Imprecise Probabilistic Beliefs

Klaus Nehring

University of California, Davis, and

Institute for Advanced Study, Princeton

First Version: November 2002

This Version: October 2003

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", and 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 imprecise probabilistic beliefs. In the final part of the paper, we formulate behaviorally general axioms relating preferences and probabilistic beliefs. If beliefs are almost-convex-ranged, these axioms imply that preferences can be represented in an Anscombe-Aumann-style framework.

Keywords: ambiguity, comparative likelihood, coherence, multiple priors, stochastic dominance, consequentialism.[1]

[1]**Address:** School of Social Science, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540 (until June 30, 2004) ; Department of Economics, University of California, Davis, CA, 95616.(after July 1, 2004) ; e-mail throughout: kdnehring@ucdavis.edu.

1. INTRODUCTION

It is by now widely recognized that the assumption that decision makers are guided in their decisions by a well-defined subjective probability measure is questionable empirically as well as normatively. Empirically, this assumption was refuted decisively by Ellsberg's (1961) celebrated experiments. While this and much other evidence provide compelling reasons for abandoning the assumption that behavior can be explained globally in terms of precise probabilistic beliefs, it does not render the notion of probabilistic belief useless if it is applied "partially", that is: if applied to some events or event comparisons. Indeed, the very formulation of Ellsberg's original experiment involves a comparison of events with probabilistic beliefs to other potentially ambiguous ones.

To model such decision making in the context of imprecise probabilistic beliefs, we shall describe a decision-maker in terms of two entities rather than one representing preferences and beliefs separately. 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 preferences. While under expected utility and, more generally, under probabilistic sophistication, a decision maker's probabilistic beliefs are canonically "revealed" by his betting behavior, this no longer holds in the presence of ambiguity, for now there are (at least) two determinants of betting preferences: beliefs –however construed– and ambiguity attitudes.[2]

[2] For different reasons, a canonical definition of "revealed subjective probability" from choice-behavior fails to be possible in the case of state-dependendent preferences; see Karni et al. (1983) and the subsequent literature.

Even in the context of Savage's SEU theory, this "canonical" definition has been criticized as not necessarily capturing the decision maker's true beliefs (Shervish, Seidenfeld and Kadane (1990), Karni (1996), Grant-Karni (2000)); this criticism assumes, however, a non-behaviorist point of view to begin with.

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 modelers, we typically attribute to economic agents particular preferences over uncertain acts *because* we have some idea about the beliefs that can be plausibly attributed to the agents in a particular situation. In both cases, we think directly in terms of beliefs rather than preferences. This is the intuitive substance of including the decision maker's probabilistic beliefs among the primitives.[3]

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), de Finetti (1931) 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 betting preferences canonically: if A is judged at least as likely as B, then betting on A must be weakly preferred to betting on B. More generally, preferences should respect any stochastic dominance relations entailed by the likelihood relation and the ranking of consequences: if, for any consequence, some act f is at least as likely to generate this consequences or

[3] In section 5, we will further comment on the implications of this move, and discuss the extent to which this paper's contribution is meaningful from a strictly behaviorist perspective as well.

a better one than another act g, f should be weakly preferred to g; in this case, we will simply say that the decision-makers are "compatible" with his beliefs. A decision maker in the context of probabilistic beliefs is thus formally given as a pair $(\succsim, \trianglerighteq)$ such that his preferences $\succsim$ are compatible with his beliefs $\trianglerighteq$. 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, respect for Stochastic Dominance accommodates 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 issues about the representation of probabilistic beliefs are more fundamental than particular behavioral assumptions, as argued compellingly by Machina-Schmeidler (1992) and Epstein-Zhang (2001). The goal of the present 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 set of admissible probability measures ("priors") according to which judging an event A as at least as likely as B is equivalent to A's probability weakly exceeding that of B, for any admissible prior 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 structural 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 for complete likelihood relations. 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. In exchange, Savage had to pay the price of restricting attention to non-atomic (strictly speaking: "convex-ranged") probability measures.

The main result of this paper, Theorem 2, is a counterpart to Savage's result for incomplete comparative likelihood relations; it appears to be the first such result in the literature. 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 in particular 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 a structural "Equidivisibility" condition that assumes that 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 random device, as assumed in the widely-used Anscombe-Aumann framework. In an important sense, Equidivisibility is thus not really restrictive at all since any coherent likelihood relation can be extended to a larger one incorporating a hypothetical random-device on a larger state space. See section 3 for details and further examples.

Uniqueness

Next to providing axiomatic foundations, a second main concern of the paper is to establish that comparative likelihood relations are adequately expressive as formal representations of "imprecise probabilistic beliefs" (understood in an 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 the canonical primitive.

Fortunately, Equidivisibility ensures not only the existence of a multi-prior representation, but also its uniqueness. In section 3, we investigate to what extent it is it possible to weaken this assumption 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 also show by example that uniqueness is lost easily when the likelihood relation is not almost-convex-ranged.

The difference between almost- and strict convex-rangedness can be substantial. For example, it is frequently appropriate to assume 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 smoothness 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.

Applications to Decision Theory

Comparative likelihood relations represented by (almost) convex-ranged sets of priors promise to be very fruitful for decision theory itself. From the mathematical point of view, this happens because convex-ranged beliefs effectively endow the event-spaces with a mixture-space structure. In particular, we show that if preferences over multi-valued acts are compatible with a convex-ranged set of priors, 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.

In a companion paper (Nehring 2001), we have used this 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 exclusively 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 sets of priors 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 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 fail to be behaviorally general. Multiple-prior representations of complete preference orderings have obtained by Gilboa-Schmeidler (1989), Ghirardato et al. (2002) and Casadesus et al. (2000); again, these are about preferences, not belief, and are behaviorally quite restrictive.

2. There is a sizeable literature on comparative likelihood relations the great majority of which focuses on the complete case; see Fishburn (1986) and Regioli (1999) for surveys. In the incomplete case, one can use standard arguments from the theory of linear inequalities to obtain a characterization of coherence for likelihood relations defined on arbitrary families of sets; see Walley (1991 p. 192-3) and related earlier results by Heath-Suddert (1972) and, in the complete, finite-state case, Kraft et al (1959). In view of the complexity and epistemic intransparency of the involved conditions, such characterizations have generally not been considered to be of significant foundational interest. A main contribution of Theorem 2 is precisely to provide an epistemic counterpart to the algebraic logic of these conditions, for which the Splitting axiom (combined with Equidivisibility) is crucial. The uniqueness issue has been studied so far only in the complete case.[4] Likewise, the central notions of convex-

[4]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

ranged and almost-convex-ranged sets of priors appear to be novel.

3. Some of the recent literature on decision making under ambiguity can be read as offering proposals for characterizing a decision maker's unconditional probabilistic beliefs directly through definitions of "unambiguous events" revealed by the preference relation; see Epstein-Zhang (2001), Ghirardato-Marinacci (2002) and Nehring (1999)). As further discussed 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 relationship of the Epstein-Zhang definition to the natural definition of unambiguous events in terms of explicit beliefs will be specifically discussed in Appendix A.1 below. In fact, the present paper originated in an attempt 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; see also Ghirardato-Maccheroni-Marinacci (2002) for related work in the case "utility sophisticated preferences".[5]

4. 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. Indeed, the already mentioned Machina (2001) inspired our interest in almost-convex-ranged sets of priors; otherwise, his contribution arose independently of

any hint of this in the literature (see, for example, the discussion of comparative likelihood relations in Walley (1991, 191-197).

[5]The notion of "utility sophisticated preferences" as introduced in Nehring (2001) describes decision-makers who depart from expected utility only for reasons of ambiguity; in particular, utility sophisticated decision-makers maximize expected utility when probabilities are known.

ours. Congenial with our work, Machina imposes epistemically motivated restrictions on preferences. However, these assumptions are imposed directly in the form of a continuity condition, while we model the postulated probabilistic beliefs explicitly as a likelihood relation, and obtain analogous preference restrictions via compatibility. Our approach can thus be viewed as generalizing and grounding Machina's. [6] Among many other things, Machina (2002) defines a global comparative likelihood relation from preferences, assuming 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.

5. There is a small literature that attempts to characterize how imprecise probabilistic beliefs *determine* rational choice under ambiguity (see Jaffray (1989), and Nehring (1991, 2000), as well as, at a more conceptual level, Levi (1980)). One can think of the beliefs as described in these works as convex sets of priors. In contrast to the present paper, the interpretation of such sets (respectively their associated likelihood relations) is an *exhaustive* one. That is, in the language of the present paper, the DM's beliefs $\trianglerighteq$ fully summarize all likelihood judgements the DM is willing to make; event comparisons for which this relation is incomplete reflect deliberate *suspensions of judgement*. The necessity of an exhaustive interpretation in these contributions reflects their different decision-theoretic goal, which was to characterize the implications of substantive rationality assumptions on choice behavior rather than to provide behaviorally general and thus minimal restrictions on choices.

[6] 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.

Overview

Section 2 characterizes coherent likelihood-relations with a convex-ranged multi-prior representation assuming in particular of the likelihood relation, and illustrates the notions of equidivisibility/convex-rangedness with a number of examples. In Section 3, we weaken equidivisibility to almost-equidivisibility, and show that coherent almost-equidivisible likehood-relations have a unique almost-convex-ranged multi-prior-representation, and show that uniqueness is easily lost when convex-rangedness is weakened further. Section 4 formulates conditions that relate beliefs to preferences, and briefly discusses some issues that arise when preferences are state-dependent. If beliefs are convex-ranged, these conditions yield a subjective version of the Anscombe-Aumann framework. The concluding section 5 contains a sketchy methodological discussion of our proposal to consider likelihood relations as independent primitives; it also points out how epistemic constraints on preferences, formalized in terms of likelihood relations, can be given a purely behaviorist interpretation. All proofs can be found in the Appendix, which also contains a discussion of the present work to Epstein-Zhang (2001) and a derivation of relation of the the Anscombe-Aumann framework in the almost-convex-ranged case.

