Decision-Making in the Context of Imprecise Probabilistic Beliefs

Klaus Nehring

University of California, Davis[1]

First Version: November 2002[2]

This Version: April 2006

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 axiom relating preferences and probabilistic beliefs which implies that preferences over unambiguous acts are probabilistically sophisticated and which entails representability of preferences over Savage acts in an Anscombe-Aumann-style framework. The motivation for an explicit and separate axiomatization of beliefs for the study of decision-making under ambiguity is discussed in some detail.

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 understood satisfactorily. 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. First, one might try to *define* beliefs from preferences following Savage (1954). While Savage's own definition can be invoked even under ambiguity at a purely formal level, 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. What is the structure of such beliefs, and how do they (rationally) constrain his preferences?"[2]

To address these two questions, we propose to model probabilistic beliefs as comparative likelihood relation $\unrhd$ over events, with "$A \unrhd B$" denoting the judgement "A is at least as likely as B", thereby

---

[1]For different reasons, a canonical definition of "revealed subjective probability" from choice-behavior fails to be possible in the case of state-dependent 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.

[2]Note that an answer to the first question naturally entails an answer to the second: to answer the latter, one simply needs to check whether the beliefs revealed by his preferences are consistent with the stated set of probabilistic beliefs. By contrast, as we shall see shortly, an answer to the second question does not necessarily entail an answer to the first. Thus, the second question is more modest and, arguably, more basic.

1

following the lead of Keynes (1921), de Finetti (1931) and Savage (1954). This likelihood relation shall be taken as an independent, non-behavioral primitive, *leaving open* the question whether/under what circumstances it can in turn be derived from preferences. We will thus describe a decision maker by two relations, a preference relation over acts together with a likelihood relation over events. The likelihood relation can represent either "objective" probabilistic information or purely subjective beliefs; these interpretations are fleshed out at the beginning of section 2. Under either interpretation, we will assume the likelihood relation to be incomplete in order to make room for preferences reflecting ambiguity, while preferences are assumed to be complete as usual.

Imprecise Probabilistic Beliefs in the Ellsberg Paradox

To illustrate the logic of the proposed preferences-plus-beliefs framework, let us consider the classical two-color, two urn 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 and B (the ball drawn from the known urn is red / black), as well as R' and B' (the ball drawn from the unknown urn is red / black). There is one fundamental piece of probabilistic information, namely that the events R and B are equally likely ($R \equiv B$). 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:

$$R \sim B \succ R' \sim B'$$

in obvious notation[3].

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 Consequentialism".

In the above example, preferences are evidently compatible with the specified information that $R \equiv B$. One may wonder, however, whether its is possible to attribute to the decision maker *in*

---

[3]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'.

2

addition a belief that red and black from the unknown urn are equally likely, $R' \equiv B'$, as implied by Savage's definition of revealed likelihood. Clearly, this can be done only at the price of sacrificing fundamental *coherence* properties of 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 judgement would be incompatible with the observed preference for betting on the known urn. A similar argument shows that the specified preferences are not compatible with attributing a belief that R' is strictly more likely than B', nor with the converse belief that B' is strictly more likely than R'. *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 a very 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 (1963) and Schmeidler (1989).

This beliefs-based approach does not represent the only possible explanation of the Ellsberg paradox. 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.[4] This alternative, preference-based account allows to maintain completeness of the likelihood relation at the price of sacrificing Compatibility/Likelihood Consequentialism. This is high price to pay as it severs radically the connection between belief and preference, whereas here 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 key task of the present paper is therefore the characterization of "coherent" likelihood relations, that is, of likelihood relations that incorporate all such entailments. For the limiting case of complete relations, Savage (1954) achieved a characterization of this kind leading to a representation by a

---

[4]Segal (1987), Ergin-Gul (2004) and Chew-Sagi (2003) can be interpreted in this vein, as well as perhaps Tversky-Wakker's (1995) notion of "source preference".

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

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 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 thus assumes a minimal degree of completeness of the likelihood relation. It is satisfied, for example, in the presence of a continuous random device, as assumed in the widely-used Anscombe-Aumann framework. In an important sense, Equidivisibility is 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 try to determine how specified imprecise probabilistic beliefs rationally constrain preferences. To do so, we propose an axiom called "Likelihood Consequentialism" that extends the compatibility requirement that has been formulated above for betting preferences in a natural way to general acts with multiple outcomes. It represents a *minimal* yet generally applicable criterion of consequentialist rationality relating preferences to probabilistic beliefs. It is minimal in that it does not constrain the DM's risk or ambiguity attitudes in any substantive way, thereby ensuring behavioral generality. In particular, it accommodates Allais- and Ellsberg-style choice patterns, and is not tied to assumptions about functional form. 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. We view the existence of substantive yet behaviorally general rationality restrictions on preferences as a crucial advantage of likelihood relations as an epistemic primitive (in contrast to, for example, a direct use of sets of priors).

