Imprecise Probabilistic Beliefs as a Context for Decision-Making under Ambiguity

Klaus Nehring

University of California, Davis[1]

First Version: November 2002;

This Version: March 27, 2008[2]

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

[2] I thank audiences at RUD in Paris (June 2002), the University of Heidelberg, Stanford, and the Institute for Advanced Study, Princeton, for helpful feedback. Precursors of some of the ideas of this paper were presented in December 1996 at LOFT2 in Torino, Italy, and in July 2001 at RUD in Venice, Italy.

Abstract

Coherent imprecise probabilistic beliefs are modelled as incomplete comparative likelihood relations admitting a multiple-prior representation. Under a structural assumption of Equidivisibility, we provide an axiomatization of such relations and show uniqueness of the representation. In the second part of the paper, we formulate a behaviorally general "Likelihood Compatibility" axiom relating preferences and probabilistic beliefs and characterize its implications for the class of "invariant biseparable" preferences that includes the MEU and CEU models among others.

1. INTRODUCTION

In the wake of Ellsberg's (1961) celebrated experiments, it is by now widely recognized that decision makers are not always guided by a well-defined subjective probability measure. Ellsberg's challenge to received decision theory is particularly profound in that it puts into question not so much particular assumptions on decision makers' preference attitudes towards uncertainty, but the very understanding of uncertainty itself. Even though much effort has gone into modelling of Ellsberg-style "ambiguity", the nature and role of probabilistic beliefs in such contexts is not yet well understood. This issue is central not just from the point of view of decision theory itself, but also from that of its economic applications, since, in large part, economic models are models of agents' beliefs, whether in macroeconomics, finance, game theory or elsewhere.

The modelling of an agents' probabilistic beliefs under ambiguity can be approached in at least two ways. On the one hand, one might try to *define* beliefs from preferences following Savage (1954). While Savage's own definition can be invoked at a purely formal level even under ambiguity, it is in general no longer associated with well-defined probabilistic beliefs, as will be illustrated shortly in the context of the Ellsberg paradox. The canonical relation between probabilistic beliefs and (betting) preferences that obtains under expected utility breaks down, since betting preferences are now determined by beliefs –however construed– and ambiguity attitudes.[1] It is an open question whether and under what circumstances Savage's definition can be generalized satisfactorily. And, in any case, it seems likely that even the "best possible" definition will be less canonical, that it will come with more strings attached than Savage's. In this paper, we therefore want to pursue a less ambitious goal:

> "Suppose that we know that the decision-maker entertains a specified set of probabilistic beliefs, and possibly others. What is the structure of such beliefs, and how do they constrain his preferences?"

In addressing these questions, we shall strive for behavioral generality: any satisfactory answer must be applicable to a wide range of choice behavior including for example Allais- and Ellsberg-

[1] 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 (2004)); this criticism assumes, however, a non-behaviorist point of view to begin with.

style choice patterns, and should not be tied to assumptions about specific functional forms. As argued compellingly by Machina-Schmeidler (1992) and Epstein-Zhang (2001), behavioral generality is important since issues about the representation of probabilistic beliefs are more fundamental than particular behavioral assumptions.

Neither of the two questions has been answered satisfactorily in the literature. To the extent that a "default answer" is available regarding the structure of imprecise probabilistic beliefs, it is presumably given by a representation in terms of sets of priors. However, the existing justifications all assume expected-utility maximization with respect to risk (as in Bewley (1986), Walley (1991)), or Gilboa-Schmeidler (1989)), and, in the latter case, also very specific attitudes towards ambiguity.

As far as we can tell, the second issue of how imprecise probabilistic beliefs (rationally) constrain preferences has not been explicitly addressed in the literature. And, indeed, even in simple cases, the answer is not completely obvious, as we illustrate in section 4 in the context of the α-Minimum Expected Utility and Choquet expected utility models.

To address the two questions in behavioral generality, we propose to model probabilistic beliefs as a comparative likelihood relation $\trianglerighteq$ over events, with "$A \trianglerighteq B$" denoting the judgement "A is at least as likely as B". In this we follow the lead of the classical contributions by Keynes (1921), de Finetti (1931) and Savage (1954). The likelihood relation shall be taken as an independent, non-behavioral datum, *leaving open* the question whether/under what circumstances it can in turn be derived from preferences. The likelihood relation can represent either "objective" probabilistic information or purely subjective beliefs; these interpretations are fleshed out at the beginning of section 2.

Imprecise Probabilistic Beliefs in the Ellsberg Paradox

While preferences will be assumed to be complete as usual, the likelihood relation will assumed to be incomplete in order to make room for ambiguity. To illustrate the role of incompleteness, let us consider the classical two-color version of the Ellsberg paradox. One ball is drawn from each of two urns both of which are composed of red and black balls only. The decision maker is told that the first ("known") urn contains as many red as black balls, but is told nothing about the composition of the second ("unknown") urn. We will focus here on the four events associated with the colors of each draw: R_{kn} and B_{kn} (the ball drawn from the known urn is red / black), as well as R_{un} and B_{un} (the ball drawn from the unknown urn is red / black). There is one fundamental piece of probabilistic information, namely that the events R_{kn} and B_{kn} are equally likely ($R_{kn} \equiv B_{kn}$). According to the typically observed choice pattern, betting on any color of the known urn is preferred to betting on any color of the unknown urn[2]:

$$R_{kn} \sim B_{kn} \succ R_{un} \sim B_{un}. \tag{1}$$

Comparative likelihood relations constrain betting preferences canonically: if A is at least as likely as B, then betting on A must be weakly preferred to betting on B. If this condition is satisfied for arbitrary events A and B, preferences and the specified information/beliefs will be said to be *compatible* with each other. We shall refer to the underlying rationality principle that extends to multi-valued acts as "Likelihood Compatibility".

In the above example, preferences are evidently compatible with the specified information that $R_{kn} \equiv B_{kn}$. One may wonder, however, whether it is possible to attribute to the decision maker *in addition* a belief that red and black from the unknown urn are equally likely, $R_{un} \equiv B_{un}$, as would be implied by Savage's definition of revealed likelihood. Yet this can be done only at the price of sacrificing the fundamental coherence properties that characterize the "logic of probability". For this logic evidently implies that if a red and black draw from the unknown urn were judged equally likely, then all four possible draws must be equally likely. But such a judgment would be incompatible with the observed preference for betting on the known urn exhibit by (1). A similar argument shows that the specified preferences are not compatible with attributing a belief that R_{un} is strictly more, or strictly less, likely than B_{un}. *Thus any coherent likelihood relation that is compatible with the specified preferences must be incomplete even though the preference relation itself is complete.*

Incompleteness of the likelihood relation alongside a complete preference relation yields an intuitive account of the Ellsberg paradox, in that the absence of a likelihood comparison between the colors from the unknown urn captures precisely the epistemic difference between the two urns that motivates the preference for betting on the known urn. Indeed, this is not a novel interpretation at all, but simply fleshes out formally the common verbal interpretation starting with Ellsberg (1961) and Schmeidler (1989).

This *beliefs-based* explanation of the Ellsberg paradox is not the only possible explanation . A frequently proposed alternative is derived from the claim that the decision maker has well-defined global subjective probabilities, but simply "dislikes" betting on the unknown urn relative to betting on the known urn.[3] This alternative, preference-based account allows to maintain completeness of the likelihood relation at the price of sacrificing Likelihood Compatibility. This is a high price to

[2]In this notation, an event E is preferred to another event E' if betting on E (receiving the better of two consequences on E, and the worse on E^c) is preferred to betting on E'.

[3]Segal (1990), Ergin-Gul (2004) and Chew-Sagi (2003) can be interpreted in this vein, as well as perhaps Tversky-

pay as it severs radically the connection between belief and preference, whereas in the belief-based account at least a unidirectional version of the classical relationship is preserved.

Representation of Coherent Likelihood Relations by Multiple Priors

The example also illustrates that the content and power of the restrictions induced by a set of likelihood judgements depends critically on the nature of entailment relationships among them. The first key task of the present paper is therefore the characterization of "coherent" likelihood relations, that is, of likelihood relations that incorporate all entailments from the logic of probability. For the limiting case of complete relations, Savage (1954) achieved a characterization of this kind leading to a representation by a numerical probability measure. This result was in fact a key step in deriving his celebrated Subjective Expected Utility Theorem. Remarkably, by an appropriate choice of auxiliary conditions, Savage was able to make do with a single 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 atomless (more precisely: "convex-ranged") probability measures.

The first main result of the present paper, Theorem 2, offers a counterpart to Savage's result for incomplete comparative likelihood relations; it appears to be the first result of its kind in the literature. Without completeness, 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 each split into a more and a less likely "subevent", and if A is judged at least as likely as B, then the more likely subevent of A must be at least as likely as the less likely subevent of B. Under appropriate auxiliary conditions, Theorem 2 shows that a likelihood relation satisfies Additivity and Splitting if and only if it has a representation in terms of a set of admissible probability measures ("priors"); according to this representation, an event A as at least as likely as B if and only if A's probability is at least as large as that of B, for any admissible prior in the set. Theorem 2 justifies a formal identification of coherence with the existence of such a multi-prior representation.

As in Savage, and indeed in a somewhat more pronounced form, there is a price to be paid for the simplicity in the rationality axioms underlying coherence due to the need for substantive

Wakker's (1995) notion of "source preference".

structural assumptions. Specifically, we assume that any event can indeed be split into two equally likely subevents (roughly as in De Finetti 1931). Besides non-atomicity, Equidivisibility 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 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 2 for details and further examples.

Importantly, Equidivisibility ensures *uniqueness* of the multi-prior representation (within the class of closed, convex sets of priors). We show by example (see section 2.4) that this assumption cannot be greatly weakened without losing uniqueness. 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 such likelihood relations as the canonical primitive representing probabilistic beliefs.

Preferences Constrained by Imprecise Probabilistic Beliefs

In the second part of the paper, we consider how specified imprecise probabilistic beliefs rationally constrain preferences. To do so, we propose an axiom called "Likelihood Compatibility" that extends the compatibility requirement formulated above for betting preferences to acts with multiple outcomes. It represents a *minimal,* generally applicable criterion of consequentialist rationality relating preferences to probabilistic beliefs expressed as likelihood relations. It is minimal in that it does not constrain the DM's risk or ambiguity attitudes in any substantive way, thereby ensuring behavioral generality.

We view the existence of behaviorally general yet substantive rationality restrictions on preferences captured by this axiom as a crucial advantage of using likelihood relations as the epistemic primitive in contrast to, for example, a direct use of sets of priors. In particular, we show that, given a likelihood relation satisfying the assumptions of Theorem 2, Likelihood Compatibility entails probabilistic sophistication in the sense of Machina-Schmeidler (1992) over risky (unambiguous) acts, that is: acts whose outcomes have well-defined probabilities derived from the likelihood relation. [4]

[4]Taking this argument further, in the working paper version of this paper we show that any such preference ordering can be represented as a preference ordering over Anscombe-Aumann (1963) acts with a mixture-operation that is defined in terms of the given likelihood relation. This construction can be viewed as a decision-theoretic, belief-based foundation for the Anscombe-Aumann (1963) framework. This derivation not only clarifies the assumptions on

In applications, it is obvious importance to determine when preferences belonging to a particular preference model are compatible with a specified likelihood relation. This need not be straightforward even in very simple cases such as the MEU preferences. In the MEU model, for example, one would like to be able to determine compatibility by comparing the set of priors representing the preference relation Ψ to the set of priors representing the likelihood information Π, checking for set inclusion. While this criterion need not always work, we show that it does work whenever preferences maximize expected utility over risky (unambiguous) acts. This result is obtained as a corollary to the main result of the second part of the paper, Theorem 1, which shows for a large class of preferences under which conditions Likelihood Compatibility can be characterized in terms of a generalized multi-prior criterion that is based on Ghirardato et al.'s (2004) and Nehring's (1996) definition of perceived ambiguity.[5]

Related Literature

1. Our main result, Theorem 2, builds on and can be viewed as the likelihood counterpart of the multiple-prior representations of partial orderings due to Bewley (1986/2002) and Walley (1991) following Smith (1961). All of these, however, use preferences as their primitive[6] and derive the multiple-prior representation together with expected-utility maximization with respect to those priors, and thus fail to be behaviorally general. Mathematically, the objects of the present paper (orderings over sets) have much less structure a priori than the objects in these contributions (orderings over random variables). This difference probably explains why, in spite of the suggestive parallelism, there do not exist counterpart results for likelihood relations in the literature up to now. The key technical insight of the present paper is the realization that it is possible to formulate simple, epistemically well-motivated axioms that allow to canonically extend a likelihood ordering over events to an ordering over real-valued functions, thereby making the existing characterizations and the associated vector-space techniques such as separation theorems applicable; the construction of the extension itself is non-trivial.[7]

preferences and beliefs implicit in the Anscombe-Aumann model, it leads to an even more powerful structure since all uncertainty is treated at the same level.

[5]For a subclass of these preferences, this definition is closely related to Siniscalchi's (2006) notion of "plausible priors".

[6]Walley and Smith do so by taking "acceptable gambles" as their primitive notion.

[7]Multiple-prior representations of complete preference orderings have been obtained by Gilboa-Schmeidler (1989), Ghirardato et al. (2004) and Casadesus et al. (2000); again, these are about preferences, not belief, and are behav-

2. In terms of overall motivation of axiomatizing an epistemic primitive, a closely related contribution in the literature is Koopman (1940a and b). Koopman presents an axiomatic treatment of comparative *conditional* likelihood relations, whose primitive compares event pairs ("A given B is at least as likely as C given D"). Koopman's results are much weaker, however, than the results of the present paper: while Koopman provides sufficient conditions for the existence of lower- and upper-probability functions that are additive on the class of events where the two coincide, he has no representation theorem and no characterization of coherence. It is also not clear how conditional likelihood comparisons are to be related to behavior.

3. There is a sizeable literature on comparative likelihood relations that is mainly focused on the complete case; see Fishburn (1986) and Regoli (1998) 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 combinatorial complexity and algebraic character of the conditions, such characterizations have generally not been considered to be of significant foundational interest (c.f. e.g. Regoli 1998).