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 $\equiv$, and the asymmetric component by $\vartriangleright$. 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 be disposed to make further probabilistic judgments not listed in $\trianglerighteq$. Note that since this interpretation assumes nothing about the DM's beliefs where $\trianglerighteq$ is incomplete, there is no issue of demarcating a precise boundary between those comparisons where the DM is prepared to make a likelihood judgment and those where he is not. For now, we shall treat the comparative likelihood relation $\trianglerighteq$ as a non-behavioral primitive; we will consider its relation to behavior below in sections 4 and 5.

2.1. Savage's Probability Theorem

As a reference point, we briefly review Savage's Probability Theorem which delivers a unique representation of complete 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) $\Omega \triangleright \emptyset$.

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* $A \cap C = B \cap C = \emptyset$.[7]

Additivity is the hallmark of comparative *likelihood*. Normatively, it can be read as saying that in comparing two events in terms of likelihood, states common to both events do not matter. It is well-known that, on finite state-spaces, Additivity is far

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

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;[8] 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 $\pi(B) = \alpha\pi(A)$.[9] 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 $A \rhd \emptyset$.

Axiom 5 (Fineness) *For any non-null A there exists a finite partition of Ω $\{C_1, ..., C_n\}$ such that for all $i \leq n$, $A \unrhd C_i$.*

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

Theorem 1 (Savage) *Let Σ be a σ-algebra. The likelihood relation $\unrhd$ 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 $A, B \in \Sigma$:*

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

An important feature of Savage's result is the uniqueness of the representing probability. It justifies the view that the comparative likelihood relation captures 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 explicit beliefs, likelihood relations will now allowed to be incomplete.

[8]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.

[9]In the countably additive case axiomatized by Villegas (1964), convex-rangedness is equivalent to the absence of probability atoms.

Axiom 7 (Partial Order)[10] *$\trianglerighteq$ 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 measure:

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

Likelihood relations satisfying (1) will be referred to as *non-contradictory,* and the associated probability measures π as *admissible,* with their set denoted by Π . While condition (1) rules out inconsistencies 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 an admissible prior implying the contrary judgment, as stated by the following condition:

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

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**. For any set of priors $\Pi \subseteq \Delta(\Sigma)$, let $\trianglerighteq_\Pi$ denote the likelihood relation induced by the unanimity condition (3). Coherence entails deductive closure in the sense that, if a set of likelihood judgments $\trianglerighteq' \subseteq \trianglerighteq$ entails another judgment "$C \trianglerighteq D$" assuming

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

completeness, i.e. if $\pi(C) \geq \pi(D)$ for all $\pi \in \Pi\ '$, the coherent relation $\trianglerighteq$ contains in fact this judgment. In practice, explicit beliefs will often be given as a non-contradictory relation $\trianglerighteq$ describing the DM's direct likelihood judgments. From these he can infer others via coherence; the resulting sum total of likelihood judgements is given by the smallest coherent superrelation of $\trianglerighteq$, its *coherent hull* $\overline{\trianglerighteq}$; it is easily verified that $\overline{\trianglerighteq} = \trianglerighteq_{(\Pi\ )}$.

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 class of closed (hence compact), convex subsets of $\Delta(\Sigma)$ be denoted by $\mathcal{K}(\Delta(\Sigma))$. Given $\Pi \in \mathcal{K}(\Delta(\Sigma))$, 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\ $.

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. Here, it will turn out to be sufficient to complement Additivity by the following second rationality axiom called "Splitting".

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

In words: If two events are split into two subevents each, then the more likely subevent of the more likely event is more likely than the less likely subevent of the less likely event. In the following Theorem, we will only make use of the special case in which the two events are split into equally likely subevents. Splitting is made powerful by the following structural 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 $B \equiv A \backslash B$.*

In terms of the multiprior representation, Equidivisibility is satisfied if given any non-null event A there exists an event B with unambiguous conditional probability one half; this would happen, for example, if there exists an event T that is viewed as independent of A and equally likely to its complement, for then $A \cap T \equiv A \cap T^c$. Note that the plausibility of this assumption does not depend on $A's$ being unambiguous itself.

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[11], 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 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 convex-rangedness of Π implies the convex-rangedness of every $\pi \in \Pi$, the converse is far from true. Also note the following Fact.

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

[11]The generality added by allowing Σ to be an algebra is significant since algebras can often be described explicitly while σ-algebras typically cannot. In this vein, Savage's Theorem has recently been extended by Kopylov (2003).

Finally, Savage's Fineness and Tightness axioms are no longer adequate. To obtain a real-valued representation, a condition expressing 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.[12] The following is the main result of the paper.

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 proof idea of Theorem 2 with a bit of "reverse engineering". The key is the derivation of a "mixture-space" structure of the event-space resulting

[12]By contrast, neither Tightness nor Fineness are 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 Savage's Theorem (Theorem 1), one can replace Tightness by Continuity. This has 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.

from the convex-rangedness of the set of priors. Specifically, one can extend every coherent likelihood relation represented by the convex-ranged set of priors Π to a partial ordering on the domain $B_0(\Sigma, [0,1])$ of finite-valued functions $Z : \Omega \to [0,1]$ by associating with each function Z an equivalence class $[Z]$ of events $A \in \Sigma$ as follows. Let $A \in [Z]$ if, for some appropriate partition of Ω $\{E_i\}$, $Z = \sum z_i 1_{E_i}$, and such that, 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:

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

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.

In principle, one could conceive of *coherent expectation orderings* as epistemic primitives. However, such an epistemic interpretation would run into the following two problems. On the one hand, the meaning of a comparison of random variables in terms of their expectation seems intuitively not clear; it seems doubtful that a genuine epistemic primitive can be based on a complex, mathematically structured implicit expectation operation. Moreover, unless one assumes expected-utility maximization (at least relative to the specified ordering), as Walley and Bewley do, the link between expectation orderings and preferences is not clear. Expectation orderings are thus not viable as a behaviorally general vehicle for describing a decision maker's imprecise probabilistic beliefs.

2.3. Examples of Equidivisibility

The key structural assumption behind Theorem 2, 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. Whereas Fineness is in substance a strong non-atomicity condition, Equidivisibility assumes in addition that the likelihood relation is sufficiently complete. 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 Π' of *non-atomic, countably additive* priors. Due to Lyapunov's (1940) celebrated convexity theorem, Π is convex-ranged.

The priors $\pi \in \Pi'$ can be interpreted as a finite set of hypotheses a decision-maker deems reasonable without being willing to assign probabilities to them. Finitely generated sets of priors also occur naturally in social belief aggregation, where $\trianglerighteq_I$

represents the unanimity likelihood ordering induced by the finite set of individuals' likelihood orderings $\trianglerighteq_i$ that are assumed to be precise 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. Then Theorem 2 implies that $\Pi_{(\;\;_I)} = co\{\mu_i\}_{i \in I}$; the "social prior" μ_I must therefore be a convex combination of individual priors.[13]

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 partition $\mathcal{P}$ 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 $A \in \mathcal{P}$, $\pi(./A) = \mu_A$ or $\pi(A) = 0$. Then Π is clearly convex-ranged, however imprecise the DM's beliefs about the events in $\mathcal{P}$ may be.

Example 3 (External Randomization Device)

As a variant of example 2, consider state-spaces with a continuous randomization device in the manner of Anscombe-Aumann. 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

[13] This corollary to Theorem 2 is related to 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.

(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 $i \in I$, $\Omega_1 \times T_i \trianglerighteq_{AA} \Omega_1 \times T_i'$.

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

The example of an external randomization device is important especially because it counters the potential 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 ARE COMPARATIVE LIKELIHOOD RELATION SUFFICIENTLY EXPRESSIVE?

Intuitively, different closed and convex sets 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

$$X \succsim_\Pi Y \text{ iff } E_\pi X \geq E_\pi Y \text{ for all } \pi \in \Pi.^{14}$$

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

Since consequence valuations have been fixed, such differences in preference associated with different sets Π are naturally attributed to differences in beliefs. Thus, for comparative likelihood relations to fully represent a DM's beliefs, their multi-prior representation must be unique.

Furthermore, in general, comparative likelihood relations $\trianglerighteq$ are to be understood as "non-exhaustive", that is: as admitting for the possibility that the DM may be willing to endorse further likelihood judgements not listed in $\trianglerighteq$. It is therefore desirable not merely that the specified relation $\trianglerighteq$ has a unique representation, but that all of its coherent superrelations have unique multi-prior representation as well. We shall refer to such likelihood relations as *guaranteeing uniqueness.* Note that a coherent likelihood relation $\trianglerighteq$ guarantees uniqueness if and only if there is a one-to-one relation between the domain $\mathcal{D}$ of all coherent superrelations of the $\trianglerighteq$ and the closed, convex subsets of Π .

Theorem 2 has already shown that Equidivisibility is sufficient for uniqueness. Since any superrelation of an equidivisible relation is equidivisible as well, Theorem 2 happily shows that Equidivisibility guarantees uniqueness. In this section, we will conversely ask to what extent Equidivisibility can be weakened while preserving guaranteed uniqueness. We will show that while this can be done to a limited extent, uniqueness cannot, in general, be guaranteed without substantial comparability assumptions.

3.1 Example of Non-Uniqueness

Just like complete likelihood relations, incomplete ones typically do not have a unique representation in finite state spaces. Without completeness, it is however no

Indeed, in general, the associated preference orderings can easily be mutually incompatible, in that a likelihood-relation can be represented by two disjoint sets $\Pi, \Pi' \in \mathcal{K}(\Delta(\Sigma))$ such that, for some X, Y, $X \succ_{\Pi} Y$ and $Y \succ_{\Pi'} X$).

longer sufficient to assume non-atomicity in some form as Savage did through his Fineness and Tightness conditions. This is shown by the following example.