While minimal, the restrictions on preferences entailed by Likelihood Consequentialism are substantial. In particular, we show that, given a likelihood relation satisfying the assumptions of Theorem 2, Likelihood Consequentialism entails probabilistic sophistication over unambiguous (risky) acts in the sense of Machina-Schmeidler (1992), that is: acts whose outcomes have well-defined probabilities derived from the likelihood relation. Taking the argument further, 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. Our derivation not only clarifies the assumptions on preferences and beliefs implicit in the Anscombe-Aumann model, it leads to an even more powerful structure since all uncertainty is treated at the same level. Moreover, since it applies to any likelihood relation with a convex-ranged representation, our derivation does not presuppose the existence of a continuous random device.

Likelihood relations represented by convex-ranged sets of priors promise to provide a fruitful setting for further decision theoretic analyses. Indeed, in companion papers (Nehring 2001, 2004), we use this framework to address three basic issues in the theory of decision making under ambiguity:

1. how to infer beliefs from preferences;

2. how to characterize decision-makers that depart from subjective expected utility exclusively for reasons of ambiguity; and

3. how to define ambiguity attitudes in terms of betting preferences only to ensure behavioral generality.

In each case, the additional structure provided by a convex-ranged sets of priors is crucial.

Related Literature

1. Our first main result, Theorem 2, builds on and can be viewed as the likelihood counterpart of the multiple-prior representations of partial orderings due to Bewley (1986) and Walley (1991) following Smith (1961). All of these, however, use preferences as their primitive[5] 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 in general much less structure than the objects in these contributions (orderings over random variables), which allow the use of vector-space techniques such as separation theorems. This difference probably explains why there do not exist counterpart results for likelihood relations in the literature up to now in spite of the suggestive parallelism. The key technical insight of the present paper is precisely 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 applicable; the construction of the extension itself is non-trivial.[6]

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

---

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

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

6

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 Regioli (1999) for surveys. In the incomplete case, one can use standard arguments from the theory of linear inequalities to obtain a characterization of coherence for likelihood relations defined on arbitrary families of sets; see Walley (1991 p. 192-3) and related earlier results by Heath-Suddert (1972) and, in the complete, finite-state case, Kraft-et-al (1959). In view of the combinatorial complexity[7] and algebraic character of the conditions, such characterizations have generally not been considered to be of significant foundational interest (c.f. e.g. Regioli 1999).

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 existence of the multi-prior representation results dating back to Smith (1961), the extant results in the literature on likelihood relations do not seem to come close to those of the present paper.

4. Some of the recent literature on decision making under ambiguity can be read as offering proposals for characterizing a decision maker's unconditional probabilistic beliefs directly through definitions of "unambiguous events" revealed by the preference relation; see Epstein-Zhang (2001), Ghirardato-Marinacci (2002) and Nehring (1999)). Of these, only Epstein-Zhang (2001) strives to be behaviorally general. Any such definition can be used to define a compatibility requirements between preferences and explicitly given unconditional probabilistic beliefs. In Appendix 1, we point out some limitations of the Epstein-Zhang definition from this perspective.[8] Thus, even for the special yet fundamental case of unconditional probabilistic beliefs, no generally applicable criterion of preference compatibility with such beliefs is available in the literature.

---

[7]In particular, the rationality axioms involve finite families of events whose cardinality is only bounded by the cardinality of the state space; moreover these families need not have any simple set-theoretic structure.

[8]Epstein-Zhang's (2001) primary goal appears to be a behavioral separation of risk and ambiguity, rather than a behavioral definition of probabilistic unconditional beliefs.

Overview

Section 2 characterizes likelihood relations with a convex-ranged multi-prior representation (Theorem 2), illustrates the central structural assumption of Equidivisibility/convex-rangedness with a number of examples, shows that uniqueness may easily be lost in its absence, and distinguishes "coherent" (deductively closed) from "consistent" (non-contradictory) likelihood relations.

In Section 3, we formulate a behaviorally general rationality requirement relating preferences to beliefs, "Likelihood Consequentialism". It entails probabilistic sophistication over unambiguous acts (Proposition 1) and leads to a representation of preferences in an Anscombe-Aumann framework (Proposition 2). Even in simple cases such as preferences with a multi-prior representation, the implications of Likelihood Consequentialism are not straightforward (Proposition **??**). While bounded rationality considerations may motivate a weakening of coherence to consistency, this may lead to major losses in the bite of Likelihood Consequentialism (Example 5).

Section 4 mentions possible rationality conditions that go beyond Likelihood Consequentialism, and discusses the applicability of this principle if preferences are state-dependent. We conclude by pointing out how the analysis of this paper can be utilized from the standard strict behaviorist position that does not admit beliefs as independent primitives.