Furthermore, the important uniqueness issue has not been addressed before outside the complete case. Indeed, it seems fairly remarkable a priori that likelihood relations can match multi-prior representations in their expressiveness at all; we are not aware of any hint of this in the literature; see, for example, the discussion in Walley (1991, pp. 191-197) which appears to suggest the opposite. In sum, in spite of the existence of the multi-prior representation results dating back to Smith (1961), the extant results in the literature on likelihood relations do not come close to those of the present paper.

4. Somewhat related to the second part of the paper is a literature in which imprecise probabilistic beliefs (co-)*determine* rather than merely *constrain* preferences, as assumed here and alluded to by the word *con*text; see Jaffray (1989), Nehring (1991, 1992, 2000), Gajdos et al. (2004, 2006), Wang (2003) and others[8]. As explained in more detail in section 3.2, these models rely on a much stronger, exhaustive interpretation of the given probabilistic beliefs. By contrast, the present approach as-

iorally quite restrictive.

[8]Other, more distantly related contributions include in particular those that model preferences over sets of lotteries such as Olsziewski (2002), Stinchcombe (2003) and Ahn (2005).

sumes only an incomplete interpretation of the likelihood relation as describing only *some* but not necessarily all of the DM's probabilistic beliefs, for example beliefs derived from probabilistic information such as the existence of a random device, the composition of an Ellsberg urn, information about frequencies in large populations; see Walley (1991, sections 2.10.3 and 9.7.4) for more on the fundamental distinction between exhaustive and incomplete interpretations of imprecision. Typically, contributions relying on an exhaustive interpretation relate preferences/choices in different epistemic situations or even allow for preferences over such situations.[9] Such frameworks allow substantially stronger conclusions on the basis of substantially stronger and presumably conceptually more controversial assumptions. Their goal is typically to characterize specific, epistemically interpretable models of decision making under ambiguity, a goal that is quite different from the present goal of determining the constraints imposed by the existence of probabilistic beliefs/information in behaviorally general manner.

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 $\triangleright$.

2.1 The Likelihood Relation as Information and as (Partial) Belief

The inclusion of beliefs among the primitives is a likely source of controversy, as it goes against the grain of the reigning Ramsey-De Finetti-Savage tradition. Precisely because we do not want to belittle the methodological and philosophical issues at stake, we defer their discussion to future work. In its place, we submit that 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

[9] An exception is Kopylov (2006).

the decision maker's probabilistic beliefs among the primitives.

The likelihood relation can be given two primary interpretations. First, the likelihood relation may summarize *information* about the unconditional, conditional or comparative probabilities available to the decision maker. Such information arises naturally in various contexts. For example, as we shall explain in section 2.5, the notion of an independent random device with known objective probabilities that is at the heart of the Anscombe-Aumann (1963) framework can be usefully modeled in this way. Similarly, information about the composition of urns in the context of Ellsberg experiments represents important probabilistic information. Likewise, if the decision maker observes independent, identical repetitions of a sampling experiment with unknown parameters (e.g. tosses of the same coin with unknown bias), this information about the structure of the sampling process can be captured by a comparative likelihood relation that embodies "exchangeability" a la de Finetti (1937))[10]. On the information interpretation, a likelihood relation will be almost always incomplete, since the decision maker will possess information only about the likelihood of some events but not of others.

Secondly and more generally, the likelihood relation can serve to represent the decision maker's *subjective beliefs*, whether or not these are based on "given" information. Here, beliefs as an independent (non-behavioral) datum are to be understood as "propositional attitudes", that is: as dispositions to affirm certain likelihood-judgments in thought or in speech, in addition to preferences which can be viewed as dispositions to act. Importantly, the beliefs need not be specified exhaustively. That is, the decision maker may "have" further beliefs that have not yet been elicited and recorded in $\trianglerighteq$, but which may be verified either by further elicitation (e.g. via interrogation) or revelation through preferences. Indeed, probabilistic information in the sense above can be understood as a special case of non-exhaustively specified probabilistic beliefs.

2.2 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) $\Omega \triangleright \emptyset$.

[10]Roughly, this likelihood relation takes any event to be equally likely to any permutation of it.

Axiom 3 (Positivity) $A \trianglerighteq \emptyset$ *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* $A \cap C = B \cap C = \emptyset$.

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 do not matter. It is well-known that, on finite state-spaces, Additivity is far from sufficient to guarantee the existence of a probability-measure representing the complete comparative likelihood relation; see Kraft-Pratt-Seidenberg (1959). Refining an earlier seminal result by de Finetti (1931), Savage (1954) provided assumptions which, together with Additivity, gave rise to a characterization of convex-ranged probability measures;[11] 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 $\pi(B) = \alpha\pi(A)$. Evidently, convex-ranged probability measures exist only when the state-space is infinite. We state a version of his result for the sake of comparison. It requires two more axioms; the event A is *non-null* if $A \vartriangleright \emptyset$.

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

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

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

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

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

2.3 Dropping Completeness

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

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

[11]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.

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 is no longer feasible. For example, while under completeness, one can use Additivity to infer that if two events are equally likely to their respective complements, they must be equally likely to each other, this no longer follows without completeness. Yet such an implication seems necessary for a proper *likelihood* interpretation of the relation. More generally, the following second rationality axiom called "Splitting" seems intuitively compelling.

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 proof of the following Theorem, we will only make use of the special case in which the two events are split into equally likely subevents.

Significantly, Splitting is not a conceptually independent addition to Additivity, but merely compensates for the missing completeness of the likelihood relation, in that any additive completion of a given likelihood relation satisfies Splitting automatically.

Fact 1 *For any weak order $\trianglerighteq$, Additivity implies Splitting.*

Fact 1 shows that Splitting appeals to the same ordinal, qualitative intuition that makes the Additivity axiom so compelling. By contrast, the linear-programming inspired conditions of Heath-Suddert (1972) and Walley (1991 p. 192-3) already appeal to a cardinal notion of subjective probability, as a result of which their foundational value seems to be rather limited.

By themselves, Additivity plus Splitting are not enough to deliver an interesting representation, as the case of complete likelihood relations on a finite state-space shows. We thus make 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 $B \equiv A \backslash B$.*

Very broadly, Equidivisibility can be viewed as an assumption that the likelihood relation is sufficiently rich in comparisons. The axiom can be motivated, for example, by the existence of a rich set of independent random events. To see this, let T be an event with an unambiguous probability of 0.5, i.e. such that $T \equiv T^c$. Then A is naturally viewed as independent from T if this judgment is maintained conditional on the occurrence of A, that is if $A \cap T \equiv A \cap T^c$. Clearly $A \cap T$ and $A \cap T^c$

split A into two equally likely parts. Note that the plausibility of the existence of such events does not depend on whether or not the event A itself is unambiguous.

Finally, Savage's Fineness and Tightness axioms are no longer adequate to obtain a real-valued representation. In their stead, 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 mutually disjoint events A_i such that $A \trianglelefteq A_i$ for all i. A sequence of events $\{A_n\}_{n=1,..,\infty}$ is *converging in probability* to the event A if, for all $K \in \mathbf{N}$ there exists $n_K \in \mathbf{N}$ such that for all $n \geq n_K$ the symmetric difference $A_n \triangle A$ is a $\frac{1}{K}-$event.

Axiom 10 (Continuity) *For any sequences $\{A_n\}_{n=1,..,\infty}$ and $\{B_n\}_{n=1,..,\infty}$ converging in probability to A and B respectively,*

$$A_n \trianglerighteq B_n \text{ for all } n \text{ implies } A \trianglerighteq B.$$

These axioms ensure the existence of a **multi-prior representation**, i.e. the existence of a set of finitely additive probability measures $\Pi \subseteq \Delta(\Sigma)$ such that, for all $A, B \in \Sigma$:

$$A \trianglerighteq B \text{ if and only if } \pi(A) \geq \pi(B) \text{ for all } \pi \in \Pi. \tag{2}$$

Likelihood relations for which such a representation exists will be called **coherent**.

Note that if $\trianglerighteq$ satisfies (2) for some set of priors Π, then it satisfies (2) also for the closed convex hull 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 all closed (hence compact), convex subsets of $\Delta(\Sigma)$ be denoted by $K(\Delta(\Sigma))$.

Note also that all axioms except Equidivisibility but including Continuity[12] are implied by the existence of a multi-prior representation. Equidivisibility imposes further constraints on the set of priors Π. On σ-algebras, it is equivalent to the following "range-convexity" condition on Π; if Σ is merely an algebra, it is equivalent to "dyadic range-convexity".[13] Let $\mathbf{D}$ denote the set of dyadic

[12]The necessity of Continuity follows from observing that, for any $\trianglerighteq$ with representing set Π, any $\pi \in \Pi$ and any $\frac{1}{K}-$event A, $\pi(A) \leq \frac{1}{K}$; if Π is convex-ranged as defined just below, the converse holds as well.

In contrast to Continuity, neither Tightness nor Fineness are entailed by coherence, even under completeness. While Tightness is implied by coherence and Equidivisibility, Fineness is not; indeed, it is not difficult to verify that a coherent and equidivisible relation is fine if and only if, for all $A \in \Sigma : \min_{\pi\in\Pi} \pi(A) = 0$ implies $\max_{\pi\in\Pi} \pi(A) = 0$, which in turn is equivalent to the condition that all admissible priors $\pi \in \Pi$ have the same null-events. While vacuously satisfied in the precise case of a singleton set Π, this condition is clearly quite restrictive when beliefs are imprecise.

[13]The generality added by allowing Σ to be an algebra is significant since algebras can often be described explicitly while σ-algebras typically cannot. We note that Savage's Theorem has only very recently been extended to algebras by Kopylov (2003).

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-convexity of Π implies the range-convexity of every $\pi \in \Pi$, the converse is not true in general unless Π has a finite number of extreme points (see example 2 below). Moreover, as established by Fact 5 in the Appendix, on σ-algebras dyadic range-convexity and range-convexity coincide.

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, Positivity, Nondegeneracy, Additivity, Splitting, Equidivisibility and Continuity.*

If the set of priors $\Pi \subseteq \Delta(\Sigma)$ *represents* $\trianglerighteq$ *(i.e. satisfies (2)), then* $\Pi' \subseteq \Delta(\Sigma)$ *represents* $\trianglerighteq$ *as well if and only if* Π *and* Π' *have the same closed convex hull.*

By the final statement of the characterization, the likelihood ordering $\trianglerighteq$ has always a *unique* closed convex set of priors representing it (in the product or "weak*"-topology which will be assumed throughout). This will henceforth be taken to be the canonical representation of $\trianglerighteq$ and denoted by $\Pi_{\trianglerighteq}$ or Π for short.[14]

We shall sketch the proof idea of Theorem 2 with a bit of "reverse engineering". The key is the derivation of a vector-space-like structure of the event-space resulting from the range-convexity 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 $\mathcal{Z}$ of finite-ranged, Σ-measurable 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

[14] In view of the Krein-Milman theorem, an alternative candidate for a canonical representation would be the set of extreme points of this closed convex set Π.

partial ordering on $\mathcal{Z}$, denoted by $\widehat{\trianglerighteq}$, by setting

$$Y \widehat{\trianglerighteq} Z \text{ if and only if } A \trianglerighteq B \text{ for some } A \in [Y] \text{ and } B \in [Z]. \tag{3}$$

It is easily verified that this ordering satisfies the following two conditions:

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

and

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

Moreover, it is positive, non-degenerate and continuous.[15] In the sequel, we shall refer to partial orderings on $\mathcal{Z}$ satisfying these five 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 a final step, we invoke the just-quoted Theorem to obtain the desired multi-prior representation. The proof is non-trivial and requires a surprising amount of work due to the gap between the ordinally formulated axioms and the cardinal character of the derived conditions.

2.4 Fully Expressive Likelihood Relations

Analogously to the complete case in which uniqueness of the representing prior is a natural heuristic yardstick of the expressive adequacy of likelihood relations as an epistemic primitive, in the more general incomplete case uniqueness of the representing closed and convex set of priors is again a natural criterion of adequate expressiveness. Thus we shall refer to coherent likelihood relations with a unique multi-prior representation as *fully expressive.*

The uniqueness statement in Theorem 2 is thus significant by ensuring that there is a large and interesting class of imprecise likelihood relations that is fully expressive and thus viable as an epistemic primitive. The existence of such a class is not at all obvious *a priori.* Indeed, in finite state spaces, likelihood relations seem to lack full expressivity 'generically', and there is no indication in

[15]See Appendix, Lemma 9, for formal definitions.

the existing literature that this situation can be remedied in a systematic fashion in infinite state spaces.[16]

While Equidivisibility is not strictly necessary for full expressivity, it does not seem possible to weaken this assumption greatly and still obtain full expressivity in a robust manner. In particular, non-atomicity-like conditions akin to Savage's Fineness condition are not nearly enough.

Example 1. 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:

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

If $K > 1$, $\trianglerighteq^K$ is not equidivisible; in particular, $\trianglerighteq^K$ does not admit any event with unambiguous probability $\frac{1}{2}$.[17] In Appendix A.1, we exhibit two distinct multiple-prior representations for each $\trianglerighteq^K$, falsifying uniqueness and thus full expressivity.

Note that, for $K > 1$, $\trianglerighteq^K$ satisfies Savage's Fineness and Tightness conditions. Moreover, if K is close to 1, all admissible priors are uniformly close to the Lebesgue measure which is convex-ranged. Thus, even though such $\trianglerighteq^K$ are close to being convex-ranged and close to being complete, uniqueness is lost.