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:

$$A \trianglerighteq^K B \text{ if and only if } \lambda(A \backslash B) \geq K\lambda(B \backslash A). \tag{5}$$

It is easily verified that the associated set of admissible priors $\Pi_{(\ \ K)}$ (which we shall also denote as Π_1^K) consists of all probability measures π with Lebesgue density ϕ such that $\operatorname{ess\,sup}_{\omega\in[0,1]} \phi(\omega) \leq K \operatorname{ess\,inf}_{\omega\in[0,1]} \phi(\omega)$;[15] in particular, the extreme points of Π_1^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 $\Pi_2^K \subseteq \Pi_1^K$ denote the closed, convex hull of $\{\pi_D | \lambda(D) = \frac{1}{K+1}\}$; the following Fact states that Π_2^K induces the same likelihood relation $\trianglerighteq^K$. On the other hand, since the set of extreme points of Π_2^K is a strict subset of that of Π_1^K, the two sets are different, demonstrating non-uniqueness. Indeed, Π_1^K and Π_2^K even induce different lower probability functions denoted by $\pi_{1,K}^-$ and $\pi_{2,K}^-$.

Fact 2 *i)* $\trianglerighteq_{(\Pi_2^K)} = \trianglerighteq^K$;

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

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

The lower probabilities $\pi_{1,K}^-(A)$ and $\pi_{2,K}^-(A)$ are shown in the following figure as functions of $\lambda(A)$ for $K = 3$.

---

[15] ess sup and ess inf denote the essential supremum and essential infimum, respectively.

Fig. 1: Two Different Lower Probabilities

For $K > 1$, $\trianglerighteq^K$ clearly satisfies Savage's Fineness and Tightness conditions. Note that if K is close to 1, all admissible probabilities are uniformly close to the Lebesgue measure; nonetheless, uniqueness is lost. Also, if K is close to 1, $\trianglerighteq^K$ is close to being equidivisible, and $\Pi_{(\ K)}$ is close to being convex-ranged[16], but, for given K, it does not come "arbitrarily closely". We will now show that if a coherent likelihood relation $\trianglerighteq$ comes arbitrarily close to being equidivisible, its multi-prior representation comes arbitrarily close to being convex-ranged, and that this suffices for uniqueness.

3.2 Almost-Equidivisibility Implies Uniqueness

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

[16] This can be made formally precise.

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 $\pi \in \Pi$.*

Note that for singleton sets $\Pi = \{\pi\}$ defined on a sigma-algebra Σ, almost-convex-rangedness and convex-rangedness coincide, since both are equivalent to convex-rangedness of π (cf. Fact 1); as shown in the next subsection, this ceases to be the case for non-singleton Π. [17]

The following result establishes a one-to-one relation between coherent almost-equidivisible likelihood relations and almost-convex-ranged sets of priors.

Theorem 3 *i) If $\trianglerighteq$ is coherent and almost-equidivisible, Π 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 second part of this result is proved by showing through an approximate limiting 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 $\trianglerighteq$ is unique as well. The continued applicability of the mixture-space construction

[17] If Σ is merely an algebra, $\{\pi\}$ is almost-convex ranged if and only if π is dense-ranged in the sense of Kopylov (2003).

suggests that it might be possible to generalize Theorem 1 to almost-equidivisible likelihood relations. We have left such an extension to future research, as it is far from straightforward.

While the example of section 3.1 suggest that almost-convex rangedness cannot be greatly weakened in general without losing uniqueness, it is easily seen not to be strictly necessary[18]; it is thus another interesting task for future research to determine exactly the conditions ensuring uniqueness respectively a guarantee of uniqueness.[19]

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

In many situations, it is reasonable to assume random quantities to be continuously distributed. For example, the subjective probability of any reasonable, contemporary human DM (assuming him to have precise probabilistic beliefs, for the sake of the argument) over the true temperature (idealized as real-valued magnitude) in Seoul on August 15, 2006, would be described by a continuous density function. Poincaré (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 the decimal expansion of ω is odd; thus A_k is the evenly spaced union of $\frac{1}{2}10^k$

[18]As a simple counterexample in the complete case, consider the likelihood relation that judges each of a finite number of states equally likely.

[19]The uniqueness issue has been studied so far only in the complete case; see, for example, Fishburn (1986) and Cohen (1991).

[20]Quoted from Machina (2001).

intervals of length $\varepsilon_k := 10^{-k}$, and has Lebesgue measure $\frac{1}{2}$. Poincaré obtained the following result:

Proposition 1 (Poincaré) *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 assumption 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 as the corollary of a result for general metric spaces.

Let Ω be a compact metric space endowed with metric d and a non-atomic, countably additive "reference-measure" λ on the Borel-σ-algebra Σ; without loss of generality, assume that λ has full support. To gain some useful generality, take the event space to be any subalgebra Σ_0 containing all open ε-balls. Let $\Delta(\Sigma_0)_{cont}$ denote the set of all measures π with continuous density; for each $\pi \in \Delta(\Sigma_0)_{cont}$, let ϕ_π denote its unique continuous density function. For example, if Ω is the unit interval, Σ_0 may the set of all finite interval-unions, the case considered in Machina (2001).

Proposition 2 *Suppose that $\Pi \subseteq \Delta(\Sigma_0)_{cont}$ has equicontinuous densities*[21] *that are uniformly bounded below strictly above zero. Then Π is almost-convex-ranged.*

Proposition 2 goes beyond Machina (2001) by modelling the postulated "uniform continuity" of probabilistic beliefs explicitly as a likelihood relation. This leads to behaviorally general restrictions on preferences analogous to 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 assumptions of

[21] That is, for all $a \in \Omega$ and $\epsilon > 0$, there exists $\eta > 0$ such that, for all $\pi \in \Pi$, $|\phi_\pi(b) - \phi_\pi(a)| \leq \epsilon$ whenever $\delta(b,a) \leq \eta$.

event-differentiability of preferences[22] together with a Euclidean structure of the state space.

A limitation common to both Proposition 2 and Machina (2001) is their imposition of a requirement on the set of admissible priors, and not *directly* on the likelihood relation as the epistemic primitive. This limitation is not necessarily fatal, in that for reasons of mathematical tractability, in practice likelihood relations will typically be specified directly in terms of sets of priors rather than the associated likelihood relation, just as in the standard complete case subjective uncertainty is typically modeled directly in terms of the subjective probability measure rather than the likelihood relation. Indeed, in many cases, it will be difficult to describe explicitly the likelihood content of a specified sets of priors. Nonetheless, in the case of fundamental qualitative assumptions on beliefs, an explicitly description in terms of the epistemic primitive seems important for a full understanding and assessment; a famous example in the single-valued case is de Finetti's (1937) celebrated exchangeability theorem.

In the present case, we will show now that such an explicit description is possible if the restriction to uniformly continuous densities takes a particular, natural form. Specifically, we will consider sets $\Pi_{\lambda,M} \subseteq \Delta(\Sigma)_{cont}$ consisting of all priors with uniformly Lipschitz log-densities with modulus of continuity M. Formally, define

$$\Pi_{\lambda,M} := \{\pi \in \Delta(\Sigma)_{cont} | \text{for all } a, b \in \Omega : |\log \phi_\pi (a) - \log \phi_\pi (b)| \leq M d(a, b)\}.$$

Note that uniform equicontinuity of log-densities implies uniform boundedness of densities above zero, and is therefore equivalent to the assumption on densities in Proposition 2 above. Uniform Lipschitz continuity is only slightly stronger; the main additional restriction on the structure of $\Pi_{\lambda,M}$ is thus the requirement that it include all densities of a given modulus of continuity M, and not just some. This seems appropriate to capture the "uniform continuity" assumption as such; due to the almost

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

convex-rangedness of $\Pi_{\lambda,M}$, further restrictions on $\Pi \subset \Pi_{\lambda,M}$ can be captured by appropriate enlargements of the likelihood relation $\trianglerighteq_{\lambda,M}$ generating $\Pi_{\lambda,M}$ that is about to be defined.

Let $\Psi(A,B) := \sup_{a\in A, b\in B} e^{Md(a,b)}$, and define a comparative likelihood relation $\trianglerighteq_{\lambda,M}$ as follows. For any A, B :

$$A \trianglerighteq_{\lambda,M} B :\Leftrightarrow \quad \lambda(A) \geq \Psi(A,B)\lambda(B). \tag{6}$$

Intuitively, $\trianglerighteq_{\lambda,M}$ expresses uniform continuity by comparing the likelihood of "small" events (as measured by their diameter) that are close to each other (in the Hausdorff metric, say) by their reference measure admitting a small degree of imprecision that vanishes in the limit. These are substantive likelihood judgments which may or may not be reasonable in a particular (idealized) setting.[23] The following Proposition shows that $\Pi_{\lambda,M}$ is indeed generated by the likelihood relation $\trianglerighteq_{\lambda,M}$.

Proposition 3 $\Pi_{\lambda,M} = \Pi_{(\ \ _{\lambda,M})}$.

It is straightforward to verify that $\trianglerighteq_{\lambda,M} \subseteq \trianglerighteq_{(\Pi_{\lambda,M})}$, i.e. that $\trianglerighteq_{\lambda,M}$ contains only comparative likelihood judgements induced by $\Pi_{\lambda,M}$; this evidently implies that $\Pi_{(\ \ _{\lambda,M})} \supseteq \Pi_{\left(\ \ _{(\Pi_{\lambda,M})}\right)} = \Pi_{\lambda,M}$. The converse, which says that $\trianglerighteq_{\lambda,M}$ is sufficiently rich to exclude all priors outside $\Pi_{\lambda,M}$, is not self-evident, and takes some work to prove.