The first part of the Appendix discusses the relationship of the present work to the work on unambiguous events by Zhang (1999) and Epstein-Zhang (2001); the second part of the Appendix contains all proofs.

2. COHERENT LIKELIHOOD RELATIONS

A decision maker's probabilistic beliefs shall be modelled in terms of a partial ordering $\unrhd$ on an algebra of events Σ in a state space Ω, his "comparative likelihood relation", with the instance $A \unrhd B$ denoting the DM's judgment that A is at least as likely as B. We shall denote the symmetric component of $\unrhd$ ("is as likely as") by $\equiv$, and the asymmetric component by $\rhd$. The comparative likelihood relation can be viewed as representing a not necessarily exhaustive set of probabilistic judgments attributed to the DM, his *explicit probabilistic beliefs.*

2.1   The Likelihood Relation as a Primitive

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 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, 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). 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, 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. The beliefs need not be specified exhaustively. That is, the decision maker may "have" further beliefs that have not yet elicited and recorded in $\unrhd$, 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. To highlight the generally non-exhaustive character of the specified likelihood relation, we sometimes refer to it as describing the decision maker's *explicit* 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) $\Omega \triangleright \emptyset$.

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). Savage (1954) realized, however, that Additivity suffices for the characterization of convex-ranged probability measures;[9] 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 $A \triangleright \emptyset$.

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 $B \triangleright A$ there exist non-null events C and D such that* $B \backslash D \triangleright A \cup C$.

---

[9]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.

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 in explicit beliefs, likelihood relations will now allowed to be incomplete.

Axiom 7 (Partial Order)[10]   *$\trianglerighteq$ is transitive and reflexive.*

A main achievement of Savage's Probability Theorem is its reliance on Additivity as the sole axiom capturing the logical syntax of probability. If the completeness assumption is dropped, this seems no longer feasible. For example, while under completeness, one can use Additivity to infer that if 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.*

---

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

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.[11]

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 likelihood 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 set of finitely additive probability measures $\Pi \subseteq \Delta(\Sigma)$ such that, for all $A, B \in \Sigma$:

$$A \trianglerighteq B \text{ if and only if } \pi(A) \geq \pi(B) \text{ for all } \pi \in \Pi. \tag{1}$$

---

[11]Equidivisibility is non-trivially weaker than the similar older assumptions of De Finetti (1931) and Koopman (1940a), according to which any event can be split in arbitrarily many (e.g. 17) equally likely parts.

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

Note that if $\trianglerighteq$ satisfies (1) for some set of priors Π, then it satisfies (1) 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 "convex-rangedness" condition on Π; if Σ is merely an algebra, it is equivalent to "dyadic convex-rangedness".[13] Let $\mathbf{D}$ denote the set of dyadic numbers between 0 and 1, i.e. of numbers of the form $\alpha = \frac{\ell}{2^k}$, where k and ℓ are non-negative integers such that ℓ does not exceed 2^k.

Definition 1 *A set of priors Π is* **convex-ranged** *if, for any event $A \in \Sigma$ and any $\alpha \in (0,1)$, there exists an event $B \in \Sigma$, $B \subseteq A$ such that $\pi(B) = \alpha\pi(A)$ for all $\pi \in \Pi$. The set Π is* **dyadically convex-ranged** *if this holds for all $\alpha \in \mathbf{D}$.*

Note that while convex-rangedness of Π implies the convex-rangedness of every $\pi \in \Pi$, the converse is far from true. Moreover, as established by Fact 5 in the Appendix[14], on σ-algebras dyadic convex-rangedness and convex-rangedness 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. The representing Π is unique in $\mathcal{K}(\Delta(\Sigma))$.*

---

[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).

[14]Fact 5 is presented as a corollary of Lemma 13.

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 convex-rangedness of the set of priors. Specifically, one can extend every coherent likelihood relation represented by the convex-ranged set of priors Π to a partial ordering on the domain $\mathcal{Z}$ of finite-valued functions $Z : \Omega \to [0,1]$ by associating with each function Z an equivalence class $[Z]$ of events $A \in \Sigma$ as follows. Let $A \in [Z]$ if, for some appropriate partition of Ω $\{E_i\}$, $Z = \sum z_i 1_{E_i}$, and such that, for all $i \in I$ and $\pi \in \Pi : \pi(A \cap E_i) = z_i \pi(E_i)$. It is easily seen that for any two $A, B \in [Z] : \pi(A) = \pi(B)$ for all $\pi \in \Pi$, and thus $A \equiv B$. One therefore arrives at a well-defined partial ordering on $\mathcal{Z}$, denoted by $\widehat{\unrhd}$, by setting

