A \widetilde{\phi}(\omega)\, d\lambda = \pi(A)$ for any $A \in \Sigma$. Thus $\widetilde{\phi}$ is a density for π as well.

To verify that $\widetilde{\phi}$ is continuous, consider a, b such that $\delta(a,b) \leq \frac{1}{n}$. One has

$$\left|\widetilde{\phi}(a) - \widetilde{\phi}(b)\right| \leq \left|\widetilde{\phi}(a) - \phi_-^n(a)\right| + \left|\phi_-^n(a) - \phi_-^n(b)\right| + \left|\phi_-^n(b) - \widetilde{\phi}(b)\right|.$$

Now $\phi_-^n(a) \leq \Psi\left(B^{\frac{2}{n}}(b), B^{\frac{2}{n}}(b)\right)\phi_-^n(b)$, since $B^{\frac{1}{n}}(a) \subseteq B^{\frac{2}{n}}(b)$ by the triangle inequality. Thus, by Fact 4, all three terms on the right-hand side converge to 0 as $n \to \infty$, which establishes the continuity of $\widetilde{\phi}$.

Step 4. $\widetilde{\phi}(a) \leq \Psi(a,b)\widetilde{\phi}(b)$ for all $a, b \in \operatorname{supp}\lambda$.

Suppose not, i.e. suppose that $\widetilde{\phi}(a) > \widetilde{\phi}(b)\Psi(a,b)(1+\delta)^2$ for some $\delta > 0$. By Fact 4, there exists $\varepsilon > 0$ such that $\Psi(B^\varepsilon(a), B^\varepsilon(a)) \leq 1+\delta$ as well as $\Psi(B^\varepsilon(b), B^\varepsilon(b)) \leq 1+\delta$.

Hence by assumption iii) on Ψ,

$$\Psi(B^\varepsilon(a), B^\varepsilon(b)) \leq \Psi(B^\varepsilon(a), B^\varepsilon(a))\Psi(a,b)\Psi(B^\varepsilon(b), B^\varepsilon(b)) \leq \Psi(a,b)(1+\delta)^2.$$

By the continuity of $\widetilde{\phi}$,
$\operatorname{ess\,sup}_{\omega \in B^\varepsilon(a)} \widetilde{\phi}(\omega) \geq \widetilde{\phi}(a)$ as well as $\operatorname{ess\,inf}_{\omega \in B^\varepsilon(b)} \widetilde{\phi}(\omega) \leq \widetilde{\phi}(b)$.
Therefore,

$$\operatorname{ess\,sup}_{\omega \in B^\varepsilon(a)} \widetilde{\phi}(\omega) \geq \widetilde{\phi}(a) > \widetilde{\phi}(b)\Psi(a,b)(1+\delta)^2 \geq \Psi(B^\varepsilon(a), B^\varepsilon(b))\operatorname{ess\,inf}_{\omega \in B^\varepsilon(b)} \widetilde{\phi}(\omega),$$

in contradiction to step 2. □

Proof of Fact 2.

Parts ii) and iii) are verified by elementary computation. To verify part i), it evidently suffices to show that for any disjoint A, B such that $\lambda(A) < K\lambda(B)$, there exists $\pi \in \widetilde{\Pi}^K$ such that $\pi(A) < \pi(B)$. To see this, note first that since $\lambda(A) + \lambda(B) \leq 1$, $\lambda(A) < \frac{K}{K+1}$. Therefore by the convex-rangedness of λ, there exists $D \subseteq A^c$ such that $\lambda(D) = \frac{1}{K+1}$ and a) $D \supseteq B$ or b) $B \supseteq D$. In the first case, $\pi_D(B) = \frac{K+1}{2}\lambda(B)$; in the second case, that is, whenever $\lambda(B) \geq \frac{1}{K+1}$, $\pi_D(B) \geq \pi_D(D) = \frac{1}{2}$. On the other hand, $\pi_D(A) = \frac{K+1}{2K}\lambda(A)$, which is less than $\frac{1}{2}$ since $\lambda(A) < \frac{K}{K+1}$. Thus, in either case, $\pi_D(A) < \pi_D(B)$, as needed to be shown.