[23] For instance, if temperature is measured digitally and digits are displayed sequentially, and if there is a positive lower probability that the diode used malfunctions by not being able to display a particular number, small events missing that number in their complete decimal expansion may have positive lower probability but zero Lesbesgue measure, leading to a contradiction of (6) respectively the non-existence of a Lesbesgue density for any admissible prior. In this modified situation, almost-convex-ranged beliefs may nonetheless be reasonable.

4. DECISION MAKING IN THE CONTEXT OF PROBABILISTIC BELIEFS

4.1. Probabilistic Consequentialism

Consider now a DM described by a preference ordering over acts and explicit beliefs over events. Let X be a set of *consequences*. An *act* is a mapping from states to consequences, $f : \Omega \to X$ that is measurable with respect to an algebra of events Σ; the set of all acts is denoted by $\mathcal{F}$; for simplicity, we will assume all acts to be finite-valued throughout. 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.

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 others not included in it. The relation $\trianglerighteq^0$ will be referred to as the "*epistemic context*" of the decision situation. Thus, a *decision-maker in an epistemic context* is described by the pair $(\succsim, \trianglerighteq^0)$. We will say that the context is convex-ranged if it has a convex-ranged multi-prior representation.

We propose as a fundamental principle of consequentialist rationality that consequence valuations and likelihood comparisons, when available, should be decisive in determining the ranking of acts; put somewhat differently, the judged (comparative) likelihood of events is teh only attribute of events that should matter in comparing the consequence incidences $f^{-1}(\{x\})$ and $g^{-1}(\{x\})$ of the various consequences of different acts; other conceivable factors such as familiarity with a type of event or felt competence in assessing it *should* not matter rationally. We shall refer to this as the Principle of Probabilistic Consequentialism. The task is to formalize this principle in terms of various axioms on the relation between preferences and beliefs.

The following axiom called "Stochastic Equivalence" does not exploit any information about consequence valuations. Say that acts f and g are *stochastically equivalent* ("$f \sim^0 g$") if $f^{-1}(\{x\}) \equiv^0 g^{-1}(\{x\})$ for all $x \in X$.

Axiom 12 (Stochastic Equivalence) *$f \sim g$ whenever f and g are stochastically equivalent.*

Stochastic Equivalence presupposes that the same consequence "x" really is the same in the different states in all relevant aspects (that consequences are "properly individuated"); it thus does not apply to state-dependent preferences. It is obvious that Stochastic Equivalence implies probabilistic sophistication of preferences over unambiguous acts. [24]

State-independent preferences are typically event-wise monotone.

Axiom 13 (Eventwise Monotonicity) *For all $x, y \in X$, $h \in \mathcal{F}$, and $A \in \Sigma$, $x \succsim y$ implies $x_A h \succsim y_A h$.*[25]

Again, with proper individuation of consequences, this represents a fundamental principle of consequentialist rationality. However, Grant (1995) has pointed out that Eventwise Monotonicity makes an implicit separability assumption that may be violated in some economic settings[26] in which preferences are nonetheless proba-

[24] Formally, let Λ^0 denote the family of unambiguous events associated with Π^0 defined in A.1, and let π^0 denote the restriction of π to Λ^0, for any $\pi \in \Pi^0$; by construction, the set-function π^0 is finitely additive on Λ_0. An act f is unambiguous if $f^{-1}(\{x\}) \in \Lambda^0$ for all $x \in X$; let $\mathcal{F}^0$ denote their set. Stochastic Equivalence implies the following property:

$$\text{For all } f, g \in \mathcal{F}^0 \text{ such that } \pi^0\left(f^{-1}(\{x\})\right) = \pi^0\left(g^{-1}(\{x\})\right) \text{ for all } x \in X,\ f \sim g.$$

[25] Note that Eventwise Monotonicity has been formulated uni-directionally for cleaner statements in the sequel; it is thus slightly weaker than Savage's axiom P3.

[26] in which the individuation is coarser than that relevant to the DM.

bilistically sophisticated. To accommodate such situations, we have formulated the Stochastic Equivalence axiom above.

If preferences are eventwise monotone, the principle of Probabilistic Consequentialism motivates the following monotone strengthening of Stochastic Equivalence. Say that f *stochastically dominates* g ("$f \succsim^0 g$") if, for all $x \in X$, $\{\theta \mid f(\theta) \succsim x\} \trianglerighteq^0 \{\theta \mid g(\theta) \succsim x\}$.

Axiom 14 (Stochastic Dominance) *$f \succsim g$ whenever f stochastically dominates g.*[27]

It is immediate that Stochastic Dominance implies in particular that preferences over unambiguous acts respect ordinary stochastic dominance and are therefore probabilistically sophisticated in the sense of Machina-Schmeidler (1992). Stochastic Dominance clearly implies Stochastic Equivalence and Eventwise Monotonicity.[28] The converse holds if explicit beliefs satisfy the following solvability condition:

$$\text{For all } A, B \text{ such that } A \trianglerighteq^0 B \text{ there exists } A' \subseteq A \text{ such that } A' \equiv^0 B. \tag{7}$$

This condition in turn is satisfied if explicit beliefs are convex-ranged and complete, but may easily fail otherwise.

[27] A formally related axiom called "Cumulative Dominance" has been introduced by Sarin-Wakker (1992); in their axiom, Savage's revealed likelihood relation takes the place of $\succsim^0$ here. In Sarin-Wakker (1992), this relation is complete but not necessarily additive. Sarin-Wakker (2000) derive probabilistic sophistication from cumulative dominance, assuming this relation to be complete and coherent.

[28] Indeed, Eventwise Monotonocity is equivalent to Stochastic Dominance with respect to the set-inclusion relation.

4.2 A Subjective Interpretation of the Anscombe-Aumann Framework

Probabilistic Consequentialism will now be combined with the mixture-space constructions of sections 2 and 3 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 outright suspicion (see, e.g., Epstein (1999)). In the body of the text, we will consider the simpler convex-ranged case, and offer a generalization to the almost-convex-ranged case in the Appendix.

The Anscombe-Aumann (1963) framework is distinguished by taking acts to be mappings from states to probability distributions of consequences, rather than simply as mappings from states to consequences as in the Savage (1954) framework. These probability distributions are interpreted as objective probabilities of the realizations of an external random device ("roulette lotteries") that is not part of the explicitly modeled state space. We will show that if a preference relation over Savage acts satisfies Stochastic Equivalence with respect to a convex-ranged likelihood relation, it can be canonically extended to a preference relation over Anscombe-Aumann (AA-) acts; this extension is *internal*, that is: it does not rely on the 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 $F = [q_1, A_1; q_2, A_2; ...]$ in analogy to the notation for Savage acts. Given a convex-ranged epistemic context $\trianglerighteq^0$, any AA-act F can be identified with a class $[F]$ of Savage acts by the following stipulation: $f \in [F]$ if, for any $x \in X$, any i such that

A_i is $\equiv^0$-non-null, and all $\pi \in \Pi^0$,

$$\pi(f^{-1}(\{x\}) \cap A_i/A_i) = q_i^x.$$

Thus $[F]$ consists of all Savage acts that yield the consequence probabilities specified by F as unambiguous conditional probabilities with respect to the given context. Convex-rangedness of the context ensures that $[F]$ is non-empty. Moreover, any two acts in $[F]$ are easily seen to be stochastically equivalent with respect to the context $\trianglerighteq^0$. Hence Stochastic Equivalence ensures that any two acts in $[F]$ are indifferent. Thus one obtains a well-defined weak order on $\mathcal{F}^{AA}$ by setting

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

Note that, 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 now replicated in multiple copies.

The above construction 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, Eventwise Monotonicity of AA-preferences is much stronger than Eventwise Monotonicity of Savage preferences.[29]

[29] To see this, one can define for each F_{-A_1} $[., A_1; q_2, A_2; q_3, A_3; ...]$ preferences over A_1-conditional lotteries q by setting $q \succsim_{F_{-A_1}} q'$ iff $[q, A_1; q_2, A_2; q_3, A_3; ...] \succsim [q, A_1; q_2, A_2; q_3, A_3; ...]$. If $\succsim$ satisfies Stochastic Dominance, $\succsim_{F_{-A_1}}$ is monotone with respect to ordinary stochastic dominance; however, unless the underlying preference relation $\succsim$ is utility sophisticated in the sense of Nehring (2001), $\succsim_{A_1}$ may well differ from preferences over unconditional lotteries $\succsim_{F_{-\Omega}}$, that is, $\succsim^{AA}$ may not satisfy Eventwise Monotonicity even though $\succsim$ does.

Earlier representations of the Anscombe-Aumann framework in a Savage setting have been obtained by Pratt-Raiffa-Schleifer (1964) and Klibanoff (2001a); in contrast to ours, the former assumes expected utility maximization, the latter utility sophistication. Our representation is more general also in that it applies to any convex-ranged context, and therefore does not assume the existence of a subjective randomization device as given by the context $\trianglerighteq_{AA}$ defined in section 2. If the epistemic context is almost-convex-ranged rather than convex-ranged, the above construction fails to work because the equivalence classes $[F]$ may easily be empty. In Appendix A.2 we show that it can be rescued by associating with a given AA-act equivalence classes of convergent sequences of Savage acts; this yields a counterpart to the construction of Machina (2001).

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. (2001); since the mixture operation in their proposal is defined in utility terms, the implied interpretation of conditions on AA-preferences in their approach may be very different from that in the present epistemically approach.