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

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

(Additivity) $Y \widehat{\unrhd} Z$ if and only if $Y + X \widehat{\unrhd} Z + X$ for any X, Y, Z ,   (2)

and

(Homogeneity) $Y \widehat{\unrhd} Z$ if and only if $\alpha Y \widehat{\unrhd} \alpha Z$ for any Y, Z 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 3 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.[16]

---

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

[16]In principle, one could conceive of coherent expectation orderings as epistemic primitives. This, however, would run into the following two problems. On the one hand, the meaning of a comparison of random variables in terms of their expectation seems intuitively not clear; it seems doubtful that a *non-behavioral* epistemic primitive can be based on a complex, mathematically structured implicit expectation operation. It is thus not suprising that Walley (1991), for example, consistently adopts a behavioral interpretation of coherent expectation orderings (respectively their counterpart in his work, "lower previsions") in terms of acceptability of gambles.Moreover, unless one assumes expected-utility maximization, as both Walley and Bewley do, the link between expectation orderings and preferences

14

2.4 Uniqueness of the Multi-prior Representation and the Expressive Adequacy of Likelihood Relations

The uniqueness part of the multi-prior representation in Theorem 2 is non-trivial and significant, as it ensures the expressive adequacy of likelihood relations as an epistemic primitive. Indeed, it is not at all obvious *a priori* that likelihood relations are sufficiently expressive as primitive carriers of imprecise probabilistic beliefs. Indeed, in finite state spaces, likelihood relations seem evidently defective in this regard, and there is no indication in the existing literature that this situation can be remedied in a systematic fashion in infinite state spaces.

Analogously to the complete case in which uniqueness of the representing prior is a natural heuristic criterion of adequate expressiveness, uniqueness of the representing closed and convex set of priors is a natural yardstick of adequate expressiveness in the more general incomplete case. While it can be shown that Equidivisibility is not strictly necessary to achieve uniqueness, it does not seem possible to weaken this assumption greatly and still obtain uniqueness in a robust manner. In particular, non-atomicity-like conditions in the manner of Savage's Fineness and Tightness conditions are not nearly enough as shown by the following example.

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

$$A \unrhd^K B \text{ if and only if } \lambda(A\backslash B) \geq K\lambda(B\backslash A). \tag{3}$$

With $K > 1$, $\unrhd^K$ is clearly not equidivisible; in particular, $\unrhd^K$ does not admit any unambiguous event with probability different from 0 or 1. It is easily verified that the associated set of admissible priors $\Pi_{(\unrhd^K)}$ (which we shall also denote as Π_1^K) consists of all probability measures π with Lebesgue density ϕ such that $\operatorname{ess\,sup}_{\omega \in [0,1]} \phi(\omega) \leq K \operatorname{ess\,inf}_{\omega \in [0,1]} \phi(\omega)$;[17] in particular, the extreme points of Π_1^K consist of all probability measures π_D with density ϕ_D, where D ranges over Σ with $0 < \lambda(D) < 1$, and ϕ_D is given by

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

Let $\Pi_2^K \subseteq \Pi_1^K$ denote the closed, convex hull of $\{\pi_D | \lambda(D) = \frac{1}{K+1}\}$; the following Fact states that Π_2^K induces the same likelihood relation $\unrhd^K$. Yet, as described in the following Fact that is easily

is not clear. Expectation orderings are thus not viable as a behaviorally general vehicle for describing a decision maker's imprecise probabilistic beliefs.

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

verified, Π_1^K and Π_2^K induce different lower probability functions denoted by $\pi_{1,K}^-$ and $\pi_{2,K}^-$.

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

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

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

The lower probabilities $\pi_{1,K}^-(A)$ and $\pi_{2,K}^-(A)$ are shown in the following figure as functions of $\lambda(A)$ for $K = 3$, with $\pi_{1,K}^-$ above $\pi_{2,K}^-$ and touching at $\lambda = \frac{3}{4}$.
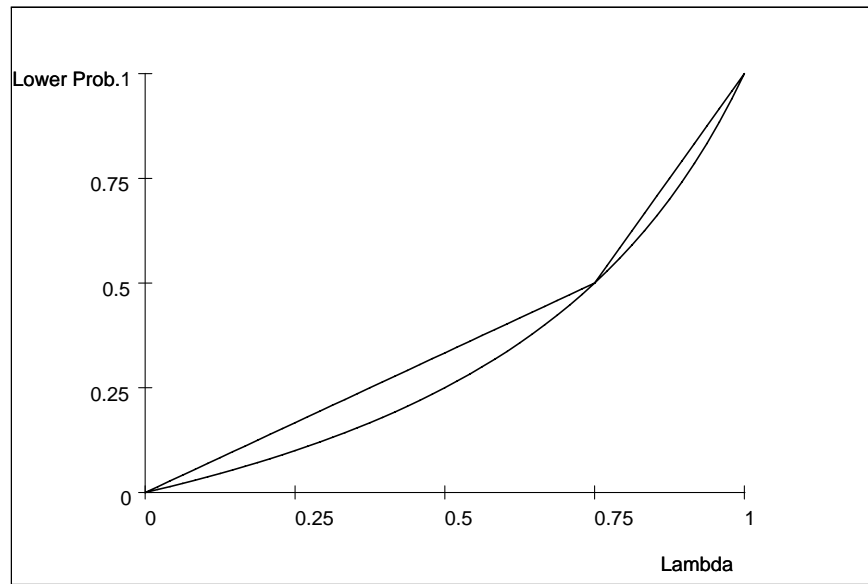