Example 1 suggests that Equidivisibility is not far from necessary for full expressivity. This is further confirmed by asking whether a likelihood relation can capture a particular kind of probability judgment, such as the judgment that "my subjective probability of A conditional on E is $\frac{3}{4}$". Suppose that the DM has already specified a (non-exhaustive) equidivisible likelihood ordering $\trianglerighteq$, and wants to express this judgment by adding some appropriate likelihood comparisons capturing this judgment. If $\trianglerighteq$ is equidivisible, he can achieve this by picking any event B with unambiguous conditional probability 0.75 given E, i.e. any B such that $\pi(B) = 0.75\pi(E)$ for all $\pi \in \Pi$, and adding the judgment "$A \equiv B$". But if $\trianglerighteq$ is not equidivisible, the DM will not in general be able to capture the judgment in terms of a comparative probability comparison, a significant failure in expressivity. In particular, note that it is not sufficient to assume that $\trianglerighteq$ possesses a rich set of

[16] See, for example, the discussion of comparative probability orderings in Walley (1991, section 4.5). Levi (1980, p. 207) goes as far as asserting that "the serious trouble is that there does not seem to be any way in which one can obtain complete information about X's credal state from data about comparative probability when that state is one which violates strict Bayesian requirements [i.e. precision]".

[17] To see this, suppose that there is an event E such that $E \equiv^K E^c$; by (5), this would imply that both $\lambda(E) \geq K\lambda(E^c)$ and $\lambda(E^c) \geq K\lambda(E)$, contradicting the assumption that $K > 1$.

unambiguous events; what matters is a sufficiently rich set of *conditionally* unambiguous events, i.e. range-convexity.

2.5 Examples of Equidivisibility

The key structural assumption behind Theorem 2, Equidivisibility, in effect assumes a sufficiently rich set of conditionally unambiguous events. Its content is further illuminated by means of the following two specific examples.

Example 2 (Limited Imprecision, Social Belief Aggregation). The first example is based on the intuitive notion of a limited extent of overall ambiguity. One way to make this intuition 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 $\Pi_{\trianglerighteq_I} = co\{\mu_i\}_{i \in I}$; the "social prior" μ_I must therefore be a convex combination of individual priors.[18]

Example 3 (External Randomization Device)

In the manner of Anscombe-Aumann (1963), consider state spaces with a continuous extraneous randomization device. 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 $\trianglerighteq_{\text{rand}}$ 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.

[18] This corollary to Theorem 2 is related to results by Mongin (1995) and Gilboa-Samet-Schmeidler (2004), who derive from social respect for unanimous indifferences a representation of the social prior as an affine linear combination of individual priors.

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

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

While the first condition ensures the existence of a convex-ranged probability measure η over random events, the second describes their stochastic independence. In view of AA1 and AA2, it is easily verified that $\trianglerighteq_{\text{rand}}$ satisfies all the assumptions of Theorem 2 including Equidivisibility. Hence there exists a unique set of priors Π^{rand} representing $\trianglerighteq_{\text{rand}}$; indeed, Π^{rand} is simply the set of all product-measures of the form $\pi_1 \times \eta$, where π_1 can be any finitely additive measure on Σ_1, reflecting the stochastic independence of the random device.

The example shows that, in an important sense, Equidivisibility is completely unrestrictive in that any coherent likelihood ordering can be extended to an equidivisible one after embedding in a larger state space. Moreover, such an embedding is imaginatively and operationally achievable. Thus, at bottom, range-convexity merely requires that the domain of beliefs (and, later, preferences) be sufficiently comprehensive.[19] Assuming the domains of preference/belief to be sufficiently rich is a standard strategy of theoretical modelling, and there seems to exist no particular reason in the present case that would appear to make this strategy unworkable or irrelevant.

3. DECISION MAKING IN THE CONTEXT OF PROBABILISTIC BELIEFS

3.1 Likelihood Compatibility: Definition

Consider now a decision maker described by a preference ordering over acts and a likelihood ordering 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 $[x_1 \text{ on } A_1; x_2 \text{ on } A_2; ...]$ for the act with consequence x_i in event A_i; for the act $[x \text{ on } A; y \text{ on } A^c]$ 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 like-

[19] Note, in particular, that we have not appealed to *objective* probabilities, which could be viewed as a heterogeneous and philosophically controversial element.

lihood relation $\trianglerighteq$ on Σ. The relation $\trianglerighteq$ 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 $(\succsim, \trianglerighteq)$. A coherent context $\trianglerighteq$ will be referred to as *convex-ranged* if it has a convex-ranged multi-prior representation Π on the event-algebra Σ.[20]

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 the *only* attribute of events that should matter in comparing the incidence sets $f^{-1}(x)$ 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 Likelihood Consequentialism.

The task is to formalize this principle in terms of axioms on the relation between preferences and beliefs in maximal behavioral generality, that is in particular: without imposing restrictions on risk-preferences. By way of motivation, begin by considering preferences over bets, i.e. comparisons of pairs of the form $([x \text{ on } A; y \text{ on } A^c], [x \text{ on } B; y \text{ on } B^c])$. Here, Likelihood Consequentialism implies canonically that betting on the weakly more likely event is to be weakly preferred, as expressed by the following condition. For all $A, B \in \Sigma$ and $x, y \in X$ such that $x \succ y$:

$$[x \text{ on } A; y \text{ on } A^c] \succsim [x \text{ on } B; y \text{ on } B^c] \text{ whenever } A \trianglerighteq B. \tag{6}$$

Note that condition (6) can be viewed as a unidirectional version of Savage's behavioral definition of revealed likelihood. Condition (6) asks to be complemented by an analogous condition entailing strict rather than weak preferences. At first sight it seems natural to formulate such a condition by simply replacing $\trianglerighteq$ with its asymmetric component $\triangleright$. However, if $\trianglerighteq$ is incomplete, the resulting condition would however be overly restrictive, as illustrated by the following example.[21]

Example 4. Let $X = \{x, y\}$ with $x \succ y$, and assume that acts (bets) are ranked according to the lower probability $\min_{\pi\in\Pi} \pi(f^{-1}(x))$ of the superior outcome, i.e. that

$$[x \text{ on } A; y \text{ on } A^c] \succsim [x \text{ on } B; y \text{ on } B^c] \text{ iff } \min_{\pi\in\Pi} \pi(A) \geq \min_{\pi\in\Pi} \pi(B).$$

[20] Range-convexity can be derived from Theorem 2 if Σ is a σ-algebra. It may also follow from the specific structure of the likelihood relation; for example, any superrelation of the "Anscombe-Aumann relation" $\trianglerighteq_{\text{rand}}$ in example 3 has a convex-ranged representation, even though the domain of $\trianglerighteq_{\text{rand}}$, the product-algebra $\Sigma_1 \times \Sigma_2$, is not a σ-algebra if the generic state space Ω_1 is infinite.

[21] I thank Simon Grant for emphasizing this point.

Suppose that $\trianglerighteq$ is such that there exists an event E with $\max_{\pi\in\Pi}\pi(E) > 0$ while $\min_{\pi\in\Pi}\pi(E) = 0$. In this case $[x \text{ on } E; y \text{ on } E^c] \sim [x \text{ on } \emptyset; y \text{ on } \Omega]$ even though $E \vartriangleright \emptyset$, violating the envisaged asymmetric counterpart to condition (6).

The difficulty illustrated in Example 4 can be overcome by making use of the *uniform* (rather than merely asymmetric) component $\vartriangleright\!\vartriangleright$ of a coherent likelihood relation defined as follows.

Definition 2 (Uniformly More Likely) *$A \vartriangleright\!\vartriangleright B$ ("A is uniformly more likely than B") if and only if there exists finite partitions of A and B^c, $A = \Sigma_{i\in I} A_i$ and $B^c = \Sigma_{j\in J} B_j$, such that $A \backslash A_i \trianglerighteq B \cup B_j$ for all $i \in I$ and $j \in J$.*

The following Fact shows that the definition indeed captures the notion of "uniformly more likely events" if the context is in fact coherent and equidivisible.

Fact 2 *For any likelihood ordering $\trianglerighteq$, $A \vartriangleright\!\vartriangleright B$ implies $\min_{\pi\in\Pi}[\pi(A) - \pi(B)] > 0$. The converse holds if $\trianglerighteq$ is equidivisible.*

Definition 2 leads to the following asymmetric counterpart of condition (6). For all $A, B \in \Sigma$ and $x, y \in X$ such that $x \succ y$:

$$[x \text{ on } A; y \text{ on } A^c] \succ [x \text{ on } B; y \text{ on } B^c] \text{ whenever } A \vartriangleright\!\vartriangleright B. \tag{7}$$

The following axiom called "Likelihood Compatibility" (LC) extends these conditions to multi-valued acts. The idea is that if two acts differ only in the states in which two particular consequences are realized, then the likelihood comparison of the corresponding events is a *decisive* criterion for their preference comparison.

Axiom 11 (Likelihood Compatibility) *For all $f \in \mathcal{F}$, $x, y \in X$ and events $A, B \in \Sigma$:*
i) $A \trianglerighteq B$ and $x \succsim y$ imply

$$[x \text{ on } A\backslash B; y \text{ on } B\backslash A;\ f(\omega) \text{ elsewhere}] \succsim [x \text{ on } B\backslash A; y \text{ on } A\backslash B;\ f(\omega) \text{ elsewhere}], \text{ and}$$

ii) $A \vartriangleright\!\vartriangleright B$ and $x \succ y$ imply

$$[x \text{ on } A\backslash B; y \text{ on } B\backslash A;\ f(\omega) \text{ elsewhere}] \succ [x \text{ on } B\backslash A; y \text{ on } A\backslash B;\ f(\omega) \text{ elsewhere}].$$

If $(\succsim, \trianglerighteq)$ satisfies LC, we shall also say that preferences are *compatible* with the context $\trianglerighteq$. Note that, considering the case $B = \emptyset$ and exploiting transitivity, LC entails the following weak version of Savage's axiom P3.[22]

[22] Also, since coherence entails Additivity, the events $A, B \in \Sigma$ could have been assumed disjoint.

Axiom 12 (Eventwise Monotonicity) *For all acts* $f \in \mathcal{F}$, *consequences* $x, y \in X$ *and events* $A \in \Sigma$:

i) $[x \text{ on } A; f(\omega) \text{ elsewhere}] \succsim [y \text{ on } A; f(\omega) \text{ elsewhere}]$ *whenever* $x \succsim y$, *and*

ii) $[x \text{ on } A; f(\omega) \text{ elsewhere}] \succ [y \text{ on } A; f(\omega) \text{ elsewhere}]$ *whenever* $x \succ y$ *and* $A \rhd\rhd \emptyset$.

3.2 Likelihood Compatibility: Discussion

1. Unidirectional Constraint on Preferences vs. Bidirectional Definition of Revealed Likelihood The unidirectionality in the statement of the Likelihood Compatibility axiom is essential. It cannot be made bidirectional, and thereby transformed into a behavioral definition of "revealed likelihood" $\trianglerighteq_*$, without sacrificing coherence. To see this, define $\trianglerighteq_*$ as follows:

$A \trianglerighteq_* B$ iff, for all $f \in \mathcal{F}$ and $x, y \in X$ such that $x \succsim y$:

$[x \text{ on } A \backslash B; y \text{ on } B \backslash A;\ f(\omega) \text{ elsewhere}] \succsim [x \text{ on } B \backslash A; y \text{ on } A \backslash B;\ f(\omega) \text{ elsewhere}]$.

Note that, in particular, $A \trianglerighteq_* A^c$ iff, for all $x, y \in X : [x \text{ on } A; y \text{ on } A^c] \succsim [x \text{ on } A^c; y \text{ on } A]$.

Example 5. Return to the 2-color Ellsberg paradox discussed in the introduction, with $X = \{0, 1\}$ and $1 \succ 0$, and

$$[1 \text{ on } R_{kn}, 0 \text{ on } B_{kn}] \sim [1 \text{ on } B_{kn}, 0 \text{ on } R_{kn}] \succ [1 \text{ on } R_{un}, 0 \text{ on } B_{un}] \sim [1 \text{ on } B_{un}, 0 \text{ on } R_{un}].$$

Evidently,

$$R_{kn} \equiv_* B_{kn} \text{ and } R_{un} \equiv_* B_{un}, \text{ but not } R_{kn} \equiv_* R_{un}.$$

Thus $\trianglerighteq_*$ is not coherent, as it violates the Splitting axiom. Likewise, preferences $\succsim$ are not compatible with any coherent superrelation $\trianglerighteq'$ of $\trianglerighteq_*$, since for any such relation $R_{kn} \equiv' R_{un}$ which is incompatible with the preference $[1 \text{ on } R_{kn}, 0 \text{ on } B_{kn}] \succ [1 \text{ on } R_{un}, 0 \text{ on } B_{un}]$. Thus $\trianglerighteq_*$ does not identify well-defined probabilistic beliefs that can be meaningfully attributed to the decision-maker.

Since any interesting model of preferences under ambiguity needs to accommodate the Ellsberg paradox, and since Example 5 pertains to any preference ordering satisfying Savage's axiom P4, the example shows that there is no hope to make $\trianglerighteq_*$ a viable definition of revealed probabilistic beliefs under ambiguity by appropriate behavioral assumptions on preferences. *It is exactly this failure of* $\trianglerighteq_*$ *that motivates our treatment of probabilistic beliefs as an independent entity.*

Also, since $\trianglerighteq_*$ would appear to be the most natural candidate for such a direct definition, its failure suggests that a workable, behaviorally general definition of a coherent revealed likelihood ordering $\trianglerighteq^{rev}$ from preferences will not be easy to attain. Likelihood Compatibility of preferences with $\trianglerighteq^{rev}$ is necessary but not sufficient. A natural line of attack is to focus on maximal coherent relations $\trianglerighteq$ with which $\succsim$ is compatible.[23]

2. Minimal Behavioral Interpretation of Likelihood Orderings Likelihood Compatibility can be viewed as a rationality axiom relating independently given beliefs and preferences, or, alternatively, as part of the very meaning of asserting a likelihood comparison in the first place. On the latter interpretation (to which the author is inclined himself), a DM cannot make a likelihood judgment without an attendant commitment to the preference rankings described by LC. Note, however, that in view of the discussion 3.2.1, a likelihood judgment cannot be *identified* with such a behavioral commitment.

On either of these two views, LC endows likelihood orderings with a direct (unidirectional) behavioral correlate, a "minimal behavioral interpretation" in the spirit (but not implementation) of Walley (1991, section 1.4.4). This constitutes an important *decision-theoretic* justification for taking likelihood orderings as primitive representations of beliefs.