4.3. Discussion

Ordinal vs. Cardinal Information about Consequence Valuations.—

From the well-known utility characterization of ordinary stochastic dominance it follows immediately that 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 $x \succsim y$. Stochastic Dominance is thus the strongest rationality requirement on weak preference that relies on ordinal information about the valuation of consequences only. Much stronger normative restrictions can be obtained if one exploits cardinal information about comparisons of utility differences. A theory of "Expected Utility in the Presence of Ambiguity" along such lines is developed in Nehring (2001, section 4). It is not pursued here, since these stronger constraints on the relation between preferences and beliefs conflict with the desideratum of behavioral generality adopted in this paper.

State-Dependent Preferences and Intrinsic Event Attitudes.—

While the non-exploitation of cardinal information suggests that Stochastic Dominance may be too weak a rationality requirement, Stochastic Dominance might conversely be criticized as to be too strong a rationality requirement, for example for reasons of state-dependence of preference. In this subsection, we will however argue that in many (perhaps in all) cases, such objections can be overcome by a more refined modeling.

Indeed, as formulated above, the Stochastic Equivalence and Stochastic Dominance axioms presuppose "proper individuation of consequences", in the sense that consequence x is the same (for the decision-maker) in any state *in all valuation-relevant aspects*[30]; we will refer to this also as the case of "state-independent preferences"[31]. State-dependence of preferences can be modeled by including aspects of the state into the description of the consequences. For example, when states describe the DM's health, gaining 1000$ when healthy may not be worth as much as gaining 1000$ when sick; this difference can be accounted for by distinguishing the consequences (1000$,healthy) and (1000$,sick). In such cases, one needs to abandon the assumption that the set of possible consequences is state-independent; that is, one would define Savage acts as mappings from $f : \omega \to \prod_{\omega \in \Omega} X_\omega$. Note that Stochastic Equivalence remains meaningful but loses some power; it becomes vacuous only in the case of "fully subjective" consequence individuation in which $X_\omega \cap X_{\omega'} = \varnothing$ whenever $\omega \neq \omega'$.

[30]As argued forcefully by Broome (1991), the validity of *any* normative axiom hinges on the proper individuation of consequences.

[31]This terminology is somewhat misleading, however, as there is no established preference-based criterion outside SEU defining state-independence; this is further discussed below.

In this latter case, one arrives at substantive restrictions on preferences by postulating a non-behavioral *judged consequence-ordering* $\succsim_{[X]}$ on $X = \cup_{\omega \in \Omega} X_\omega$ that permits cross-state comparisons of consequences. A number of works on state-dependent preferences from Fishburn (1973) to Karni (2003) employ devices that entail such an ordering. The Stochastic Dominance condition continues to be meaningful, with $\succsim_{[X]}$ taking the role of $\succsim_{|X}$.

State-dependence of Preference is of interest in the present context especially since it allows to capture many – and arguably all – instances of an apparent "intrinsic" attitude towards events. For instance, Chew and Sagi (2003), suggest that decision makers may have a taste for betting on particular types of events over others that override their likelihood assessments. For example, on February 1, 2003, a DM may have attributed equal probability to Saddam Hussein's surviving a US led invasion of Iraq, and to the Iraqui soccer team winning against Brazil. However, the DM may have preferred to bet on the Iraqui soccer team rather than on Saddam Hussein; he may, for example, have expected taking special joy from winning a bet on an underdog team, but regretting having profited from an unjust cause. On their face, such preferences might appear to challenge the normative generality of the principle of Probabilistic Consequentialism in that non-likelihood features of events seem to be legitimately valued by the DM. However, this challenge loses its force once one recognizes that the bets do not really involve the same (properly individuated) consequences. For clearly, the bets entail different psychological outcomes in various states (the joy, the regret); while these matter to the decision maker, they are not captured by a description ("individuation") of consequences in terms of net wealth alone. Again, Stochastic Equivalence and Dominance continue to apply once acts that are described in terms of properly individuated consequences.

We have appealed to the informal notion of "proper individuation"/"state-independent preferences" without characterizing it in behavioral terms. This seems necessary, since it is unclear what behavioral conditions could take its place. Clearly, Eventwise Monotonicity is not sufficient as a behavioral criterion, since it captures at most the ordinal implications of state independence. In the context of the sure-thing principle, assuming in addition Savage's axiom P4 "works", as shown by Savage's representation theorem. However, in situations of ambiguity in which the sure-thing principle fails, P4 seems neither necessary nor sufficient. It is not necessary because P4 may fail due to prize-dependent ambiguity attitudes[32]; it does not seem sufficient, either, for in the absence of the sure-thing principle, there may be state-dependence effects that show up in multi- but not two-valued acts. Despite these reservations, it would appear that the conjunction of Eventwise Monotonicity and P4 represents as reasonable "prima facie" criterion of proper individuation, and hence of the applicability of the Stochastic Dominance axiom.[33]

5. EPISTEMIC VERSUS BEHAVIORIST INTERPRETATION

In this section, we will flesh out the notion of decision-making in the context probabilistic beliefs first in terms of the epistemic interpretation adopted as the "official" one in this paper, and subsequently point out how this approach can be understood in purely behaviorist terms. On the latter view, probabilistic beliefs must be definable in terms of behavior, while on the former, they represent independently meaningful entities. The purpose of the following discussion is not to settle the deep and long-standing philosophical disputes around these views, but to clarify what they involve.

[32] See Klibanoff et al. (2002) for a worked-out model with this feature.

[33] Adapting a recent argument by Karni (1996), it can be argued that more than a "prima-facie" criterion can be given in principle, even under expected utility: roughly speaking, Karni argued that under SEU, preferences identify statewise utilities only up to positive affine transformations *state-by-state*; thus, if the "true" consequence utilities are not constant across states, Savage's "revealed likelihood" relation differs from the agent's true likelihood. In the present setting, this amounts to saying the Stochastic Equivalence will be violated.

5.1 The Epistemic Interpretation

On a properly *epistemic view*, the context $\trianglerighteq^0$ represents a (non-exhaustive) list of comparative likelihood judgements attributed to the decision maker. More specifically, the instance "$A \trianglerighteq^0 B$" represents the DMs *disposition* to judge that A is at least as likely as B, that is: to affirm this proposition when asked for the comparison. The epistemic approach therefore presumes that the DM is able to understand and use the notion of "comparative likelihood"; it would thus appear to exclude rats, for example, but include humans almost without exception – we are not aware of neurological or anthropological evidence to the contrary. The epistemic approach allows one to think of beliefs as *reasons* for choice; the beliefs themselves may in turn be derived from underlying evidence, or may be taken as subjective judgments without identified basis. Note, however, that it does not assume beliefs to be temporally or causally *prior* to preferences. As formulated, the epistemic framework is atemporal (in the way Savage's framework is), and axioms such as Stochastic Equivalence and Stochastic Dominance are consistency conditions relating beliefs and preferences; if these axioms are violated, a DM may well come to revise his beliefs in light of his preferences, concluding, for example, that the former are based on sloppy or wishful thinking, while the latter reveal his true convictions. Thus the epistemic approach is entirely consistent with the *psychological intuitions* that historically have been part of the appeal – one might even say, of the wisdom – of the behaviorist view, the intuitions that "actions (may) speak louder than words" and that "you may not know what you believe before you act" .

Is an epistemic approach less "scientific" than a purely behavioral one? While this is not the place to address this issue in a satisfactory manner, we want to suggest that the answer is not *obviously* "yes", and that indeed a case can be made for the converse. First of all, since an epistemic context consists in judgments that can be recorded in language, that is: consists in *speech behavior* of a particular kind, it does not rely on introspection, and is no less verifiable in principle than choice behavior.[34] Of course, in practice there are elicitation issues of truthfulness, incentives and correct understanding, but these issues arise in choice experiments as well, especially if the stakes are low or moderate.

Moreover, the epistemic approach endeavours to model *rational* choice and belief; Probabilistic Consequentialism is therefore part and parcel of it. If Probabilistic Consequentialism is adopted as a maintained hypothesis, belief attributions become *falsifiable* on the basis of choice behavior through their Stochastic Dominance implications.[35] In effect, assumptions about beliefs translate into assumptions about preferences of the form

$$\succsim \supseteq \succsim^0 = \succsim^0 (\trianglerighteq), \tag{8}$$

where $\succsim^0 = \succsim^0 (\trianglerighteq)$ denotes the Stochastic Dominance implications of the epistemic context $\trianglerighteq$ together with the preference ordering over consequences $\succsim_{|X}$.

5.2 The Behaviorist Interpretation

On the behaviorist view of decision theory, the only meaningful attribute of a decision maker is his choice behavior, captured here by his preferences. Attributions of belief are considered meaningful only to the extent that they can be formulated in terms of conditions on agents' preferences. Evidently, the epistemic constraints on preferences (8) are of this form. Thus, the notion of decision making in the context of probabilistic beliefs is meaningful also from a behaviorist viewpoint. In contrast to the epistemic interpretation, the relation $\trianglerighteq$ (or $\succsim^0 = \succsim^0 (\trianglerighteq)$) is no longer an indepen-

[34]See also Karni (1996) for a defense of the use of verbal testimony in the decision sciences.

[35]There is likely to be disagreement about how damaging reliance on a "maintained hyposthesis" is; suffice it here to say that, according to the celebrated if controversial Duhem-Quine thesis, all testing in science is the testing of joint hypotheses.

dently meaningful entity about which truth-claims can be made; instead, it plays now the role of a "parameter" that indexes the purely behavioral conditions "$\succsim \supseteq \succsim^0 (\trianglerighteq)$", whose value $\trianglerighteq = \trianglerighteq^0$ is selected by the outside analyst. Examples of behaviorally formulated epistemic constraints on preferences of this type are de Finetti's (1937) classical notion of exchangeability and Machina's (2001) recent notion of "almost-objective uncertainty". Indeed, exchangeability which is commonly stated as a symmetry condition on $\succsim$ can easily be reformulated as an epistemic constraint of the form (8), with the underlying context $\trianglerighteq = \trianglerighteq_\Pi$ generated by the set of all exchangeable priors. Likewise, while Machina formulated his notion of "almost-objective uncertainty" ("uniformly continuous densities", cf. section 3.3) as a smoothness-condition on preferences, it follows from our analysis in sections 3.3 and A.2 that one can capture this notion by conditions of the form (8) as well.