Fig. 1: Two Different Lower Probabilities

For $K > 1$, $\trianglerighteq^K$ clearly satisfies Savage's Fineness and Tightness conditions. Note that if K is close to 1, all admissible probabilities are uniformly close to the Lebesgue measure; nonetheless, uniqueness is lost.

2.5 Examples of Equidivisibility

The key structural assumption behind Theorem 2, Equidivisibility, is fairly strong. While it implies Fineness in the presence of Continuity, the converse is not close to being true as just illustrated, unless the likelihood relation is complete. Whereas Fineness is in substance a strong non-atomicity

condition, Equidivisibility assumes in addition that the likelihood relation is sufficiently complete. It is further illuminated by means of the following specific examples.

Example 1 (Limited Imprecision, Social Belief Aggregation). One way to make the intuition of a limited extent of overall ambiguity precise is to assume that Σ is a σ-algebra and that Π is the convex hull of a finite set Π' of non-atomic, countably additive priors. Due to Lyapunov's (1940) celebrated convexity theorem, Π is convex-ranged. The priors $\pi \in \Pi'$ can be interpreted as a finite set of hypotheses a decision-maker deems reasonable without being willing to assign probabilities to them.

Finitely generated sets of priors also occur naturally in social belief aggregation, where $\unrhd_I$ represents the unanimity likelihood ordering induced by the finite set of individuals' likelihood orderings $\unrhd_i$ that are assumed to be precise with representing measures μ_i. Assume that social decisions are based on a precise likelihood ordering $\unrhd_I$ represented by some measure μ_I that respects unanimity in beliefs. Then Theorem 2 implies that $\Pi_{(\unrhd_I)} = co\{\mu_i\}_{i \in I}$; the "social prior" μ_I must therefore be a convex combination of individual priors.[18]

Example 2 (Missing Information).

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

Example 3 (External Randomization Device)

As a variant of example 2, consider state-spaces with a continuous randomization device in the manner of Anscombe-Aumann. Specifically, consider a state space that can be written as $\Omega = \Omega_1 \times \Omega_2$, where the space Ω_1 is the space of "generic states" , and Ω_2 that of independent "random states" with associated algebras Σ_1 and σ-algebra Σ_2. The "continuity" and stochastic independence of the random device are captured by a coherent likelihood relation $\unrhd_{AA}$ defined on the product algebra $\Sigma = \Sigma_1 \times \Sigma_2$ that satisfies the following two conditions, noting that any $A \in \Sigma_1 \times \Sigma_2$ can be written as $A = \sum_i S_i \times T_i$, where the $\{S_i\}$ form a finite partition of Ω_1.

---

[18]This corollary to Theorem 2 is related to a result by 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 $\unrhd_{AA}$ to $\{\Omega_1\} \times \Sigma_2$ satisfies all of Savage's axioms (axioms 1 through 6).

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

While the first condition ensures the existence of a convex-ranged probability measure $\overline{\pi}_2$ over random events, the second describes their stochastic independence. By AA1 and AA2, it is easily verified that $\unrhd_{AA}$ satisfies all the assumptions of Theorem 2 including Equidivisibility. Hence there exists a unique set of priors Π_{AA} representing $\unrhd_{AA}$; indeed, Π_{AA} is simply the set of all product-measures of the form $\pi_1 \times \overline{\pi}_2$, where π_1 can be any finitely additive measure on Σ_1, reflecting the stochastic independence of the random device.