By contrast, the behavioral interpretation of sets of priors is not clear, especially if behavioral generality is desired. This deficiency is in fact Walley's (1991, section 3.8.6) main argument against taking sets of priors as primitive representations of imprecise probabilistic beliefs.[24]

Similarly, it is sometimes suggested in discussion that one might model beliefs as complete but non-additive orderings of events. To the extent that "beliefs" in this context are more than a mere relabeling of preferences and are taken to have genuine epistemic content, the question arises how to formulate counterpart conditions to coherence. But even if this question is resolved (or bracketed or dissolved), it is not clear how such non-additive orderings of events would constrain behavior.[25]

[23]For a concrete proposal in this direction, see Nehring (2001, section 3).

[24]Siniscalchi (2006) makes an interesting proposal to behaviorally identify individual priors in the representing set. His work can be seen as a refined of the notion of "perceived ambiguity" proposed by Ghirardato et al. (2004) discussed in section 4.2 below, and relies on significant additional behavioral assumption which rule out, for example, probabilistically sophisticated preferences that are not SEU.

[25]Consider, for example, "lower probability orderings" $\geq$ on Σ with a representation $A \geq B$ iff $\min_{\pi \in \Pi} \pi(A) \geq \min_{\pi \in \Pi} \pi(B)$; the natural constraint on behavior would be to require LC relative to $\trianglerighteq(\Pi)$, but this fails to provide specific behavioral restrictions associated with *single* $\geq$-comparisons.

3. Behavioral Generality vs. Bernoullian Rationality To maintain behavioral generality, Likelihood Compatibility exploits ordinal information about consequences only, and entails no restrictions on preferences over unambiguous (risky) acts besides monotonicity with respect to stochastic dominance; see the next subsection for details. Much stronger and potentially more controversial normative restrictions can be obtained if one exploits cardinal information about comparisons of utility differences. A theory of "utility sophisticated preferences" along such lines is developed in the companion paper Nehring (2007); some relevant connections are established in section 4.2 below.

4. Incomplete vs. Exhaustive Interpretation of the Likelihood Relation Besides 'ordinality', a second important aspect of the minimalism of LC is its "hereditariness", i.e. the property that compatibility of $\succsim$ with $\trianglerighteq$ implies compatibility of $\succsim$ with any subrelation $\trianglerighteq' \subseteq \trianglerighteq$. This makes it possible to interpret $\trianglerighteq$ "incompletely" (non-exhaustively) as describing only *some* of the DM's probabilistic beliefs, for example beliefs derived from probabilistic information such as the existence of a random device, the composition of an Ellsberg urn, information about frequencies in large populations. In many cases, such information can plausibly be attributed to the DM from the outside, without eliciting them from the DM directly, and LC provides a behavioral criterion that allows to test (falsify) such an attribution.

By contrast, almost all of the other literature that appeals to independently given beliefs attempts to provide models in which beliefs (co-)*determine* preferences respectively final choices; see Jaffray (1989), Nehring (1991, 1992, 2000), Gajdos et al. (2004, 2006) and others quoted in the introduction. A simple example is any model which implies that, given a set of priors Π, the DM maximizes the minimal expected utility $\min_{\pi\in\Pi} E_\pi u\circ f$ relative to the given set Π; clearly, going from Π to a larger set Π' representing less information would typically reverse some preference comparisons. Any such model thus fails to be hereditary. It follows that for such non-hereditary models to be meaningful, it must be assumed that the set Π represents the DM's probabilistic beliefs *exhaustively;* that is, if two events are not compared in terms of $\trianglerighteq$, they must be interpreted as non-comparable. An exhaustive interpretation of $\trianglerighteq$ is clearly much more demanding; for example, the "parametric", fully behavioral use of likelihood relations described in the following paragraphs does not seem sensible on an exhaustive interpretation.

The wider applicability of incompletely interpreted likelihood relation seems especially valuable if the likelihood relation is supposed to capture objective probabilistic information. For in that case, an exhaustive interpretation would amount to assuming not merely that the DM's information has been exhaustively described (which will already be hard to achieve in practice), but that the DM has no subjective beliefs whatever beyond those that are directly entailed by this objective information. This assumption seems quite strong and restrictive, and need not be made on an incomplete interpretation.

5. First-Person vs. Third-Person Point of View: The Likelihood Relation as a Primitive or as a Parameter In the above discussion, we have interpreted the likelihood relation $\trianglerighteq$ as describing a (possibly non-exhaustive) list of the decision-maker's dispositions to affirm particular likelihood comparisons. From the first-person point of view of the decision-maker himself, it seems eminently sensible to posit probabilistic beliefs as distinct entities in this way, for otherwise it is difficult to see how the decision-maker can invoke particular beliefs as *grounds* for the evaluation of uncertain prospects. Indeed, a substantial part of the discipline of "decision analysis" is devoted to articulating the decision-maker's beliefs and bringing them to bear on the decision-problem at hand.

By contrast, economics as an empirical discipline takes the point of view of an outside observer. We would submit that also from this "third-person" point of view, the study of a decision-maker's beliefs via direct questioning should not be taboo, notwithstanding its clear limitations[26]. Nonetheless, in contrast to this position, many economists subscribe to the *behaviorist* view according to which statements about beliefs as independent propositional attitudes are non-observable and thus lack empirical content. Does a behaviorist position render the notion of decision-making in an epistemic context empirically empty?

Evidently Likelihood Compatibility as a relation between preferences and likelihood judgements loses empirical content once the latter cease to be an empirically meaningful entity on their own. Empirical content can be regained, however, if Likelihood Compatibility is absorbed into a behavioral definition of compatibility with an epistemic context: simply say that a preference relation $\succsim$ satisfies *Compatibility-with-*$\trianglerighteq$ if the pair $(\succsim, \trianglerighteq)$ satisfies LC. Here the likelihood relation is "imputed" by the analyst without any truth-claims regarding the beliefs as such.

The analogy with continuity conditions on preferences may be helpful. Just as "*Compatibility-with-*$\trianglerighteq$" conditions, continuity conditions refer in their statement to an "imputed" topology τ that is itself not derived from behavior. Just like the truth-value of "continuity-relative-to τ", that of "Compatibility-with-$\trianglerighteq$" is determined by preferences alone; the behavioral content of either type of

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

condition is therefore clear-cut.[27]

One can summarize this behaviorist use of likelihood relations by saying that the context $\trianglerighteq$ represents an epistemic *parameter* constraining preferences rather than an independent epistemic *primitive*. From this point of view, results on likelihood relations such as Theorem 2 can be viewed as meta-propositions that demarcate for which relations $\trianglerighteq$ the preference condition "Compatibility-with-$\trianglerighteq$" is epistemically meaningful.

A parametric interpretation matches quite well the role played by probabilistic constraints in economic applications. In existing applied work, be it in microeconomics, macroeconomics, finance or game theory, such constraints are motivated mostly by direct appeal to intuitive descriptions of the probabilistic uncertainty faced by the agents, often with a strikingly free use of frequentist language that refers to randomly drawn types, underlying process of asset returns, etc. to the distress of anyone raised in the Savage tradition. The resulting model then makes behavioral predictions that can be confronted with empirical data. (Only in very rare cases will the model be motivated directly by informed assumptions about preferences over Savage acts or something similar.) In the same way, the imputation of likelihood relations can be understood parametrically to entail falsifiable predictions about an agent's behavior without the need to ascertain the truth-value of these imputations in themselves.

3.3 Probabilistic Sophistication on Unambiguous Events

Of particular interest are preferences over acts whose outcomes have well-defined probabilities; such acts will be called unambiguous. Compatibility of preferences with a convex-ranged likelihood relation implies probabilistic sophistication of preferences over unambiguous acts in the sense of Machina-Schmeidler (1992). To make this precise, we need the following definitions. Say that $B \in \Sigma$ is *unambiguous given* A if, for some $\alpha \in [0,1]$, $\pi(B) = \alpha\pi(A)$ for all $\pi \in \Pi$. Let Λ_A denote the family of events $B \in \Sigma$ that are unambiguous given A; clearly, Λ_A is closed under finite disjoint union and complementation, but not necessarily under intersection. In the terminology of Zhang (1999), each Λ_A is a λ-*system* with the property that $B \in \Lambda_A$ iff $B \cap A \in \Lambda_A$. An event A is *null* if $A \equiv \varnothing$, or, equivalently, if $\pi(A) = 0$ for all $\pi \in \Pi$. For any non-null A and any arbitrary $\pi \in \Pi$,

[27] Indeed, it is easily verified that, given an ordering over outcomes (constant acts), "Compatibility-with-$\trianglerighteq$" boils down to the requirement that the preference relation $\succsim$ contain a partial ordering $\succsim_{\trianglerighteq}$ that mirrors the structure of $\trianglerighteq$.

let $\overline{\pi}(.|A)$ denote the restriction of the conditional probability measure $\pi(.|A)$ to Λ_A. We will say that B is *unambiguous* if it is "unambiguous given Ω", and write Λ for Λ_Ω, as well as $\overline{\pi}$ for $\overline{\pi}(.|\Omega)$. An act $f \in \mathcal{F}$ is *unambiguous* if, for all $x \in X$, $\{\omega \in \Omega \mid f(\omega) = x\}$ is unambiguous; let $\mathcal{F}^{ua}$ denote their set. A "lottery" q is a probability distribution on X with finite support, and will be written as $q = (q^x)_{x \in X}$, where q^x denotes the probability of obtaining x under q; let $\mathcal{L}$ denote their set. The unambiguous act f induces the lottery $\overline{\pi} \circ f^{-1}$ with $\left(\overline{\pi} \circ f^{-1}\right)^x = \overline{\pi}\left(\{\omega \in \Omega \mid f(\omega) = x\}\right)$. The lottery p *stochastically dominates* the lottery q if, for all $y \in X$, $\sum_{x:x \succsim y} p^x \geq \sum_{x:x \succsim y} q^x$; p *stochastically dominates* q *strictly* if at least one of these inequalities is strict. An ordering $\succsim_\mathcal{L}$ is *monotone* (with respect to stochastic dominance) if, for all $p, q \in \mathcal{L}$, $p \succsim_\mathcal{L} q$ whenever p stochastically dominates the lottery q, and $p \succ_\mathcal{L} q$ whenever p stochastically dominates the lottery q strictly.

Definition 3 (Probabilistic Sophistication on Unambiguous Events) *The preference ordering $\succsim$ is probabilistically sophisticated on unambiguous events if there exists a monotone ordering $\succsim_\mathcal{L}$ on $\mathcal{L}$ such that, for all $f, g \in \mathcal{F}^{ua}$,*

$$f \succsim g \text{ if and only if } \overline{\pi} \circ f^{-1} \succsim_\mathcal{L} \overline{\pi} \circ g^{-1}.$$

Note that, by the range-convexity of $\trianglerighteq$, the mapping $f \mapsto \overline{\pi} \circ f^{-1}$ is onto; the ordering $\succsim_\mathcal{L}$ in this representation is therefore uniquely defined. Following Machina-Schmeidler (1992), $\succsim_\mathcal{L}$ can be viewed as capturing the decision-makers' risk preferences that become analytically separate from his beliefs and, in the present more general context, from his preferences over non-unambiguous acts.

Proposition 1 *If the weak order $\succsim$ is compatible with the coherent and convex-ranged likelihood relation $\trianglerighteq$, it is probabilistically sophisticated on unambiguous events.*

If the set of unambiguous events was an algebra rather than a λ-system, Proposition 1 could be derived straightforwardly by copying from the proof of Machina-Schmeidler's (1992) main result. Their proof does not apply as is, since the set of unambiguous events is not necessarily closed under intersection. However, range-convexity entails "enough" intersection closedness to make use of their proof nonetheless.[28]

[28] Proposition 1 and its proof have sigificant parallels to a recent (and prior) purely behavioral result of Kopylov (2003).

4. MULTI-PRIOR CHARACTERIZATION OF LIKELIHOOD COMPATIBILITY

The implications of Likelihood Compatibility are non-trivial, even in the case of very well-behaved preferences. Consider, for example, the Minimum EU model with preferences given by

$$f \succsim g \text{ iff } \min_{\pi \in \Psi} E_\pi u \circ f \geq \min_{\pi \in \Psi} E_\pi u \circ g, \tag{8}$$

for some utility function $u : X \to \mathbf{R}$ and some closed, convex set of probability measures Ψ; for axiomatizations of MEU preferences over Savage acts (as pertinent here), see Casadesus et al. (2000) and Ghirardato et al. (2003). *When are MEU preferences with representation* (u, Ψ) *compatible with imprecise probabilistic beliefs represented by the set* Π? The answer to this question is not obvious, either on direct "intuitive grounds", nor given the formal definition of compatibility proposed in this paper; whatever the correct answer is, it cannot be taken ready-made from the literature.

In particular, while the inclusion $\Psi \subseteq \Pi$ implies compatibility of preferences with the associated likelihood relation $\trianglerighteq_\Pi$, the converse may fail. To see this, consider the case of a singleton $\Pi = \{\pi\}$, with π convex-ranged. According to Proposition 1, compatibility with the associated likelihood relation $\trianglerighteq_{\{\pi\}}$ is equivalent to probabilistic sophistication with respect to π. On the other hand, if $\Psi \subseteq \Pi = \{\pi\}$, that is if $\Psi = \{\pi\}$, MEU preferences must maximize expected utility. But it is well-known that in the MEU model probabilistic sophistication with respect to π does not entail expected utility maximization.[29]

So when does the eminently applicable and heuristically attractive multi-prior characterization of Likelihood Compatibility as equivalent to the inclusion $\Psi \subseteq \Pi$ hold? We shall obtain the answer as a corollary of a characterization result on the much broader class of "invariant biseparable" (i.b.) preferences to be introduced shortly, Theorem 3. To facilitate the reader's orientation, it helps to give away the punchline for the MEU case. Here, the characterization turns out to be particularly simple. Indeed, it is straightforward from the definition of the set of unambiguous events Λ that for the desired characterization $\Psi \subseteq \Pi$ to obtain in the MEU model, preferences must maximize expected utility maximizer under "risk" (i.e. with respect to unambiguous acts); more interestingly, it follows from 3 that this condition is also sufficient.