The entailed epistemic constraints (8) serve as a common denominator of the epistemic and behavioral interpretations, and show that the difference between them need not be significant *in practice*. Instead, the difference seems at heart mainly philosophical in nature. While behaviorism has the clear advantage of ontological parsimony, it suffers from an equally clear *explanatory deficit*: why should one assume that a decision-maker's preferences look as-if they were consistent with a particular hypothesized set of likelihood comparisons, if was meaningless *in principle* to attribute any, and thus a fortiori these, likelihood comparisons to the DM as his beliefs. It is chiefly due to this explanatory deficit that we have adopted an epistemic interpretation as the "official" viewpoint of this paper.

One potential avenue for overcoming the explanatory deficit within a behaviorist viewpoint is to try*define* a DM's "revealed beliefs" $\trianglerighteq^*$ from his preferences in an attempt to generalize Savage (1954). Savage substantially strengthened the force of the behaviorist view by identifying a DM's beliefs with a particular aspect of his preferences. A natural way of trying to achieve the same in the present, more general setting is to use condition (8) not merely as a necessary, but as a necessary *and sufficient* condition for attributing particular beliefs to the DM, that is, to define $\trianglerighteq^*$ as the union of all contexts such that the given preference relation $\succsim$ is compatible with that context. This works if $\succsim$ is a Savage-style SEU ordering, and yields exactly the revealed likelihood relation defined by Savage in this case. In the presence of ambiguity, however, this definition will often fail, even if preferences are well-behaved, since $\trianglerighteq^*$ may fail to be coherent or even non-contradictory. This is illustrated by the following example.

Example. Consider a product space with an external random device $\Omega = \Omega_1 \times \Omega_2$ in the manner of section 2.3 with $\#\Omega_1 = \{\alpha, \beta\}$, and assume that preferences have a minimum expected-utility representation, with

$$f \succsim g \text{ iff } \min_{\pi \in \Pi} E_\pi (u \circ f) \geq \min_{\pi \in \Pi} E_\pi (u \circ g),$$

taking Π to be any non-singleton subset of Π_{AA} that is symmetric in α and β. By construction, $\succsim$ satisfies (8) with respect to $\trianglerighteq_{AA}$. Setting $A = \{\alpha\} \times \Omega_2$, clearly $\min_{\pi \in \Pi} \pi(A) = \min_{\pi \in \Pi} \pi(A^c) < \frac{1}{2}$. It is easily verified that $\succsim$ is satisfies (8) also with respect to $\{\overline{(A, A^c)}\}$, representing the judgment that A is equally likely to its complement. The coherent hull of $\trianglerighteq_{AA}$ and $\overline{\{(A, A^c)\}}$ is complete and represented by the product measure of $\pi_1 = (\frac{1}{2}, \frac{1}{2})$ and $\overline{\pi_2}$. However, $\succsim$ clearly violates (8) with respect to this relation ($\overline{\trianglerighteq_{AA} \cup \{(A, A^c)\}}$), since $\succsim$ is not even probabilistically sophisticated. Thus $\trianglerighteq^*$ cannot be coherent.

The example reveals a fundamental difficulty in defining revealed beliefs from preference information alone. It is precisely this difficulty that motivated the more modest goal of this paper of formulating epistemic constraints on preferences based on given epistemic contexts.[36] A proposal of how to overcome it is offered in Nehring (2001, section 3).

[36] Responding to related issue, Chew-Sagi (2003) argue in a recent paper that it is not possible

Finally, while it is sometimes argued that a behaviorist approach to decision theory is *de rigueur* (see, e.g., Epstein/Zhang 2001), it should be kept in mind that its reach is limited. In particular, there seems to be general agreement (including by Epstein/Zhang 2001) that the notions of "state" and "act" in the Savage-framework are not fully behavioral. We would conclude that, whatever its appeal as a methodological guideline, the behaviorist view is unlikely to be coherent as an over-arching philosophical view of what decision theory is or should be.

to identify a canoncial domain of probabilistic beliefs (probabilistically sophisticated behavior), and argue for the coexistence of multiple domains.

APPENDIX

A.1 Unambiguous Events

Comparison to Zhang (1999).—

An event A is *unambiguous* ($A \in \Lambda$) if all admissible priors assign the same probability to it: $A \in \Lambda$ if $\pi(A) = \pi'(A)$ for all $\pi, \pi' \in \Pi = \Pi$. Clearly, the family of unambiguous events Λ^0 is closed under complementation and finite disjoint union; requiring closure under countable disjoint unions and adopting measure-theoretic terminology , Zhang (1999) called such families "lamba-systems". Consider the restriction $\trianglerighteq_{|\Lambda}$ of any likelihood relation $\trianglerighteq$ satisfying the assumptions of Theorem 2 on Σ to the family of unambiguous events $\trianglerighteq_{|\Lambda}$. By construction, $\trianglerighteq_{|\Lambda}$ is complete (on Λ) ; Theorem 2 implies that $\trianglerighteq_{|\Lambda}$ can be represented by a (dyadically) convex-ranged finitely additive set function π^0 on Λ; moreover, π^0 can be extended to a finitely additive probability measure π on all of Σ.

Zhang (1999) considered likelihood relations defined on arbitrary lambda-systems as primitives and characterized those relations that are representable by a convex-ranged, countably additive set function on Λ. Zhang's result is a key ingredient in Epstein-Zhang's (2001) characterization of revealed unambiguous events discussed below. His result is not directly comparable to the corollary to Theorem 2 described in the preceding paragraph, as it derives a weaker conclusion from weaker premises, speaking broadly. Zhang's assumptions are weaker in that they apply only to Λ and not to (and incomplete relation defined on) some super-algebra Σ; on the other hand, his result conclusion is weaker as well in that it does imply representation by an additive set-function that can be extended to all of Σ. It is not known under which conditions such an extension exists;[37] Epstein (1999) and Nehring (1999) provide

[37] Beyond applying the general result of Heath-Suddert (1972), which applies to arbitrary families of sets and therefore does not exploit the structure of lamba-systems at all.

counterexamples in finite state-spaces. In cases in which such an extension does not exist, then the likelihoood relation $\trianglerighteq_\Lambda$ (viewed as an incomplete relation on Σ) is contradictory in the terminology of section 2.2; it stands to reason that such likelihood relations do not represent a well-defined set of probabilistic beliefs.

Comparison to Epstein-Zhang (2001).—

With this background, we will now compare the definition of unambiguous events based on explicit beliefs as Λ^0 ("explicitly unambiguous events") to the preference-based definition proposed by Epstein-Zhang (2001) Λ^{EZ}. A central issue arising from any behavioral definition is the extent to and sense in which preferences over unambiguous acts reveal the decision-maker's unconditional probabilistic beliefs. Epstein-Zhang (2001) provide a partial answer by showing under fairly weak assumptions on preferences and richness assumptions on the endogenously defined family Λ^{EZ} that preferences over unambiguous acts are probabilistically sophisticated with respect to an additive set-function on Λ^{EZ}. Note, however, that, the Epstein-Zhang (2001) definition of "revealed unambiguous" applies conceptually (and is meant to apply by Epstein and Zhang) whether or not the resulting family Λ^{EZ} is rich; in particular, it applies on finite state-spaces as well.

For expositional simplicity, we shall confine the following discussion to the case of two consequences ($\#X = \{x, y\}$ with $x \succ y$). In this setting, the act "betting on A" [x on A, y on A^c] will be denoted simply by the set A. Using this notation, an event T is *EZ-unambiguous* ($T \in \Lambda^{EZ}$) if, for all A, B disjoint from T, $A \succsim B$ if and only $A + T \succsim B + T$, and if the same holds for T^c instead of T.

Two questions arise naturally. First, if preferences satisfy Stochastic Dominance with respect to explict beliefs, are the explicitly unambiguous events also EZ-unambiguous; in other words: will "truly" unambiguous events be revealed as such by the EZ definition? Not necessarily; indeed, this happens only if betting preferences and beliefs are related by the following "Union Invariance" condition: for all $T \in \Lambda^0$ and all A, B disjoint from T, $A \succsim B$ if and only $A + T \succsim B + T$. While this condition looks appealing, it is clearly a substantive restriction on ambiguity attitudes.[38]

Conversely, given a preference relation $\succsim$ with EZ-unambiguous events Λ^{EZ}, does there exist necessarily exist a coherent epistemic context $\trianglerighteq$ such that $\Lambda^{EZ} \subseteq \Lambda^0$ and such that $\succsim$ respects Stochastic Dominance with respect to $\trianglerighteq$? Again the answer appears to be negative in general , and is definitively negative in finite settings. First, in view of the discussion of Zhang (1999) above, the likelihood relation revealed on Λ^{EZ} may be contradictory, hence there simply may not exist any coherent context $\trianglerighteq$ with associated Λ^0 such that $\Lambda^{EZ} \subseteq \Lambda^0$. Second, even if such a context exists, it will contain a large number of likelihood comparisons over ambiguous events, with entailed restrictions on preferences over bets on ambiguous events. While the EZ definition entails some restrictions on preferences over bets on ambiguous events, it seems doubtful that these encompass all of $\trianglerighteq$ in general. Thus, it remains an interesting question for future research to determine under which conditions a DM's preferences over EZ-unambiguous events reflect genuine (= coherent) probabilistic beliefs over these events.