Proof of Proposition 4.

Let $\overline{\mathcal{F}}$ denote the set of Cauchy-sequences in $\mathcal{F}$ viewed as a superset of $\mathcal{F}$ endowed withe the canonical extension of δ, $\delta(\{f_n\},\{g_n\}) = \limsup_{n\to\infty} \delta(f_n, g_n)$. Since $\mathcal{F}$ is dense in $\overline{\mathcal{F}}$, by a classical result on metric spaces (see, e.g. Aliprantis/Border (1999), Lemma 3.8, p. 77), V has a unique, uniformly continuous extension to $\overline{\mathcal{F}}$ likewise denoted by V.

Lemma 13 *For any $f, f' \in \mathcal{F}$ and $F \in \mathcal{F}^{AA}$, $\delta(f, f') \leq \delta'(f, F) + \delta(f', F)$.*

The verification of the lemma is routine. Indeed, for any $\pi \in \Pi$, $d(\pi \circ f^{-1}, \pi \circ g^{-1}) \leq \sup_{i\in I} d\left(\pi(./F_i) \circ f^{-1}, \pi(./F_i) \circ g^{-1}\right) \leq \sup_{i\in I} \left(d\left(\pi(./F_i) \circ f^{-1}, q_i\right) + d\left(\pi(./F_i) \circ g^{-1}, q_i\right)\right) \leq \sup_{i\in I} d\left(\pi(./F_i) \circ f^{-1}, q_i\right) + \sup_{i\in I} d\left(\pi(./F_i) \circ g^{-1}, q_i\right) \leq \delta'(f, F) + \delta(f', F)$. □

It is immediate from the lemma that any sequence $\{f_n\} \in [F]$ is a δ-Cauchy sequence, and, for any two sequences $\{f_n\}, \{g_n\} \in [F]$ $\delta(\{f_n\},\{g_n\}) = 0$, whence by the continuity of V, $V(\{f_n\}) = V(\{g_n\})$; hence one can define $V(F) := V(\{f_n\})$ for any $\{f_n\} \in [F]$.

Lemma 14 *$F \succsim^{AA} G$ if and only if $V(F) \geq V(G)$.*

To verify the if-part, if $F \succsim^{AA} G$ then there exists by definition $\{f_n\} \in [F]$ and $\{g_n\} \in [G]$ such that $f_n \succsim g_n$ for all n. By continuity,

$$V(F) = V(\{f_n\}) = \lim_{n\to\infty} V(f_n) \geq \lim_{n\to\infty} V(g_n) \geq V(\{f_n\}) = V(G).$$

Conversely, suppose that $V(F) \geq V(G)$. By Almost-Convex-Rangedness, one can find a sequence $\{f_n\} \in [F]$ such that f_m stochastically dominates f_n whenever $m < n$, and therefore by Generalized Stochastic Dominance such that $V(f_n)$ does not increase and converges to $V(F)$. By the same token, using Almost-Convex-Rangedness, one can find a sequence $\{g_n\} \in [G]$ such that g_m stochastically dominates g_n whenever $m > n$, and therefore by Generalized Stochastic Dominance such that $V(g_n)$ does not decrease and converges to $V(G)$. It follows that, for all n, $V(f_n) \geq V(F) \geq V(G) \geq V(g_n)$, whence $F \succsim^{AA} G$. □

This lemma implies immediately that $\succsim^{AA}$ is a weak order extending $\succsim$. Continuity with respect δ'' follows from the inequality

$$\delta''(F, G) \geq \delta(\{f_n\},\{g_n\}) \text{ for any } \{f_n\} \in [F] \text{ and } \{g_n\} \in [G].$$

REFERENCES

[1] Allais, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Américaine", *Econometrica* 21, 503-546.

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

[3] Bewley, T. F. (1986): "Knightian Decision Theory, Part I," Cowles Foundation Discussion Paper No. 807.

[4] Casadesus, R. , P. Klibanoff and E. Ozdenoren: "Maxmin Expected Utility over Savage Acts with a Set of Priors", *Journal of Economic Theory* 92, 35-65.

[5] Choquet, G. (1953): "Theory of Capacities", *Ann. Instit. Fourier* (Grenoble) 5, 131-295.