The relation $\unrhd_{AA}$ may play the role of a background context that allows the specification of beliefs about "generic" events (Anscombe-Aumann's horse-races) which may be the only events of direct interest. For example, the belief that an event A has unconditional probability α can be expressed as the likelihood judgement $A \equiv \Omega_1 \times T$, where T is any random event with $\overline{\pi}_2(T) = \alpha$. More generally and interestingly, for two generic events $A \subseteq E \in \Sigma_1$, the belief that A has probability α *given the* E can be expressed as the likelihood judgement $A \times \Omega_2 \equiv E \times T$, where T is again any random event with $\overline{\pi}_2(T) = \alpha$.[19]

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

2.6   Coherence and Consistency: Deductive Closure versus Non-Contradiction

By requiring deductive closure under inferences from the logic of probability, coherence is a strong notion of epistemic rationality . A weaker notion of "consistency" would merely require the absence of contradictions with respect to the logic of probability.[20] Equipped with a notion of coherence, one can define a likelihood relation as consistent if it contains some coherent superrelation. Given the

---

[19]Unconditional and conditional probabilities can be expressed as likelihood judgements in a similar manner whenever the likelihood relation has a convex-ranged representation. In fact, the possibility of doing so for arbitrary events is essentially equivalent to the definition of convex-rangedness, and underscores the need for Equidivisibility (or something very close to it) to ensure adequate expressiveness of the likelihood relation.

[20]The notions of coherence and consistency are analogues of the notions of "coherence" and "avoiding sure loss" for lower previsions defined in Walley (1991).

formal identification of coherence with the existence of a multi-prior representation, this amounts to a formal definition of $\unrhd$ as **consistent** if and only if there exists a finitely additive probability measure π such that, for all $A, B \in \Sigma$,

$$A \unrhd B \text{ implies } \pi(A) \geq \pi(B). \tag{4}$$

It is natural to think of the probabilistic information as presented to the decision maker in the form of a consistent likelihood relation $\unrhd$ (as sketched in the penultimate paragraph of Example 3 above), and its epistemic content being given by the body of all deductive inferences from it. This can formally be defined as the smallest coherent superrelation of $\unrhd$ denoted by $\unrhd^{coh}$. Letting $\Pi_{\unrhd}$ denote the set of all priors satisfying (4), it is easily verified that

$$A \unrhd^{coh} B \text{ if and only if } \pi(A) \geq \pi(B) \text{ for all } \pi \in \Pi_{\unrhd}.$$

If $\unrhd$ is equidivisible, in view of Theorem 2, $\unrhd^{coh}$ coincides with the smallest superrelation satisfying Transitivity, Positivity, Additivity, Splitting and Continuity $\unrhd^{t\&p\&a\&s\&c}$. Moreover, $\unrhd$ is consistent if and only if $\unrhd^{t\&p\&a\&s\&c}$ is non-degenerate. We thus have:

Corollary 1 *An equidivisible likelihood relation $\unrhd$ is consistent if and only if $\unrhd^{t\&p\&a\&s\&c}$ is non-degenerate; in this case $\unrhd^{t\&p\&a\&s\&c} = \unrhd^{coh}$.*

3. DECISION MAKING IN THE CONTEXT OF PROBABILISTIC BELIEFS

3.1. Likelihood Consequentialism

Consider now a decision maker described by a preference ordering over acts and explicit beliefs over events. Let X be a set of *consequences*. An *act* is a mapping from states to consequences, $f : \Omega \to X$ that is measurable with respect to an algebra of events Σ; the set of all acts is denoted by $\mathcal{F}$; for simplicity, we will assume all acts to be finite-valued throughout. A *preference ordering* $\succsim$ is a weak order (complete and transitive relation) on $\mathcal{F}$. We shall write $[x_1 \text{ on } A_1; x_2 \text{ on } A_2; ...]$ 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 consistent (typically: coherent) comparative likelihood relation $\unrhd$ on Σ. The relation $\unrhd$ will be referred to as the epistemic *context* of the decision situation. Thus, a *decision-maker in an epistemic context* is described by

19

the pair $(\succsim, \unrhd)$. A coherent context $\unrhd$ will be referred to as convex-ranged if it has a convex-ranged multi-prior representation on the event-algebra Σ.[21]

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 consequence incidences $f^{-1}(x)$ and $g^{-1}(x)$ of the various consequences of different acts; other conceivable factors such as familiarity with a type of event or felt competence in assessing it *should* not matter rationally. We shall refer to this as the Principle of 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 $x \succ y$:

$$[x \text{ on } A; y \text{ on } A^c] \succsim [x \text{ on } B; y \text{ on } B^c] \text{ whenever } A \unrhd B. \tag{5}$$

Note that condition (5) can be viewed as a unidirectional version of Savage's behavioral definition of revealed likelihood. Condition (5) 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 $\unrhd$ with its asymmetric component $\rhd$. However, if $\unrhd$ is incomplete, the resulting condition would however be overly restrictive, as illustrated by the following example.[22]

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

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

---

[21]Convex-rangedness 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" $\unrhd_{AA}$ in example 3 has a convex-ranged representation, even though the domain of $\unrhd_{AA}$, the product-algebra $\Sigma_1 \times \Sigma_2$, is not a σ-algebra if the generic state space Ω_1 is infinite.