At a broader level, we believe that the limited match between EZ's behavioral definition of unambiguous ones and the direct epistemic one proposed here indicates inherent limitations in defining beliefs from preferences *directly,* rather than inadequacies in the particular definition proposes by Epstein-Zhang, or a need to reconsider the notion of coherence. This conlusion agrees with and confirms the argument in section 5.2 in favor of a dual preference-belief framework.

---

[38] In Nehring (2001), this condition is derived from utility sophistication and Savage's P4.

In a similar vein, Klibanoff et al. (2002) have pointed out that the Epstein-Zhang definition makes substantive implicit assumptions about the decision-maker's ambiguity attitudes.

A.2 Derivation of the Anscombe-Aumann Framework for Almost-Convex-Ranged Contexts

If the context $\trianglerighteq^0$ is merely almost-convex-ranged, the construction of AA-acts and preferences given in section 4.2 fails because AA-acts may have no exact Savage counterpart, that is: $[F]$ may be empty for some F. The natural remedy is to interpret AA-acts $F = [q_i, A_i]_{i \in I}$ 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 A_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 appropriate 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'.

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

where $\pi(./A_i) \circ f^{-1} \in \Delta(X)$ is given by $(\pi(./A_i) \circ f^{-1})(x) := \pi\left(\{\omega | f(\omega) = x\}/A_i\right)$ 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 $F \in \mathcal{F}^{AA}$.

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, consider the following quasi-metric on Savage acts δ : [39]

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

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 equivalent 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 15 (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 Stochastic Dominance with respect to an almost-convex-ranged context $\trianglerighteq$, $\succsim^{AA}$ is a continuous[40] weak order extending $\succsim$.*

[39] For simplicity, this metric is defined in terms of the multi-prior representation Π^0 rather than the associated likelihood relation $\trianglerighteq^0$; the latter should be possible, although we have not worked out the details.

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

The present derivation of the Anscombe-Aumann framework is closely related to Machina's (2001) work on "Almost-Objective Uncertainty", which also reproduces the power of the Anscombe-Aumann framework in an enriched Savage setting; Proposition 4 in particular can be viewed as the counterpart of his Theorem 4. We note the following differences. 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 smoothness condition, while we model the postulated probabilistic belief explicitly as a likelihood relation, and obtain analogous preference restrictions via Continuity in Distribution and Stochastic Dominance. While our derivation is behaviorally general, Machina assumes that preferences are "event-differentiable" which is behaviorally somewhat restrictive; event-differentiability excludes, for example, the minimum expected-utility model due to Gilboa-Schmeidler (1989).[41] Finally, Machina exploits the specific structure of uniformly continuous densities on the real line (or, more generally, on Euclidean manifolds) as discussed in section 3.3; by contrast, our approach applies to arbitrary almost-convex-ranged belief contexts. Machina's more specific assumptions allow him to mimic quite precisely the product structure of the original AA-setup "in the limit"; while this aids the intended interpretation in terms of "almost-objective uncertainty", it seems inessential for the decision-theoretic purposes pursued here.

A.3 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

[41] Indeed, it follows from the analysis in Nehring (2001) that event-differentiability exculdes all utility-sophisticated preferences satisfying Savage's P4.

increasing sequence of dyadic numbers $\{\frac{\ell_k}{2^k}\}_{k=1,..,\infty}$ such that

$$\frac{\ell_k+1}{2^k} \geq \alpha. \tag{9}$$

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 (9), a decreasing sequence $\{D_k\}$ of members of $\mathcal{A}_k$ such that

$$B_{k'} \subseteq B_k \cup D_k \tag{10}$$

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 (10), $B \subseteq B_k \cup D_k$ 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.[42]

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

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 $A \in \frac{1}{2}E'$.*

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 $A \trianglerighteq B$.*

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 partitions 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 is Σ-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

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

use of completeness.

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 $\{D_{i1,\ldots,} D_{i2^{k_i}}\}$, 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 $C \backslash (\sum_i B_i) = \sum_i (C_i \backslash B_i) \in [X]$, as desired. $\square$

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)* $\mathbf{1} \widehat{\triangleright} \mathbf{0}$.
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 $\| X' - X \| \leq \frac{1}{K}$ and $\| Y' - Y \| \leq \frac{1}{K}$, it is not the case that $X' \widehat{\trianglerighteq} Y'$.

To verify this claim, take any X', Y' such that $\| X' - X \| \leq \frac{1}{K}$ and $\| Y' - Y \| \leq \frac{1}{K}$. 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. $\square$

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 definition, 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 $\mathcal{B}$. □

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 $A_i = \sum_{j \le \ell_i} E_{ij}$. 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 $\pi(A) = \sum_i \frac{\ell_i}{2^{k_i}} \pi(E_i) = E_\pi X$. $\square$

<u>Proof of Theorem 3.</u>

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 $\frac{1}{2} \le \pi^-(A/E)$ and $\pi^+(A/E) \le \frac{1}{2} + \frac{1}{n}$.

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 $j \in \{1, ..., 2n-1\}$, $\pi^+(A_j/E) \le \frac{1}{n}$. This suffices, since then by assumption $\pi^-(\sum_{i=1,..,n} A_i/E) \le \frac{1}{2}$ as well as $\pi^+(\sum_{i=1,..,n} A_i/E) \le \pi^+(\sum_{i=1,..,n-1} A_i/E) + \pi^+(A_n/E) \le \frac{1}{2} + \frac{1}{n}$.

To verify this claim, suppose that, by contradiction, for some $\pi \in \Pi$ 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, $\frac{1}{2n} < \pi^-(A_i/E) \le \pi^+(A_i/E) < \frac{1}{2(n-1)}$, 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 conclude 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 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 $\pi \in \Pi_2^K$ 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 2.

Consider any (non-null) event $A \in \Sigma_0$ and any $\gamma \in (0,1)$ and $\delta > 0$; we will show that there exists an event $B \in \Sigma_0$, $B \subseteq A$ such that $\frac{1}{(1+\delta)^3}\gamma\pi(A) \leq \pi(B) \leq (1+\delta)^3\gamma\pi(A)$ for all $\pi \in \Pi$.

Since $\Phi := \{\phi_\pi | \pi \in \Pi\}$ is equicontinuous and Ω is compact, Φ is uniformly equicontinuous. Since Φ is also uniformly bounded below above zero, there exists $\varepsilon > 0$ such that

$$\frac{\phi(a)}{\phi(b)} \leq 1+\delta \text{ for all } \phi \in \Phi \text{ and all } a, b \text{ such that } d(a,b) \leq 2\varepsilon. \tag{11}$$

By compactness, Ω can be covered by a finite number of open ε-balls; let $\mathcal{A} = \{D_i\} \subseteq \Sigma_0$ 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) on Σ. Since the Borel algebra Σ is generated by Σ_0 by the compactness (hence second-countability) of Ω, one can thus infer from Carathéodory's extension procedure (cf., for example, Aliprantis-Border (1999, p. 343)), that, for each i, there exists a set $B_i \in \Sigma_0$ with $B_i \subseteq A \cap D_i$ and γ_i such that

$$\begin{aligned} \lambda(B_i) &= \gamma_i \lambda(A \cap D_i), \text{ and} \\ \frac{\gamma}{1+\delta} &\leq \gamma_i \leq \gamma(1+\delta). \end{aligned} \tag{12}$$

We claim that $B = \sum_i B_i$ is the desired set.

By construction, the diameter of each D_i is no greater than 2ε. Hence, by (11), taking any $x_i \in D_i$, one has

$$\frac{\phi(x_i)\lambda(B_i)}{1+\delta} \leq \pi(B_i) = \int_{B_i} \phi(x)\, d\lambda \leq (1+\delta)\phi(x_i)\lambda(B_i), \tag{13}$$

and similarly

$$\frac{\phi(x_i)\lambda(A \cap D_i)}{1+\delta} \leq \pi(A \cap D_i) \leq (1+\delta)\phi(x_i)\lambda(A \cap D_i),$$

which implies

$$\frac{1}{(1+\delta)\phi\left(x_i\right)}\pi(A\cap D_i)\leq\lambda(A\cap D_i)\leq\frac{(1+\delta)}{\phi\left(x_i\right)}\pi(A\cap D_i). \tag{14}$$

Combining equations (12), (13) and (14) therefore

$$\frac{1}{(1+\delta)^3}\gamma\pi(A\cap D_i)\leq\pi(B_i)\leq(1+\delta)^3\gamma\pi(A\cap D_i).$$

Summing over i, one obtains

$$\frac{1}{(1+\delta)^3}\gamma\pi(A)\leq\pi(B)\leq(1+\delta)^3\gamma\pi(A),$$

as desired. □

Proof of Proposition 3.

As indicated in the text, to show that $\Pi_{\lambda,M}\supseteq\Pi$, it suffices to show that $\trianglerighteq_{\lambda,M}\subseteq\trianglerighteq_{\left(\Pi_{\lambda,M}\right)}$. To see this, take any A,B and such that $\lambda(A)\geq\Psi(A,B)\lambda(B)$ and any $\pi\in\Pi_{\lambda,M}$; we need to verify that $\pi\left(A\right)\geq\pi\left(B\right)$. Indeed,

$$\pi\left(A\right)=\int_A\phi\left(\omega\right)d\lambda\geq\lambda\left(A\right)\inf_{\omega\in A}\phi\left(\omega\right)\ \geq\lambda\left(B\right)\Psi(A,B)\frac{\sup_{\omega\in B}\phi\left(\omega\right)}{\Psi(A,B)}\geq\int_B\phi\left(\omega\right)d\lambda=\pi\left(B\right).$$

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

Step 1. π is absolutely continuous with respect to λ as a finitely additive measure; hence, in particular, π is in fact countably additive.