As indicated, the main goal of this section will be a characterization of Likelihood Compatibility for the class of preferences over Savage acts with an "invariant biseparable" (i.b.) representation which has been introduced axiomatically by Ghirardato et al. (2003,2004) and includes, for example,

[29] See in particular Marinacci (2002).

the CEU and the α-MEU models.[30] A preference ordering over $\mathcal{F}$ is *invariant biseparable* if there exists an onto utility function $u : X \to [0,1]$ and an evaluation functional $I : \mathcal{Z} \to [0,1]$ that is monotone, positively homogeneous ($I(cZ) = cI(Z)$ for all $c \in [0,1]$), and constant additive ($I(Z + b1_\Omega) = I(Z) + b$ for all applicable $b \in [0,1]$) such that, for all $f, g \in \mathcal{F}$,

$$f \succsim g \text{ if and only if } I(u \circ f) \geq I(u \circ g). \tag{9}$$

The class of invariant biseparable preferences is a natural choice for the study of Likelihood Comparability since it is the largest known class for which an associated set of priors can be defined. Specifically, in earlier work (Nehring 1996), and subsequently in Nehring (2001) and Ghirardato et al. (2004, henceforth: GMM), it has been shown that, with preference in this class, there exists a maximal independent subrelation a la Bewley (1986/2002) with an associated set of priors $\Psi^* \in \mathcal{K}(\Delta(\Sigma))$.[31] For expository economy, here we define this set directly using the following notion of utility sophistication that has been given an axiomatic behavioral foundation in Nehring (2001,2007) and that expresses a notion of Bernoullian rationality in the presence of ambiguity.[32] An invariant biseparable preference ordering $\succsim$ with representation (I, u) is *utility sophisticated* with respect to a set of priors $\Pi \in \mathcal{K}(\Delta(\Sigma))$ if, $f \succsim g$ whenever $E_\pi u \circ f \geq E_\pi u \circ g$ for all $\pi \in \Pi$. Note that utility sophistication with respect to Π implies Likelihood Compatibility with respect to $\trianglerighteq_\Pi$, but the converse is far from true. For example, $\succsim$ is utility-sophisticated with respect to a singleton set $\{\mu\}$ if and only if it maximizes expected utility with respect to μ.

We define the *set of Bernoulli priors* associated with $\succsim$ Ψ^* as the unique minimal set $\Pi \in \mathcal{K}(\Delta(\Sigma))$ such that $\succsim$ is utility-sophisticated with respect to Π. Propositions 4 and 5 in GMM imply that such a unique minimal set Ψ^* exists for any i.b. preference ordering. GMM further characterize Ψ^* as the Clarke derivative of I at 0. To illustrate, in the MEU model, it is easily seen that $\Psi^* = \Psi$[33].

[30] Preference models that violate Invariant Biseparability such as the smooth preferences of Klibanoff et al. (2005) and the variational preferences of Maccheroni et al. (2006) typically violate Savage's axiom P4 which requires that preferences over bets do not depend on the prizes associated with those bets.

[31] Nehring (1996) used the traditional Anscombe-Aumann (1963) framework, while Nehring (2001) reformulated the original proposal in the present setting, demonstrating applicability to a finite setting via an embedding argument similar to that behind Proposition 4 below; GMM formulate their treatment assuming a general convex consequence space with an associated mixture operation. Our formal definition of Ψ^* in this section corresponds to an interpretation of this mixture operation as a mixture of utitilities as described in Ghirardato et al. (2003) rather than as a mixture of lotteries.

[32] Gajdos et al. (2006) formulate a Dominance axiom relative to a set Π, which in the context of their other axioms amounts to utility sophistication relative to Π as defined here.

[33] Cf. GMM (p. 151).

The situation is somewhat more complicated in the case of α-MEU preferences; these are given by an (I, u) representation of the form

$$I(Z) = \alpha \min_{\pi \in \Psi} E_\pi Z + (1 - \alpha) \max_{\pi \in \Psi} E_\pi Z,$$

with $\alpha \in [0, 1]$. Here it is easily verified that $\Psi^* \subseteq \Psi$. Yet, as pointed out by Eichberger et al. (2006), the converse implication fails frequently if $\alpha \notin \{0, 1\}$.

We want to know when it is possible to determine whether a preference ordering is compatible with some likelihood relation $\trianglerighteq_\Pi$ by looking at its set of Bernoulli priors Ψ^*. Specifically, under which conditions Likelihood Compatibility is equivalent to the inclusion $\Psi^* \subseteq \Pi$? To the extent that the literature suggests a method to impose probabilistic constraints on preferences, it is probably captured by this inclusion.

From the definition of Ψ^*, it follows immediately that if $\Psi^* \subseteq \Pi$, then preferences maximize EU under risk. Thus EU maximization under risk is necessary for the multi-prior characterization of Likelihood Compatibility as equivalent to the inclusion $\Psi^* \subseteq \Pi$ to hold. In the MEU and, more generally, α-MEU models with $\alpha \neq \frac{1}{2}$, this is also sufficient. More generally, for example in the case of CEU preferences, EU maximization under risk may fail to ensure the multi-prior characterization to be valid. Nonetheless, in those cases, a somewhat stronger condition still suffices.

This condition requires that all unambiguous events be unambiguous with respect to Ψ^* ("Bernoulli unambiguous"). An event A is *Bernoulli unambiguous* iff all Bernoulli priors agree on its probability: $\pi(A) = \pi'(A)$ for all $\pi, \pi' \in \Psi^*$; their class is denoted by Λ^*.[34] The sufficiency of this stronger condition is asserted by the following result; it also provides for a second sufficient condition for the multi-prior characterization to apply whose usefulness will be explained further below.

Theorem 3 *Let $\succsim$ be an invariant biseparable preference ordering with a set of Bernoulli priors Ψ^*. Let $\trianglerighteq$ be a coherent and convex-ranged likelihood relation. Then the following four statements are equivalent:*

1. $\Psi^* \subseteq \Pi$ *;*

2. $\succsim$ *is utility-sophisticated with respect to* $\trianglerighteq$*;*

3. $\succsim$ *is compatible with* $\trianglerighteq$*, and all unambiguous events are Bernoulli unambiguous* $(\Lambda \subseteq \Lambda^*)$*;*

[34]The class of Bernoulli unambiguous events has been introduced in Nehring (1999) under a different definition; that definition is shown to be equivalent to the one given here in Nehring (2001). Bernoulli unambiguous events are further studied in the companion paper to GMM Ghirardato et al. (2005).

4. $\succsim$ *is compatible with* $\trianglerighteq$ *and utility-sophisticated relative to some convex-ranged subrelation* $\trianglerighteq_0$ *.*

The equivalence 1)$\Longleftrightarrow$2) and the implication 2)$\Longrightarrow$3) are definitional. By contrast, the implications 3)$\Longrightarrow$4) and 4)$\Longrightarrow$2) are substantive and not obvious; in both, range-convexity plays a crucial role.[35]

Criterion 1: EU maximization under risk.—

The characterization in terms of Bernoulli unambiguous events is especially useful if the latter have a simple characterization themselves. A particularly simple and useful one is given by the definition of unambiguous events proposed by Ghirardato-Marinacci (2002); in terms of the (I, u)-representation, an event A is *GM unambiguous* ($A \in \Lambda^{GM}$) if $I(1_A) + I(1_{A^c}) = 1$, or, equivalently, if preferences maximize expected utility on $\{A, A^c\}$-measurable acts.[36] Say that an i.b. preference ordering $\succsim$ has the *GM property* if $\Lambda^{GM} = \Lambda^*$. The usefulness of this definition emerges from the following observation.

Observation 1 *If* $\succsim$ *maximizes expected utility under risk and has the GM property,* $\Lambda \subseteq \Lambda^*$*.*

To see this, simply note that if $\succsim$ maximizes EU under risk, then $\Lambda \subseteq \Lambda^{GM}$. Hence $\Lambda \subseteq \Lambda^*$ by the GM property. The observation yields the following immediate corollary to Theorem 3:

Proposition 2 *Let* $\succsim$ *be an invariant biseparable preference ordering satisfying the GM property, and let* $\trianglerighteq$ *be a coherent and convex-ranged likelihood relation. Then* $\Psi^* \subseteq \Pi$ *if and only if* $\succsim$ *is compatible with* $\trianglerighteq$ *and* $\succsim$ *maximizes expected utility under risk.*

It is obvious that the GM property holds for MEU preferences: simply note that $\min_{\pi \in \Psi} \pi(A) + \min_{\pi \in \Psi} \pi(A^c) = 1$ if and only if $\pi(A) = \pi'(A)$ for all $\pi \in \Psi$. More generally, the following Fact shows that the GM property holds for α-MEU preferences whenever α is different from $\frac{1}{2}$. By contrast, if $\alpha = \frac{1}{2}$, $\Lambda^{GM} = \Sigma$, and the GM property fails to hold.

Fact 3 *For any* α*-MEU preference with* $\alpha \neq \frac{1}{2}$, $\Lambda^{GM} = \Lambda_\Psi = \Lambda^*$*.*

Ghirardato et al. (2005) have shown before that $\Lambda^{GM} = \Lambda^*$ for the subset of α-MEU preferences that satisfy the fix-point property $\Psi = \Psi^*$.[37] The validity of the more generally valid Fact 3 hinges on

[35]The last is in fact derived from a basic result in the companion paper Nehring (2007).

[36]Inspite of this notation, Λ^{GM} need not be a λ-system.

[37]This is the class of α-MEU preferences axiomatized in GMM. Ghirardato et al. (2005) obtained their result as a consequence of a more general result on i.b. preferences that however does not imply Fact 3.

the insight – which is interesting in its own right– that Ψ-unambiguous events are faithfully revealed, i.e. that $\Lambda_\Psi = \Lambda^*$, even when Ψ^* is contained in Ψ strictly, violating the fix-point property.

In the absence of the GM property, EU maximization under risk will not, in general, be enough to ensure validity of the multi-prior characterization. Consider, for instance, CEU preferences that are compatible with some convex-ranged $\trianglerighteq$ (and maximize EU under risk) but are not SEU overall. The set of Bernoulli priors associated with such preferences Ψ^* can never be contained in Π: indeed, if it was, Ψ^* would inherit the range-convexity of Π, i.e. $\succsim$ would be utility-sophisticated relative to some convex-ranged set of priors; but by Proposition 6 in Nehring (2007), $\succsim$ would be SEU. So the multi-prior characterization of Likelihood Compatibility is never applicable to CEU preferences, whether or not they maximize EU under risk.[38]

Criterion 2: Utility Sophistication under Randomization.—

The second characterization in Theorem 3 is especially useful in the case of $\trianglerighteq_0 = \trianglerighteq_{\text{rand}}$, since many ambiguity models that have been proposed in the literature within the Anscombe-Aumann framework, when appropriately translated into the present setting,[39] are utility-sophisticated with respect to $\trianglerighteq_{\text{rand}}$.

In this setting, the multi-prior characterization can be sharpened further, enhancing its applicability substantially. Let $\mathcal{F}_1$ denote the set of acts that depend on the uncertainty about generic states only (Σ_1-measurable or "primary" acts). Note first that any $\Pi \subseteq \Pi_{\text{rand}}$ can be written as $\Pi_1 \times \eta := \{\pi_1 \times \eta : \pi_1 \in \Pi_1\}$, where $\Pi_1 \in \mathcal{K}(\Delta(\Sigma_1))$ is a set of priors on generic events. In like manner, if $\succsim$ is utility-sophisticated with respect to Π_{rand} ($\succsim$ is" utility-sophisticated under randomization", we shall say), Ψ^* can be written as $\Psi_1 \times \eta$. The following Proposition 3 shows that in fact

[38]The natural diagnosis is that for such CEU preferences, Ψ^* *necessarily* overstates the 'true' ambiguity. Range-convexity of $\trianglerighteq$ is not essential for the thrust of this conclusion, only for its sharpness. Example 6 below may serve as an illustration.

In the case of CEU preferences, the inapplicability of the multi-prior characterization is not bothersome in any case since Likelihood Compatibility can easily be read off the representing capacity ν. Indeed it is easily verified that a CEU preference with representing capacity ν is compatible with the likelihood ordering $\trianglerighteq$ if and only if, for all $A, B \in \Sigma$, $\nu(A) \geq \nu(B)$ whenever $A \trianglerighteq B$.

[39]In one translates preferences formulated for the standard AA-model into the present framework, utility sophistication under randomization is equivalent to satisfaction of Monotonicity and Lottery Independence of the original AA preferences; see Nehring (2007b). These assumptions are satisfied by almost all models of preferences under ambiguity formulated in this framework; an exception is Seo (2006). For utility sophisticated preferences in this setting, Invariant Biseparabilty is equivalent to Savage's P4.

$\Psi_1 = \Psi_1^*$, where Ψ_1^* is the set of priors revealed by the restriction of the given preference relation $\succsim_{|\mathcal{F}_1}$ to the set of "primary acts".

Proposition 3 *Let $\succsim$ be an invariant biseparable preference ordering that is utility-sophisticated under randomization. Then $\succsim$ is compatible with $\trianglerighteq$ if and only if $\Psi_1^* \subseteq \Pi_1$.*

Proposition 3 provides a usable criterion of compatibility that can be employed in arbitrary (in particular: in finite) state spaces. To illustrate, assume that preferences over the set of primary acts $\succsim_{|\mathcal{F}_1}$ are observed 'in practice' (in the field), but that preferences over general acts $\succsim$ are observable 'in principle' (in the lab, say). Moreover, suppose that it is known (or hypothesized) that the DM is utility-sophisticated with respect to Π_{rand}. Proposition 3 then provides then a usable criterion to check on the basis of the observed preferences $\succsim_{|\mathcal{F}_1}$ alone whether these are compatible with some attributed likelihood relation $\trianglerighteq$.