[22]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 $[x$ on $E; y$ on $E^c] \sim [x$ on $\emptyset; y$ on $\Omega]$ even though $E \rhd \emptyset$, violating the envisaged asymmetric counterpart to condition (5).

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

Definition 2 (Uniformly More Likely) $A \rhd\rhd B$ *("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 $\in I$ and $j \in J$.*

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

Fact 3 *For any consistent likelihood relation $\trianglerighteq$, $A \rhd\rhd B$ implies $\min_{\pi \in \Pi_\trianglerighteq} [\pi(A) - \pi(B)] > 0$. The converse holds if $\trianglerighteq$ is coherent and equidivisible.*

Definition 2 leads to the following asymmetric counterpart of condition (5). For all $A, B \in \Sigma$ and $x, y \in X$ such that $x \succ y$:

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

The following axiom called "Likelihood Consequentialism" 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 Consequentialism) *For all $f \in \mathcal{F}$, $x, y \in X$ and events $A, B \in \Sigma$:*

i) $[x$ on $A\backslash B; y$ on $B\backslash A; f(\omega)$ elsewhere$]$ $\succsim$ $[x$ on $B\backslash A; y$ on $A\backslash B; f(\omega)$ elsewhere$]$ if $A \trianglerighteq B$ and $x \succsim y$, and

ii) $[x$ on $A\backslash B; y$ on $B\backslash A; f(\omega)$ elsewhere$]$ $\succ$ $[x$ on $B\backslash A; y$ on $A\backslash B; f(\omega)$ elsewhere$]$ if $A \rhd\rhd B$ and $x \succ y$.

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

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* $A \rhd\rhd \emptyset$.

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 $\pi \in \Pi$, 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 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 convex-rangedness of $\rhd$, 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* $\rhd$ *, 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, convex-rangedness entails "enough" intersection closedness to make use of their proof nonetheless.[23]

3.2   Application to Multi-Prior Preferences

Even in very simple cases, the constraints entailed by imprecise probabilistic beliefs are non-trivial. Consider, for example, the multi-prior model with preferences given by

$$f \succsim g \text{ iff } \min_{\pi \in \Psi} \sum_{x \in X} u(x) \pi(\{\omega : f(\omega) = x\}) \geq \min_{\pi \in \Psi} \sum_{x \in X} u(x) \pi(\{\omega : g(\omega) = x\}), \qquad (7)$$

for some utility function $u : X \to \mathbf{R}$ and some closed, convex set of probability measures Ψ; for axiomatizations of the multi-prior model in a Savage context which is pertinent here, see Casadesus et al. (2000) and Ghirardato et al. (2003). *When are multi-prior preferences with representation* (u, Ψ) *compatible with imprecise probabalistic beliefs represented by the set* Π? The answer to this question is not obvious, either on direct "intutitive 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, compatibility with Π (in the sense of this paper) does not imply that $\Psi \subseteq \Pi$. To see this, consider the singleton case $\Pi = \{\pi\}$, with π convex-ranged. According to Proposition 1, Compatibility with $\unrhd_{\{\pi\}}$ is equivalent to probabilistic sophistication with respect to π. On the other hand, for $\Psi \subseteq \Pi = \{\pi\}$, that is: for Ψ to be equal to $\{\pi\}$, preferences must maximize expected utility. But it is well-known[24] that probabilistic sophistication with respect to π does not entail expected utility maximization in the multi-prior model.

Quite generally, for the "natural characterization" $\Psi \subseteq \Pi$ to obtain with convex-ranged Π, the decision maker must be an expected utility maximizer with respect to unambiguous acts; more interestingly, this condition also turns out to be sufficient for this characterization.

Proposition 2 *Let* $\succsim$ *be a preference relation with multi-prior representation* (u, Ψ) *given by (7) and* $\unrhd$ *be a coherent and convex-ranged likelihood relation with representation* Π. *Then* $\Psi \subseteq \Pi$ *if*

---

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

[24]See in particular Marinacci (2002).

and only if $\succsim$ *is compatible with* $\trianglerighteq$ *and satisfies Savage's Sure-Thing Principle P2*[25] *on the set of unambiguous acts* $\mathcal{F}^{ua}$.

The characterization of Likelihood Consequentialism in the general multi-prior model without the assumption of expected utility maximization on unambiguous acts is an interesting open question.

3.3   Likelihood Consequentialism when the Context is Merely Consistent

The behavioral implications of probabilistic beliefs such as Proposition 1 depend crucially on their coherence. Since coherence is a demanding assumption on the decision-maker's rationality (or better, his "epistemic competence") in view of its built-in deductive closure requirement, from a bounded rationality perspective it is interesting to study the preference implications of probabilistic beliefs that are consistent rather than coherent. As the following example shows, this leads to major losses in the predictive and normative power of Likelihood Consequentialism.