Fix any natural number L such that $L\geq e^{M\sup_{a,b\in\Omega}d(a,b)}$. For any natural number K, we will show that $\pi(A)\leq\frac{1}{K}$ for any A such that $\lambda(A)\leq\frac{1}{KL}$, which suffices. Take any K and A such that $\lambda(A)\leq\frac{1}{KL}$. By the convex-rangedness of λ, there exists a partition $\{A_i\}_{i\leq KL}$ of Ω such that $A_1\supseteq A$ and $\lambda\left(A_i\right)=\frac{1}{KL}$. 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 $L \geq \Psi(A_i, B_j)$ by construction, one has for all i, j that $B_j \trianglerighteq_{\lambda,M} A_i$, 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 open ε-ball around a.

Step 2. For all a, b 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. If the claim is false, then, in view of the full support of λ, there exist λ-non-null sets $A \subseteq B^{\varepsilon}(a)$, $B \subseteq B^{\varepsilon}(b)$ such that

$$\inf_{\omega\in A} \phi(\omega) > \sup_{\omega\in B} \phi(\omega) \Psi(B^{\varepsilon}(a), B^{\varepsilon}(b)). \tag{15}$$

By the convex-rangedness of λ, taking appropriate subsets if necessary, it can moreover be assumed that

$$\lambda(B) = \Psi(B^{\varepsilon}(a), B^{\varepsilon}(b)) \lambda(A). \tag{16}$$

By the definition of $\trianglerighteq_{\lambda,M}$ and the admissibility of π with respect to $\trianglerighteq_{\lambda,M,}$(16) implies $\pi(B) \geq \pi(A)$. On the other hand, from (15) and (16) one obtains

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

the desired contradiction.

Step 3. π has a density $\widetilde{\phi}$ such that $\log \widetilde{\phi}$ is Lipschitz with modulus of continuity M.

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

$$\phi_-^n(a) \quad : \quad = \operatorname{ess\,inf}_{\omega\in B^{\frac{1}{n}}(a)} \phi(\omega) \text{ for } a \in \Omega;$$
$$\phi_+^n(a) \quad : \quad = \operatorname{ess\,sup}_{\omega\in B^{\frac{1}{n}}(a)} \phi(\omega) \text{ for } a \in \Omega.$$

By step 2,

$$\phi_{+}^{n}(a) \geq \phi_{-}^{n}(a) \geq \phi_{+}^{n}(a)\, e^{-\frac{2}{n}M}. \tag{17}$$

Hence the increasing and decreasing sequences $\{\phi_{-}^{n}\}$ and $\{\phi_{+}^{n}\}$ converge pointwise to the same function $\widetilde{\phi}$. Now, for any $A \in \Sigma$,

$$\int_A \phi_{-}^{n}(\omega)\, d\lambda \leq \int_A \phi(\omega)\, d\lambda = \pi(A) \leq \int_A \phi_{+}^{n}(\omega)\, d\lambda. \tag{18}$$

Since by the Monotone Convergence Theorem $\int_A \phi_{-}^{n}(\omega)\, d\lambda$ and $\int_A \phi_{+}^{n}(\omega)\, d\lambda$ converge to $\int_A \widetilde{\phi}(\omega)\, d\lambda$, one obtains $\int_A \widetilde{\phi}(\omega)\, d\lambda = \pi(A)$ from (18), for any $A \in \Sigma$. Thus $\widetilde{\phi}$ is a density for π.

To verify that $\widetilde{\phi}$ has the asserted property, consider any $a, b \in \Omega$.

By step 2, for any n,

$$\widetilde{\phi}(a) \leq \phi_{+}^{n}(a) \leq \Psi(B^{\frac{1}{n}}(a), B^{\frac{1}{n}}(b))\phi_{-}^{n}(b) \leq \widetilde{\phi}(b).$$

Since by the triangle inequality,

$$\sup_{a' \in B^{\frac{1}{n}}(a), b' \in B^{\frac{1}{n}}(b)} d(a', b') \leq d(a, b) + \frac{2}{n},$$

this implies $\log \widetilde{\phi}(a) - \log \widetilde{\phi}(b) \leq M\left(d(a, b) + \frac{2}{n}\right)$. Likewise $\log \widetilde{\phi}(b) - \log \widetilde{\phi}(a) \leq M\left(d(a, b) + \frac{2}{n}\right)$ by interchanging a and b, and therefore

$$\left|\log \widetilde{\phi}(a) - \log \widetilde{\phi}(b)\right| \leq M\left(d(a, b) + \frac{2}{n}\right).$$

Since n is arbitrary, this establishes the claim. □

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 with 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 $\pi \in \Pi^0$, $d(\pi \circ f^{-1}, \pi \circ g^{-1}) \leq \sup_{i \in I} d\left(\pi(./F_i) \circ f^{-1}, \pi(./F_i) \circ g^{-1}\right) \leq \sup_{i \in I}\left(d\left(\pi(./F_i) \circ f^{-1}, q_i\right) + d\left(\pi(./F_i) \circ g^{-1}, q_i\right)\right) \leq \sup_{i \in I} d\left(\pi(./F_i) \circ f^{-1}, q_i\right) + \sup_{i \in I} d\left(\pi(./F_i) \circ g^{-1}, q_i\right) \leq \delta'(f, F) + \delta(f', F)$. $\square$

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 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 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 $F \succsim^{AA} G$. $\square$

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] Aliprantis, C. and K. Border (1999,): *Infinite Dimensional Analyis,* 2nd ed. Berlin: Springer.

[2] 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.

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

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

[5] Broome, J. (1991), *Weighing Goods,* Basil Blackwell, Oxford.

[6] Casadesus, R. , P. Klibanoff and E. Ozdenoren (2000): "Maxmin Expected Utility over Savage Acts with a Set of Priors", *Journal of Economic Theory* 92, 35-65.

[7] Chateauneuf, A. (1985): "On the Existence of a Probability Measure Compatible with a Total Preorder on a Boolean Algebra," *Journal of Mathematical Economics* 14, 43-52.

[8] Chateauneuf, A. and J.Y. Jaffray (1984): "Archimedean Qualitative Probabilities," *Journal of Mathematical Psychology* 28, 191-204.

[9] Chew, S.-H. and J. Sagi (2003): "Domains", *mimeo.*

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

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

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

[13] Finetti, B. de (1931): "Sul significato doggetivo della probabilità," *Fundamenta Mathematica* 17, 298-329.

[14] Finetti, B. de (1937): "La prevision, ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincare* 7, 1-68.

[15] Fishburn, P. (1970), *Utility Theory for Decision Making,* New York: John Wiley and Sons.

[16] Fishburn, P. (1973): "A Mixture-Set Axiomatization of Conditional Subjective Expected Utility," *Econometrica* 41: 1-25.

[17] Fishburn, P. (1986): "The Axioms of Subjective Probability," *Statistical Science* 1, 335-358.

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

[19] Gilboa, I., D. Samet and D. Schmeidler (2001): "Utilitarian Aggregation of Beliefs and Tastes," *mimeo.*

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

[21] Ghirardato, P., F. Maccheroni and M. Marinacci (2002): "Ambiguity from the Differential Viewpoint," *mimeo.*

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

[23] Grant, S. (1995): "Subjective Probability Without Monotonicity: Or how Machina's Mom may also be Probabilistically Sophisticated", *Econometrica* 63, 159-189.

[24] Heath, D. and W. Sudderth (1972): "On a Theorem of De Finetti, Oddsmaking, and Game Theory," *Annals of Mathematical Statistics* 43, 2072-2077.

[25] Jaffray, J.-Y. (1989): "Linear Utility Theory for Belief Functions," *Operations Research Letters*, 9, pp. 107-112.

[26] Karni, E. , D. Schmeidler and K. Vind (1983): "On State Dependent Preference and Subjective Probabilities", *Econometrica* 51, 1021-1032.

[27] Karni, E. (1996): "Probabilities and Beliefs", *Journal of Risk and Uncertainty* 13, 249-262.

[28] Karni, E. and S. Grant (2000), "A Theory of Quantifiable Beliefs", forthcoming in *Journal of Mathematical Economics*

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

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

[31] Kopylov, I. (2003): "Subjective Probability on 'Small Domains"', chapter of Ph.D. thesis, Rochester.

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

[33] Levi, I. (1980): *The Enterprise of Knowledge.* Cambridge, MA: MIT Press.

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

[35] Machina, M. (2001): "Almost Objective Uncertainty", working paper UC San Diego.

[36] Machina, M. (2002): "Robustifying the Classical Model of Risk-Preferences and Beliefs", working paper UC San Diego.

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

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

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

[40] Nehring, K. (2001): "Ambiguity in the Context of Probabilist Beliefs," working paper UC Davis,
available at http://www.econ.ucdavis.edu/faculty/nehring/papers.html

[41] Poincaré, H. (1912), *Calcul des Probabilités.* 2nd ed. Paris: Gauthiers-Villars.

[42] Pratt, J., H. Raiffa and R. Schlaifer (1964): "The Foundations of Decisions under Uncertainty: An Elementary Exposition," *Journal of the American Statistical Association* 59, 353-375.

[43] 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.

[44] Regoli, G. (1998): "Comparative Probability Orderings," available at homepage of the Imprecise Probability Project (http://ippserv.rug.ac.be)

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

[46] Sarin, R. and P. Wakker (2000): "Cumulative Dominance and Probabilistic Sophistication," *Mathematical Social Sciences* 40, 191-196.

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

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

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

[50] Villegas, C. (1964): "On Qualitative Probability Sigma-Algebras," *Annals of Mathematical Statistics* 35, 1787-1796.

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

[52] Zhang, J. (1999): "Qualitative Probabilities on λ-systems," *Mathematical Social Sciences* 38, 11-20.