[6] Dempster, A. (1967), "Upper and Lower Probabilities Induced by a Multi-Valued Mapping", *Annals of Mathematical Statistics* 38, 325-339.

[7] Eichberger, J. and D. Kelsey (1996): "Uncertainty-Aversion and Preference for Randomisation", *Journal of Economic Theory 71,* 31-43.

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

[9] Epstein, L. (1999): "A Definition of Uncertainty Aversion", *Review of Economic Studies* 66, 579-608.

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

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

[12] Ghirardato, P. and M. Marinacci (2001a): "Ambiguity Made Precise: A Comparative Foundation", *Journal of Economic Theory,* forthcoming.

[13] Ghirardato, P. and M. Marinacci (2001b): "Risk, Ambiguity, and the Separation of Utility and Beliefs", *Mathematics of Operations Research,* forthcoming.

[14] Ghirardato, P., F. Maccheroni and M. Marinacci (2001c): "A Fully Subjective Perspective on Ambiguity," in progress October 2001.

[15] Ghirardato, P., F. Maccheroni, M. Marinacci and M. Siniscalchi (2001d): "Subjective Foundations for Objective Randomization: A New Spin on Roulette Wheels", mimeo.

[16] Kannai, Y. (1992): "The Core and Balancedness", in: R. Aumann and S. Hart (eds.), *Handbook of Game Theory,* North Holland, Amsterdam, 355-395.

[17] Klibanoff, P. (2001a): "Stochastically Independent Randomization and Uncertainty Aversion", *Economic Theory* 18, 605-620.

[18] Klibanoff, P. (2001b): "Characterizing Uncertainty Aversion Through Preference for Mixtures", *Social Choice and Welfare* 18, 289-301.

[19] Kraft, C. , Pratt, J. and A. Seidenberg (1959): "Intuitive Probability on Finite Sets," *Annals of Mathematical Statistics* 30, 408-419.

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

[21] Nehring, K. (1991): *A Theory of Rational Decision with Vague Beliefs.* Ph.D. dissertation, Harvard University.

[22] Nehring, K. (1994): "On the Interpretation of Sarin and Wakker's 'A Simple Axiomatization of Nonadditive Expected Utility Theory"', *Econometrica* 62, 935-938.

[23] Nehring, K. (1996): "Preference and Belief without the Independence Axiom", talk presented at LOFT2 in Torino, Italy.

[24] Nehring, K. (1999): "Capacities and Probabilistic Beliefs: A Precarious Coexistence", *Mathematical Social Sciences* 38, 197-213.

[25] Nehring, K. (2000): "Rational Choice under Ignorance", *Theory and Decision* 48, 205-240.

[26] Nehring, K. (2001): "Imprecise Qualitative Probability", in progress.

[27] Ramsey, F. (1931): "Truth and Probability", in *The Foundations of Mathematics and other Logical Essays,* reprinted in: H. Kyburg and H. Smokler (eds., 1964), *Studies in Subjective Probability,* Wiley, New York, 61-92.

[28] Sarin, R. and P. Wakker (1992): "A Simple Axiomatization of Nonadditive Expected Utility Theory", *Econometrica* 60, 1255-1272.

[29] Sarin, R. and P. Wakker (1995): "On the Interpretation of Likelihood in Choquet Expected Utility", mimeo, UCLA and University of Leiden.

[30] Savage, L.J. (1954). *The Foundations of Statistics.* New York: Wiley. Second edition 1972, Dover.

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

[32] Smith, C. (1961): "Consistency in Statistical Inference and Decision," *Journal of the Royal Statistical Society*, Series B, 22, pp. 1-25.

[33] Shafer, G. (1976). *A Mathematical Theory of Evidence*, Princeton: Princeton U.P.

[34] Wakker, P. (1989): *Additive Representations of Preferences.* Dordrecht: Kluwer.

[35] Wakker, P. (2001): "On the Composition of Risk Preferences and Belief", *Psychological Review*, forthcoming.

[36] Walley, P. (1991): *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

[37] Zhang, J.-K. (1997): "Subjective Ambiguity, Probability and Capacity", mimeo, University of Toronto.