As an example illuminating the force of Proposition 3, consider the special case in which preferences over *primary* acts are CEU. Indeed, such preferences are the counterpart in the present framework to Schmeidler's (1989) model of CEU preferences in the Anscombe-Aumann setting. By Proposition 3, the set of Bernoulli priors over generic events Ψ_1^* can be used to characterized Likelihood Compatibility. This is useful especially since this set can be explicitly using the general characterization of sets of Bernoulli priors for CEU preferences stated in GMM (example 17)[40]. As shown in Nehring (1999), the CEU functional form severely restricts the range of values that the set of Bernoulli priors may take. In particular, it has been shown there that the associated set of Bernoulli unambiguous events must necessarily be an algebra. This severely constrains the type of ambiguity situations that can be accommodated by Schmeidlerian CEU preferences, as illustrated by the following example.

Example 6. Consider a four-color urn with 100 balls.[41] The decision-maker knows that 50 balls are white or yellow, and that 50 balls are white or red. The generic states are given by the color of the drawn ball, $\Omega_1 = \{W, Y, R, B\}$ and $\Sigma_1 = 2^{\Omega_1}$. This probabilistic information is captured by the set of priors $\Pi_1 = \{\pi \in \Delta(\Sigma_1) : \pi(W \cup Y) = \pi(W \cup R) = \frac{1}{2}\}$. In particular, the associated set of unambiguous events Λ_1 is given by the family of events $\{\{W, Y\}, \{R, B\}, \{W, R\}, \{Y, B\}, \emptyset, \Omega_1\}$. Suppose, conterfactually, that the DM had proper (non-SEU) Schmeidlerian preferences that were compatible with the likelihood ordering $\trianglerighteq_{\Pi_1 \times \eta}$. By Proposition 3, $\Psi_1^* \subseteq \Pi_1$, hence $\Lambda_1^* \supseteq \Lambda_1$. By

[40]GMM's result extends an earlier result reported in Nehring (1996).

[41]Cf. Zhang (1999) and Nehring (1999).

Theorem 2 in Nehring (1999), Λ_1^* is an algebra, and must thus be equal to Σ_1; but this means that $\succsim$ must be SEU, the desired contradiction.

5. OUTLOOK

In the second part of the paper, we have then explored the implications of coherent convex-ranged probabilistic beliefs for decision making in a behaviorally general, hence minimalistic spirit. Not just in this case but quite a bit more generally, it may well turn out that richness assumptions such as range-convexity are unavoidable in the clarification of fundamental issues in the theory of decision-making under ambiguity. Paraphrasing Aliprantis and Border's (1999, p.13) justification of the Axiom of Choice, life without them may well prove to be nasty, brutish and short.

The proposed framework promises to be fruitful setting for further decision theoretic analyses. Indeed, in companion papers (Nehring 2001, 2007), this framework is used 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.

Other uses seem likely. In particular, since the Anscombe-Aumann framework can be viewed as a reduced form of the present one as shown in an earlier working paper version (Nehring 2007c), much of the existing work on decision making under ambiguity can be translated into the present framework of decision-making in the context of probabilistic beliefs. This translation alone will frequently entail gains in clarity and insight.[42]

[42] In the companion paper Nehring (2007), we illustrate this in the context of Gilboa-Schmeidler's original MEU model.

APPENDIX

A.1 Counterexample to Full Expressivity

Let Σ denote the Borel-σ-algebra on the unit interval with Lebesgue measure λ, fix $K > 1$, and define a coherent likelihood relation $\trianglerighteq^K$ as in section 2.4 as follows:

$$A \trianglerighteq^K B \text{ if and only if } \lambda(A \backslash B) \geq K\lambda(B \backslash A). \tag{10}$$

It is easily verified that the associated set of admissible priors $\Pi_{\trianglerighteq^K}$ (which we shall also denote as Π_1^K) consists of all probability measures π with Lebesgue density ϕ such that[43]

$$\text{ess} \sup_{\omega \in [0,1]} \phi(\omega) \leq K \,\text{ess} \inf_{\omega \in [0,1]} \phi(\omega). \tag{11}$$

To see the necessity, if it is not the case that $\operatorname{ess\,sup}_{\omega \in [0,1]} \phi(\omega) \leq K \operatorname{ess\,inf}_{\omega \in [0,1]} \phi(\omega)$, there exists $\epsilon > 0$ and $\omega', \omega'' \in [0,1]$ such that

$$\int_{\omega'}^{\omega'+\epsilon} \phi(\omega)\, d\omega \geq \epsilon (K - \epsilon) \,\text{ess} \sup_{\omega \in [0,1]} \phi(\omega) > \epsilon (K + \epsilon) \,\text{ess} \inf_{\omega \in [0,1]} \phi(\omega) \geq \int_{\omega''}^{\omega''+K\epsilon} \phi(\omega)\, d\omega.$$

Setting $A := [\omega', \omega' + \epsilon]$ and $B := [\omega'', \omega'' + K\epsilon]$, $\lambda(B \backslash A) \geq \lambda(A \backslash B) K$, hence $B \trianglerighteq^K A$.

But for the prior $\pi \in \Pi_{\trianglerighteq^K}$ associated with ϕ,

$$\pi(A) = \int_{\omega'}^{\omega'+\epsilon} \phi(\omega)\, d\omega > \int_{\omega''}^{\omega''+K\epsilon} \phi(\omega)\, d\omega = \pi(B),$$

implying $A \triangleright^K B$, a contradiction.

Sufficiency is straightforward.

By (11), 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 $\trianglerighteq^K$. Yet, as described in the following Fact that is easily verified, Π_1^K and Π_2^K induce different lower probability functions denoted by $\pi_{1,K}^-$ and $\pi_{2,K}^-$.

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

Fact 4 *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 $K = 3$, with $\pi_{1,K}^{-}$ above $\pi_{2,K}^{-}$ and touching at $\lambda = \frac{3}{4}$.

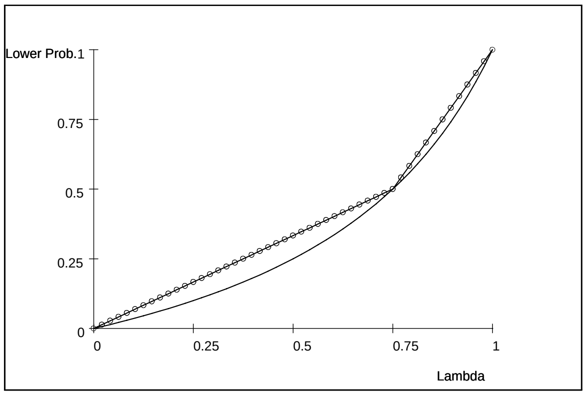

Fig. 1: Two Different Lower Probabilities

A.2 Proofs

Proof of Fact 1.

Take any events A_1,A_2, B_1, B_2 such that $A_1 + A_2 \trianglerighteq B_1 + B_2$, $A_1 \trianglerighteq A_2$ and $B_1 \trianglerighteq B_2$, while not $A_1 \trianglerighteq B_2$. By completeness, $B_2 \vartriangleright A_1$, and thus by transitivity, $B_2 \vartriangleright A_2$ and $B_1 \vartriangleright A_1$. Thus by Strong Additivity (Lemma 1), $B_1 + B_2 \vartriangleright A_1 + A_2$, the desired contradiction.

Lemma 1 *(Additivity implies Strong Additivity) $A \trianglerighteq B$ and $C \trianglerighteq D$ such that $A \cap C = B \cap D = \emptyset$ implies $A + C \trianglerighteq B + D$; moreover, $A + C \vartriangleright B + D$ if in addition $A \vartriangleright B$ or $C \vartriangleright D$.*

This Lemma is standard in derivations of Savage's Theorem; our proof is an adaptation of Fishburn (1970, p. 196). From Additivity, one infers immediately that

$$A + (C \backslash B) \trianglerighteq B + (C \backslash B) = B \cup C = C + (B \backslash C) \trianglerighteq D + (B \backslash C),$$

hence $A + (C \backslash B) \trianglerighteq D + (B \backslash C)$ by transitivity. Applying Additivity and transitivity once more and noting that $B \cap C$ is disjoint from both A and D, one obtains the desired conclusion:

$$A + C = A + (C \backslash B) + (B \cap C) \trianglerighteq D + (B \backslash C) + (B \cap C) = D + B.$$

The second part of the Lemma follows from an exactly parallel argument. □

Proof of Theorem 2.

Necessity of all axioms is straightforward. For sufficiency, 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 in the multi-prior representation to be obtained will 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 $A_i \in \Sigma$ 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 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 (Lemma 1).

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 *$A \in \frac{1}{2^k}C$, $B \in \frac{1}{2^k}D$, and $C \trianglerighteq D$ imply $A \trianglerighteq B$.*

For $k = 0$, the claim is trivial. Suppose it to hold for all $k' < k$. By Lemma 2, there exist events $A' \in \frac{1}{2^{k-1}}C$ such that $A \in \frac{1}{2}A'$ and $B' \in \frac{1}{2^{k-1}}D$ such that $B \in \frac{1}{2}B'$. By induction assumption $A' \trianglerighteq B'$, hence by Splitting $A \trianglerighteq B$.

Lemma 5 *For all $\alpha, \beta \in \mathbf{D}$, $A \in \alpha C$, $B \in \beta D : \alpha \geq \beta$ and $C \trianglerighteq D$ imply $A \trianglerighteq B$.*

Write $\alpha = \frac{\ell}{2^k}$ and $\beta = \frac{\ell'}{2^k}$ with $\ell \geq \ell'$. By definition, there exist partitions $\{A_i\}_{i \leq 2^k}$ and $\{B_i\}_{i \leq 2^k}$of C respectively D into 2^k equally likely elements such that $A = \sum_{i \leq \ell} A_i$ and $B = \sum_{i \leq \ell'} B_i$. Since $A_i \in \frac{1}{2^k}C$ and $B_i \in \frac{1}{2^k}D$, one has $A_i \trianglerighteq B_i$ by Lemma 4. The assertion follows therefore from 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 and has finite range$\}$. Any finite-ranged Z can be canonically written as $\sum_i z_i 1_{E_i}$, where $E_i = Z^{-1}(\{z_i\})$ for all i. For any $Z = \sum z_i 1_{E_i} \in \mathcal{D}$, define

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

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

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

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

Lemma 6 *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 5, $A_i \equiv B_i$. Hence $A \equiv B$ by Strong Additivity.

Lemma 7 *For all $\alpha \in \mathbf{D}$, families of mutually disjoint events $\{E_i\}_{i \in I}$ and families $\{A_i\}_{i \in I}$ such that $A_i \in \alpha E_i$ 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 8 *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 $C \in [Z]$ disjoint from A and B 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 $B \in [Y]$ disjoint from A 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 $z_i = \frac{\ell_i}{2^{k_i}}, x_i = \frac{\ell_i'}{2^{k_i}}$, and $y_i = \frac{\ell_i''}{2^{k_i}}$. Split D_i into 2^{k_i} equally likely events $\{D_{i1,...,}D_{i2^{k_i}}\}$, and set $C_i := \sum_{j \leq \ell_i} D_{ij} \in z_i D_i$, $A_i = \sum_{j=\ell_i+1}^{\ell_i+\ell_i'} D_{ij} \in x_i D_i$, and $B_i = \sum_{j=\ell_i+1}^{\ell_i+\ell_i''} D_{ij} \in y_i D_i$. Note that the sets A_i and B_i are well-defined since $\ell_i + \ell_i' \leq 2^{k_i}$ and $\ell_i + \ell_i'' \leq 2^{k_i}$ because $X + Z \in \mathcal{D}$ and $Y + Z \in \mathcal{D}$. Using Lemma 7, 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 7, 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 7, one infers that $\sum_i B_i \in [Y]$, as well as $C \backslash \left(\sum_i B_i \right) = \sum_i (C_i \backslash B_i) \in [X]$, as desired. □

Lemma 9 *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* $\mathcal{D} \times \mathcal{D}$*) with respect to the sup-norm topology.*

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 6, $B \equiv B'$. 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 5.

To verify Additivity, consider any X, Y, Z such that $X + Z, Y + Z \in \mathcal{D}$. According to Lemma 8i), 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 6, 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, is proved similarly. In view of Lemma 8i), there exist events $A \in [X]$, $A' \in [X']$ such that $A + A' \in [X + X']$, and events $B \in [Y]$, $B' \in [Y']$ such that $B + B' \in [Y + Y']$. By Lemma 6, $A \trianglerighteq B$ and $A' \trianglerighteq B'$, whence by Strong Additivity of $\trianglerighteq$, $A + A' \trianglerighteq B + B'$, and therefore $X + X' \widehat{\trianglerighteq} Y + Y'$.

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 $\frac{1}{2^K}-$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 $\| X' - X \| \leq \frac{1}{2^K}$ and $\| Y' - Y \| \leq \frac{1}{2^K}$, it is not the case that $X' \widehat{\trianglerighteq} Y'$.

To verify this claim, take any $X', Y' \in \mathcal{D}$ such that $\| X' - X \| \leq \frac{1}{2^K}$ and $\| Y' - Y \| \leq \frac{1}{2^K}$. 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 8ii) $A' \in [X' - X]$ such that $A + A' \in [X']$; likewise, by Lemma 8iii), there exist $B' \in [Y - Y']$ and $B'' \in [Y']$ such that $B' + B'' = B$.

Now, A' and B' are $\frac{1}{2^K}-$events. – Indeed, since $X' - X \leq \frac{1}{2^K} 1_\Omega$ (respectively $Y - Y' \leq \frac{1}{2^K} 1_\Omega$), using Lemma 8ii), one can infer the existence of a partition of Ω into 2^K equally likely events such that one of them contains A' (respectively B') – It is thus not the case that $A + A' \trianglerighteq B \backslash B' = B''$. Therefore, in view of Lemma 6, 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 $\widehat{\trianglerighteq}$ is closed in $\mathcal{D}\times\mathcal{D}$, the closure $cl\widehat{\trianglerighteq}$ of $\widehat{\trianglerighteq}$ in $\mathcal{B}\times\mathcal{B}$ restricted to $\mathcal{D}\times\mathcal{D}$ is simply $\widehat{\trianglerighteq}$. 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 10 *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 $\alpha_n\leq\alpha\min\left(\frac{\|X\|}{\|X_n\|},\frac{\|Y\|}{\|Y_n\|}\right)$. This ensures that, for all n, $\ell\alpha_nX_n\in\mathcal{D}$ and $\ell\alpha_nY_n\in\mathcal{D}$. By Weak Homogeneity of $\widehat{\trianglerighteq}$ on $\mathcal{D}$, $\alpha_nX_n\widehat{\trianglerighteq}\alpha_nY_n$ for all n. Hence by $(\ell-1)$-fold application of Strong Additivity of $\widehat{\trianglerighteq}$ on $\mathcal{D}$, also $\ell\alpha_nX_n\widehat{\trianglerighteq}\ell\alpha_nY_n$ for all n. By Continuity on $\mathcal{B}$, $\ell\alpha X\widehat{\trianglerighteq}\ell\alpha Y$, as desired.

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

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

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

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

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

In a final step, extend $\widehat{\trianglerighteq}$ on $\mathcal{B}$ to the set of all bounded random-variables $\mathcal{R} := B(\Sigma, \mathbf{R})$ by defining $\widehat{\trianglerighteq}$ on $B(\Sigma, \mathbf{R})$ as the unique relation $\widetilde{\trianglerighteq}$ on $B(\Sigma, \mathbf{R})$ that coincides on $\mathcal{B}$ with $\widehat{\trianglerighteq}$ on $\mathcal{B}$ and that satisfies Additivity and Homogeneity. (The uniqueness of this extension is immediate; existence follows easily form the Additivity and Homogeneity properties of $\widehat{\trianglerighteq}$ on $\mathcal{B}$). As in section 2.2, say that a relation $\widehat{\trianglerighteq}$ on $\mathcal{R}$ is a *coherent expectation ordering* if it satisfies Transitivity, Reflexivity, Positivity, Non-degeneracy, Homogeneity, Additivity, and Continuity. The following Lemma summarizes the construction, and follows immediately from Lemma 10.

Lemma 11 *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 11. 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. By Additivity and Homogeneity of $\widehat{\trianglerighteq}$, it can be assumed that $X,Y \in \mathcal{B}$. By continuity and monotonicity of $\widehat{\trianglerighteq}$ 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 12 just below, $\pi(A) < \pi(B)$, contradicting the assumption that Π' represents $\trianglerighteq$.

Lemma 12 *For any $\pi \in \Pi'$ such that $\widehat{\trianglerighteq}_{\Pi'} = \widehat{\trianglerighteq}$, and any $X \in \mathcal{D}$ and $A \in [X] : E_\pi X = \pi(A)$.*

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

Proof of Fact 2.

Suppose that there exists finite partitions of A and B^c, $A = \Sigma_{i\in I} A_i$ and $B^c = \Sigma_{j \in J} B_j$ such that $A \backslash A_i \trianglerighteq B \cup B_j$ for all $i \in I$, $j \in J$. By consistency, $\Pi_{\trianglerighteq} \neq \emptyset$. For all $\pi \in \Pi_{\trianglerighteq}$, $\pi(A\backslash A_i) \geq \pi(B)$ for all $i \in I$, hence

$$\pi(A) = \frac{1}{\#I-1}\Sigma_{i\in I}\pi(A\backslash A_i) \geq \frac{\#I}{\#I-1}\pi(B). \tag{12}$$

By the same reasoning, for all $\pi \in \Pi_{\trianglerighteq}$, $\pi(B^c) \geq \frac{\#J}{\#J-1}\pi(A^c)$, and therefore

$$\min_{\pi\in\Pi_{\trianglerighteq}} \pi(A) \geq \frac{1}{\#J}. \tag{13}$$

By (12), $\pi(B) \leq \frac{\#I-1}{\#I}\pi(A)$ for all $\pi \in \Pi_{\trianglerighteq}$, and thus by (13)

$$\min_{\pi\in\Pi_{\trianglerighteq}} [\pi(A) - \pi(B)] \geq \frac{1}{\#I} \min_{\pi\in\Pi_{\trianglerighteq}} \pi(A) \geq \frac{1}{\#I}\frac{1}{\#J}.$$

Conversely, suppose that $\min_{\pi\in\Pi}[\pi(A) - \pi(B)] \geq \frac{1}{2^n}$ for some $n \in \mathbf{N}$. By Equidivisibility, there exists partitions of A and B^c into 2^{n+1} equally likely events $\{A_i\}$ and $\{B_j\}$, respectively. Clearly, for any $\pi \in \Pi$ and any i,j, $\pi(A\backslash A_i) - \pi(B \cup B_j) \geq \pi(A) - \pi(B) - \frac{1}{2^n} \geq 0$, hence $A\backslash A_i \trianglerighteq B \cup B_j$ by coherence. □

Fact 5 *If Σ is a σ-algebra, Π is convex-ranged if and only if it is dyadically convex-ranged.*

Proof. The only-if part is immediate; to verify the if-part, take any non-null event $A \in \Sigma$, and $\alpha \in (0,1)$. By Lemma 13 below applied to the λ-system Λ_A, there exists an event $B \in \Lambda_A$ such that $\overline{\pi}(B|A) = \alpha$, verifying range-convexity. □

Lemma 13 *If Σ is a σ-algebra and $\overline{\pi}$ on Λ is dyadically convex-ranged, then Λ contains an algebra $\mathcal{A}$ on which $\overline{\pi}$ is convex-ranged.*

Proof. By dyadic range-convexity, there exists a nested sequence of algebras $\{\mathcal{A}_k\}$ such that $\mathcal{A}_k \subseteq \mathcal{A}_{k'}$ whenever $k \leq k'$ and such that $\overline{\pi}(A) = \frac{1}{2^k}$ for each atom of $\mathcal{A}_k$.

For any $A \in \Sigma$, let $A_{[k]}$ denote the largest subset of A that is an element of $\mathcal{A}_k$, and write $A^c_{[k]}$for $(A^c)_{[k]}$. Let $\mathcal{A}$ denote the set of all events $A \in \Sigma$ such that

$$sup_k \overline{\pi}(A_{[k]}) + sup_k \overline{\pi}(A^c_{[k]}) = 1. \tag{14}$$

We need to show $\mathcal{A}$ is an algebra contained in Λ on which $\overline{\pi}$ is convex-ranged.

1. For any $A \in \mathcal{A}$, $A \in \Lambda$ with $\overline{\pi}(A) = sup_k \overline{\pi}(A_{[k]})$.

By definition, for any $\pi \in \Pi$, $\pi(A_{[k]}) \leq \pi(A) = 1 - \pi(A^c) \leq 1 - \pi(A^c_{[k]})$. Taking sup's and account of (14), it follows that $\pi(A) = sup_k \overline{\pi}(A_{[k]})$, as desired.

2. $\mathcal{A}$ is an algebra

Closure under complementation is immediate. To verify closure under intersection, consider $A, B \in \mathcal{A}$.

Clearly $(A \cap B)_{[k]} = A_{[k]} \cap B_{[k]}$ and $(A \cap B)^c_{[k]} = (A^c \cup B^c)_{[k]} \supseteq A^c_{[k]} \cup B^c_{[k]}$.

Therefore in particular $\left((A \cap B)_{[k]} \cup (A \cap B)^c_{[k]}\right)^c \subseteq \left((A_{[k]} \cap B_{[k]}) \cup \left(A^c_{[k]} \cup B^c_{[k]}\right)\right)^c \subseteq \left(A_{[k]} \cup A^c_{[k]}\right)^c \cup \left(B_{[k]} \cup B^c_{[k]}\right)^c$.

By assumption, $\lim_{k\to\infty} \overline{\pi}\left(A_{[k]} \cup A^c_{[k]}\right)^c = 0$ and $\lim_{k\to\infty} \overline{\pi}\left(B_{[k]} \cup B^c_{[k]}\right)^c = 0$. Therefore also $\lim_{k\to\infty} \overline{\pi}\left((A \cap B)_{[k]} \cup (A \cap B)^c_{[k]}\right)^c$, as needs to be shown.

3. $\overline{\pi}$ is convex-ranged on $\mathcal{A}$.

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

$$\frac{\ell_j + 1}{2^j} \geq \alpha. \tag{15}$$

For any $k > 1$, let $A'_{[k]} = A_{[k]} \backslash A_{[k-1]}$, and let $A'_{[1]} = A_{[1]}$. Note that since the $A_{[k]}$ are nested, $A_{[k]} = \sum_{j \leq k} A'_{[j]}$; moreover, $A'_{[k]}$ is either empty or an atom of $\mathcal{A}_k$.

For each $k \geq 1$, and each $j \geq 1$, split $A'_{[k]}$ (if non-empty) into 2^j equally likely atoms of $\mathcal{A}_{k+j}$, and let B_{jk} be a union of ℓ_j such atoms, and C_{jk} a disjoint union of $2^j - \ell_j - 1$ such atoms. Clearly, for given k, the B_{jk} and C_{jk} and be chosen to be increasing in k.

Let $B_j = \sum_{k \leq j} B_{jk}$, $B = \cup_{j=1,..,\infty} B_j$, and likewise $C_j = \sum_{k \leq j} C_{jk}$, $C = \cup_{j=1,..,\infty} C_j$. Note that the sequences $\{B_j\}$ and $\{C_j\}$ are increasing in j. Now

$$\overline{\pi}(B_j) = \sum_{k \leq j} \overline{\pi}(B_{jk}) = \sum_{k \leq j} \alpha_j \overline{\pi}\left(A'_{[k]}\right) = \alpha_j \overline{\pi}\left(A_{[j]}\right).$$

Therefore, using step 1,

$$\sup_{j \to \infty} \overline{\pi}(B_j) = \alpha \overline{\pi}(A).$$

Since for any j, $B_j \in \mathcal{A}_{2j}$, $B_{[2j]} \supseteq B_j$, and therefore

$$\sup_{j \to \infty} \overline{\pi}\left(B_{[j]}\right) \geq \sup_{j \to \infty} \overline{\pi}(B_j) = \alpha \overline{\pi}(A). \tag{16}$$

By analogous reasoning, $\overline{\pi}(C_j) = \left(1 - \alpha_j - \frac{1}{2^j}\right) \overline{\pi}\left(A_{[j]}\right)$ and therefore $\sup_{j \to \infty} \overline{\pi}(C_j) = (1 - \alpha) \overline{\pi}(A)$. Moreover,

$$B^c_{[2j]} \supseteq C_j + A^c_{[2j]}.$$

Hence

$$\sup_{j \to \infty} \overline{\pi}\left(B^c_{[j]}\right) \geq \sup_{j \to \infty} \overline{\pi}(C_j) + \sup_{j \to \infty} \overline{\pi}\left(A^c_{[j]}\right) = (1 - \alpha) \overline{\pi}(A) + (1 - \overline{\pi}(A)) = 1 - \alpha \overline{\pi}(A). \tag{17}$$

Combining (16) and (17), it follows that $B \in \mathcal{A}$ and $\overline{\pi}(B) = \alpha \overline{\pi}(A)$, demonstrating range-convexity. □

Proof of Proposition 1.

If Λ is a σ-algebra, or if more generally Λ is an algebra with $\overline{\pi}$ convex-ranged, then LC restricted to betting preferences implies that the revealed likelihood relation $\succsim_\ell$ agrees with $\trianglerighteq$ on Λ, and LC for multi-valued acts entails Machina-Schmeidler's Strong Comparative Probability axiom. Thus the proof of Machina-Schmeidler's (1992) Theorem 1, step 5, and Theorem 2, step 2, can be used verbatim to obtain the desired conclusion.

This can be generalized to the general case in which Λ may fail to be an algebra as follows. Take any $f, g \in \mathcal{F}^{ua}$ such that $\overline{\pi} \circ f^{-1}$ stochastically dominates $\overline{\pi} \circ g^{-1}$ (weakly or strictly). Let

$\mathcal{B}_f$ (respectively $\mathcal{B}_g$ or $\mathcal{B}_{f,g}$) denote the smallest algebra containing all sets of the form $f^{-1}(x)$ (respectively $g^{-1}(x)$ or both $f^{-1}(x)$ and $g^{-1}(x)$), and let $\mathcal{B}_f^0, \mathcal{B}_g^0$ and $\mathcal{B}_{f,g}^0$ denote the families of their respective atoms. Clearly, all these are finite due to the assumed finite-rangedness of f and g.

For each $B \in \mathcal{B}_{f,g}^0$, Lemma 13 delivers the existence of an algebra $\mathcal{A}_B$ contained in Λ_B such that $\overline{\pi}(.|B)$ is convex-ranged on $\mathcal{A}_B$. Let $\mathcal{A}$ denote the algebra generated by their union, i.e. the family of all sets of the form $\sum_{B \in \mathcal{B}_{f,g}^0} A_B$, where $A_B \in \mathcal{A}_B$. Let $\mathcal{A}^{\perp}$ the subalgebra of events $A \in \mathcal{A}$ defined by the additional condition that $\overline{\pi}(A|B) = \overline{\pi}(A|B')$ for all $B, B' \in \mathcal{B}_{f,g}^0$; similarly, let $\mathcal{A}_f^{\perp}$ and $\mathcal{A}_g^{\perp}$ subalgebras of events $A \in \mathcal{A}$ defined by the weaker condition that $\overline{\pi}(A|B) = \overline{\pi}(A|B')$ for all those $B, B' \in \mathcal{B}_{f,g}^0$ that are contained in the same atom of $\mathcal{B}_f^0$ (respectively $\mathcal{B}_g^0$). By construction clearly $\mathcal{A}_f^{\perp} \supseteq \mathcal{B}_f \cap \mathcal{A}^{\perp}$ and $\mathcal{A}_g^{\perp} \supseteq \mathcal{B}_g \cap \mathcal{A}^{\perp}$.

Moreover, since $\mathcal{B}_f \cup \mathcal{B}_g \cup \mathcal{A}^{\perp} \subseteq \Lambda$, elementary reasoning shows that both $\mathcal{A}_f^{\perp}$ and $\mathcal{A}_g^{\perp}$ are contained in Λ, and that $\overline{\pi}$ is convex-ranged on both of these and on $\mathcal{A}^{\perp}$. By the latter, there exists an $\mathcal{A}^{\perp}$-measurable act h such that $\overline{\pi} \circ h^{-1} = \overline{\pi} \circ g^{-1}$, and such that by implication $\overline{\pi} \circ f^{-1}$ stochastically dominates $\overline{\pi} \circ h^{-1}$. By the Machina-Schmeidler argument for algebras (the first part of the proof), $h \sim g$ and $f \succsim h$ (respectively $f \succ h$ if the stochastic dominance is strict). Hence by transitivity $f \succsim g$ respectively $f \succ g$. $\square$

Proof of Theorem 3.

1) equivalent to 2). Immediate from the definition of Ψ^*.

1) implies 3).

Again, immediate from the definitions.

3) implies 4).

Let $\widetilde{\Pi} := \{\pi \in \Delta(\Sigma) : \pi(A) = \overline{\pi}(A) \text{ for all } A \in \Lambda\}$, and likewise $\widetilde{\Psi} := \{\pi \in \Delta(\Sigma) : \pi(A) = \overline{\psi}(A)$ for all $A \in \Lambda^*\}$, where $\overline{\psi}$ denote the unambiguous probability measure on Λ^* induced by Ψ^*.

The assertion follows from the following two Lemmas.

Lemma 14 *$\succsim$ is utility-sophisticated with respect to $\widetilde{\Pi}$.*

By definition of Ψ^*, $\overline{\psi}(A) = I(1_A)$ for all $A \in \Lambda^*$**,** hence by the assumption that $\Lambda^* \supseteq \Lambda$,

$$\overline{\psi}(A) = I(1_A) \text{ for all } A \in \Lambda. \tag{18}$$

By compatibility of $\succsim$ with $\trianglerighteq$, $I(1_A)$ is a monotone transform of $\overline{\pi}(A)$ on Λ**.** Since $\overline{\pi}$ is convex-ranged and both $\overline{\pi}$ and $\overline{\psi}$ are additive on Λ, by (18) in fact

$$\overline{\pi}(A) = \overline{\psi}(A) \text{ for all } A \in \Lambda.$$

Combining this with the assumption that $\Lambda^* \supseteq \Lambda$, we can infer that $\widetilde{\Pi} \supseteq \widetilde{\Psi}$. Since $\widetilde{\Psi} \supseteq \Psi^*$ by definition of $\widetilde{\Psi}$, $\widetilde{\Pi} \supseteq \Psi^*$; by the definition of Ψ^*, $\succsim$ is utility-sophisticated relative to $\widetilde{\Pi}$ as claimed. $\square$

Lemma 15 $\widetilde{\Pi}$ *is convex-ranged.*

We will show that, for all non-null $A \in \Sigma$, $B \in \Lambda_A$ and all $\pi \in \widetilde{\Pi}$,

$$\pi(B|A) = \overline{\pi}(B|A). \tag{19}$$

In fact, we will demonstrate (19) directly in the special case of $\overline{\pi}(B|A) = \frac{1}{2}$. This implies immediately the validity of (19) in the case of dyadic-valued $\overline{\pi}(B|A)$. Validity for arbitrary $\overline{\pi}(B|A) \in [0,1]$ then follows from the range-convexity of $\overline{\pi}(.|A)$.

Take any non-null $A \in \Sigma$ and $B \in \Lambda_A$ such that $\overline{\pi}(B|A) = \frac{1}{2}$. If A^c is null, $A \in \Lambda$, and the claim follows from the definition of $\widetilde{\Pi}$. Assume thus that A^c is non-null. By the range-convexity of Π, there exists $C \in \Lambda_{A^c}$ contained in A^c such that $\overline{\pi}(C|A^c) = \frac{1}{2}$.

By construction, for any $\pi \in \widetilde{\Pi}$, $\pi(B+C) = \frac{1}{2} = \pi((A \backslash B) + C)$, hence by additivity of π also $\pi(B) = \pi(A \backslash B)$, i.e. $\pi(B|A) = \frac{1}{2}$. $\square$

<u>4) implies 2).</u>

This is a straightforward consequence from the following Corollary to Proposition 2 in Nehring (2007).[44]

Proposition 4 *Let $\succsim$ be an invariant biseparable preference ordering that is utility-sophisticated relative to the convex-ranged likelihood ordering $\trianglerighteq_0$. Let $\trianglerighteq$ be any coherent superrelation. Then $\succsim$ is utility-sophisticated with respect to $\trianglerighteq$ if it is compatible with respect to $\trianglerighteq$.*

Suppose that $\succsim$ is compatible with $\trianglerighteq$ and $\succsim$ is utility-sophisticated relative to $\trianglerighteq_0$. Then by Proposition 4, $\succsim$ is utility-sophisticated relative to $\trianglerighteq$ respectively Π. $\square\blacksquare$

[44] The "regularity" condition assumed in the statement of Propostion 2 in Nehring (2007) is implied by invariant biseparability.

Proof of Fact 3.

Suppose that $A \in \Lambda^{GM}$, i.e. that $\alpha\pi^-(A) + (1-\alpha)\pi^+(A) + \alpha\pi^-(A^c) + (1-\alpha)\pi^+(A^c) = 1$. Since $\pi^-(A^c) = 1 - \pi^+(A)$ and $\pi^+(A^c) = 1 - \pi^-(A)$, we obtain by elementary transformations,

$$\begin{aligned} & \alpha\pi^-(A) + (1-\alpha)\pi^+(A) + \alpha\pi^-(A^c) + (1-\alpha)\pi^+(A^c) - 1 \\ = \; & \alpha\pi^-(A) + (1-\alpha)\pi^+(A) + \alpha\left(1 - \pi^+(A)\right) + (1-\alpha)\left(1 - \pi^-(A)\right) - 1 \\ = \; & (1-2\alpha)\left(\pi^+(A) - \pi^-(A)\right). \end{aligned}$$

Evidently, with $\alpha \neq \frac{1}{2}$, the latter can be equal to zero only if $\pi^+(A) = \pi^-(A)$, i.e. if $A \in \Lambda_\Psi$.

Thus we have shown that $\Lambda^{GM} \subseteq \Lambda_\Psi$. Now $\succsim$ is utility-sophisticated with respect to Ψ by construction. Hence $\Psi^* \subseteq \Psi$, and thus $\Lambda_\Psi \subseteq \Lambda^*$. Since finally for any i.b. preference, $\Lambda^* \subseteq \Lambda^{GM}$, we conclude that $\Lambda^{GM} = \Lambda_\Psi = \Lambda^*$. $\square$

Proof of Proposition 3.

In view of Theorem 3, it remains to show that $\Psi_1 = \Psi_1^*$.

Since $\succsim$ is utility sophisticated with respect to $\trianglerighteq_{\text{rand}}$, $\Psi^* \subseteq \Pi_{\text{rand}}$, whence $\Psi_1 \times \eta$. Trivially $\succsim_{|\mathcal{F}_1}$ is utility sophisticated with respect to Ψ_1, hence $\Psi_1^* \subseteq \Psi_1$. Conversely, suppose that $\succsim_{|\mathcal{F}_1}$ is utility sophisticated relative to Ψ_1'. We need to show that $\Psi_1' \supseteq \Psi_1$. Consider any $f, g \in \mathcal{F}$ and any $f', g' \in \mathcal{F}_1$ such that f', g' are measurable with respect to some finite partition $\{S_i\} \subseteq \Sigma_1$ and such that $E_\pi(u \circ f|S_i) = E_\pi(u \circ f'|S_i)$ and $E_\pi(u \circ g|S_i) = E_\pi(u \circ g'|S_i)$ for all i and all $\pi \in \Pi_{\text{rand}}$. Given $f, g \in \mathcal{F}$, such $f', g' \in \mathcal{F}_1$ by the definition of Π_{rand}. By utility sophistication with respect to Π_{rand},

$$f \sim f' \text{ and } g \sim g'.$$

Suppose that $E_\pi u \circ f \geq E_\pi u \circ g$ for all $\pi \in \Psi_1' \times \eta$. Since $\Psi_1' \times \eta \subseteq \Pi_{\text{rand}}$, for all $\pi \in \Psi_1' \times \eta$

$$E_\pi u \circ f' = E_\pi u \circ f \geq E_\pi u \circ g = E_\pi u \circ g'.$$

By the assumed utility sophistication of $\succsim_{|\mathcal{F}_1}$ with respect to Ψ_1', $f' \succsim g'$ and thus

$$f \sim f' \succsim g' \sim g,$$

verifying utility sophistication of $\succsim$ with respect to $\Psi_1' \times \eta$. By the minimality property characterizing Ψ^*, $\Psi_1' \times \eta \supseteq \Psi^* = \Psi_1 \times \eta$ and thus $\Psi_1' \supseteq \Psi_1$. $\square$

REFERENCES

[1] Ahn, D. (2005): "Ambiguity without a state space," mimeo, UC Berkeley.

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

[3] 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.

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

[5] Bewley, T. F. (2002, first version 1986): "Knightian Decision Theory, Part I," *Decisions in Economics and Finance* 25, 79–110.

[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] Eichberger, J., S. Grant and D. Kelsey (2006): "Differentiating Ambiguity – A Comment," mimeo.

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

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

[13] Ergin, H. and F. Gul (2004): "A Subjective Theory of Compound Lotteries", mimeo.

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

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

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

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

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

[19] Gajdos, T., J.-M. Tallon, and J.-C. Vergnaud (2004): "Decision making with imprecise probabilistic information," Journal of Mathematical Economics, 40(6), 647-681.

[20] Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2006): "Attitude toward imprecise information", mimeo.

[21] Ghirardato, P., F. Maccheroni, M. Marinacci and M. Siniscalchi (2003): "A Subjective Spin on Roulette Wheels", *Econometrica* 71, 1897-1908.

[22] Ghirardato, P., Maccheroni, F., and M. Marinacci (2004), "Differentiating Ambiguity. and Ambiguity Attitude", Journal of Economic Theory 118, 133-173.2005.

[23] Ghirardato, P., Maccheroni, F., and M. Marinacci (2005), "Revealed Ambiguity and Its Consequences I: Unambiguous Acts and Events," mimeo.

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

[25] Gilboa, I. , Samet, D. and D. Schmeidler (2004), "Utilitarian Aggregation of Beliefs and Tastes," *Journal of Political Economy*, 112, 932-938.

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

[27] Grant, S. and E. Karni (2004), "A Theory of Quantifiable Beliefs," *Journal of Mathematical Economics* 40, 515-546.

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

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

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

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

[32] Klibanoff, P., M. Marinacci and S. Mukerjii (2005): "A Smooth Model of Decision Making under Ambiguity," *Econometrica* 73, 1849–1892.

[33] Koopman, B. O. (1940a): "The Bases of Probability," *Bulletin of the American Mathematical Society* 46, 763-774.

[34] Koopman, B. O. (1940b): "The Axioms and Algebra of Intuitive Probability," *Annals of Mathematics* 41, 269-292.

[35] Kopylov, I. (2003): "Subjective Probability on 'Small Domains"', forthcoming, *Journal of Economic Theory.*

[36] Kopylov, I. (2006): "A parametric model of ambiguity hedging," mimeo, UC Irvine.

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

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

[39] Lyapunov, A. A. (1940): "Sur les fonctions vecteurs completement additives," *Izvestija Akademija Nauk SSR. Seria Matematiceskaja* 4, 465-478.

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

[41] Machina, M. (2004): "Almost Objective Uncertainty", Economic Theory 24, 1–54.

[42] Marinacci, M. (2002): "Probabilistic Sophistication and Multiple Priors," *Econometrica* 70, 755-764.

[43] Mongin, P. (1995). "Consistent Bayesian Aggregation," *Journal of Economic Theory* 66, 313–351.

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

[45] Nehring, K. (1992): "Foundations for the Theory of Rational Choice with Vague Priors", in: J. Geweke (Ed.): Decision Making under Uncertainty - New Models and Empirical Findings. Kluwer, Dordrecht 1992, pp. 231-242.

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

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

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

[49] Nehring, K. (2001): "Ambiguity in the Context of Probabilistic Beliefs", mimeo.

[50] Nehring, K. (2003): "Imprecise Probabilistic Beliefs", mimeo (earlier WP version).

[51] Nehring, K. (2007): "Bernoulli without Bayes: A Theory of Utility-Sophisticated Preference", mimeo, UC Davis.

[52] Nehring, K. (2007b): "Utility Sophistication in the Anscombe-Aumann Framework", mimeo, UC Davis.

[53] Nehring, K. (2007c): "Imprecise Probabilistic Beliefs as a Context for Decision-Making under Ambiguity", earlier working paper version, UC Davis.

[54] Olszewski, W. (2007): "Preferences Over Sets of Lotteries," *Review of Economic Studies* 74, 567–595.

[55] 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.

[56] 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.

[57] Regoli, G. (1998): "Comparative Probability Orderings," Documentation section, Society for Imprecise Probability Theory and Applications (SIPTA), URL http://ippserv.rug.ac.be

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

[59] Sarin, R. and P. Wakker (1997): "A Single-Stage Approach to Anscombe and Aumann's Expected Utility", *Review of Economic Studies* 64, 399-409.

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

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

[62] Schervish, M., T. Seidenfeld and J. Kadane (1990): "State-Dependent Utilities," *Journal of the American Statistical Association* 85, 840-847.

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

[64] Segal, U. (1990): "Two stage lotteries without the reduction axiom," *Econometrica* 58, 349-378.

[65] Seo, K. (2006): "Ambiguity and Second Order Beliefs.", mimeo, University of Rochester.

[66] Siniscalchi, M. (2006): "A Behavioral Characterization of Plausible Priors", *Journal of Economic Theory* 128, 91-135.

[67] Smith, C. (1961): "Consistency in Statistical Inference and Decision," *Journal of the Royal Statistical Society*, Series B, 22, 1-15.

[68] Stinchcombe, M. (2003): "Choice and games with ambiguity as sets of probabilities," mimeo, University of Texas, Austin.

[69] Tapking, J. (2004): "Axioms for preferences revealing subjective uncertainty and uncertainty aversion," *Journal of Mathematical Economics* 40, 771-797.

[70] Tversky, A. and P. Wakker (1995), "Risk. Attitudes and Decision Weights," Econometrica 63, 1255-1280.

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

[72] Wang, T. (2003): "A class of multi-prior preferences," Discussion paper, University British Columbia.

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