Example 5.   Consider a state space with external randomization $(\Omega_1 \times \Omega_2, \Sigma_1 \times \Sigma_2)$ as in Example 3. Let $\trianglerighteq_{AA}$ the likelihood relation representing the information about the random device, with $\overline{\pi}_2$ denoting the probability measure over random events Σ_2. Also, let $\trianglerighteq_1$ denote a likelihood relation over $\Sigma_1 \times \{\Omega_2\}$ satisfying all the Savage axioms (axioms 1 through 6) with representing prior $\overline{\pi}_1$. It is easily seen that the coherent hull of the combined relation $\trianglerighteq_{AA} \cup \trianglerighteq_1$ denoted by $(\trianglerighteq_{AA} \cup \trianglerighteq_1)^{coh}$ is complete and represented by the product measure $\overline{\pi}_1 \times \overline{\pi}_2$. Hence, in view of Proposition 1, Likelihood Consequentialism applied to the relation $(\trianglerighteq_{AA} \cup \trianglerighteq_1)^{coh}$ boils down to probabilistic sophistication with respect to the measure $\overline{\pi}_1 \times \overline{\pi}_2$.

By contrast, Likelihood Consequentialism applied to the relation $\trianglerighteq_{AA} \cup \trianglerighteq_1$ entails far less. In particular, there exist preferences that are separately probabilistically sophisticated over the set of acts that are measurable with respect to $\Sigma_1 \times \{\Omega_2\}(=: \mathcal{F}_1)$ and $\{\Omega_1\} \times \Sigma_2 \ (=: \mathcal{F}_2)$, but that are *not* probabilistically sophisticated over $\mathcal{F}_1 \cup \mathcal{F}_2$.

Fact 4 *There exist preference relations that are compatible with the relation* $\trianglerighteq_{AA} \cup \trianglerighteq_1$ *but fail to be probabilistically sophisticated on* $\mathcal{F}_1 \cup \mathcal{F}_2$.

Fact 4 is verified in the Appendix. Indeed, the preferences constructed there exhibit the Ellsberg paradox on $\mathcal{F}_1 \cup \mathcal{F}_2$. More specifically, we show that the property asserted by Fact 4 holds for

---

[25]Savage's Sure-Thing Principle P2 is the following condition:

For all acts $f, f', g, g' \in \mathcal{F}$ and all $A \in \Sigma$, $[f$ on A, g on $A^c] \succsim [f'$ on A, g on $A^c]$ if and only if $[f$ on A, g' on $A^c] \succsim [f'$ on A, g' on $A^c]$.

any non-SEU "second-order probabilistically sophisticated CEU" preference ordering as defined in Ergin-Gul (2004). Preferences of this kind are paradigmatic examples of the preference- rather than belief-based account of the Ellsberg paradox that was mentioned in the introduction and will be discussed further below in section 4.3.

Example 5 shows that probabilistic beliefs lose much of their explanatory/predictive power if they are allowed to be consistent yet incoherent. The example also shows that in the absence of coherence there may not be a well-defined notion of risk. Indeed, it is not even obvious how to define the set of unambiguous events without coherence. In the present example, it seems natural to consider at least all events in the union of $\Sigma_1 \times \{\Omega_2\}$ and $\{\Omega_1\} \times \Sigma_2$ as unambiguous. But then preferences over unambiguous acts are not probabilistically sophisticated.

3.4   An Epistemic Interpretation of the Anscombe-Aumann Framework

Likelihood Consequentialism will now be combined with the mixture-space construction of section 2 to obtain an epistemic interpretation of the Anscombe-Aumann (1963) framework. While this framework is often used in the analysis of decision making under ambiguity, it is generally viewed as theoretically less fundamental and transparent than the Savage framework; sometimes it is even viewed with outright suspicion (see, e.g., Epstein (1999)).

The Anscombe-Aumann (1963) framework is distinguished by taking acts to be mappings from states to probability distributions of consequences, rather than simply as mappings from states to consequences as in the Savage (1954) framework. These probability distributions are interpreted as objective probabilities of the realizations of an external random device ("roulette lotteries") that is not part of the explicitly modeled state space. We will show that if a preference relation over Savage acts is compatible with a coherent convex-ranged epistemic context $\unrhd$, it can be canonically extended to a preference relation over Anscombe-Aumann (AA-) acts over the *same* state space; in particular, an external randomization device need not be added.

Formally, an AA-act F is a finite-valued Σ-measurable mapping from the state space Ω to the set of probability distributions on X with finite support $\mathcal{L}$. Let $\mathcal{F}^{AA}$ denote the set of such acts. Denoting elements of $\Delta(X)$ by $q = (q^x)_{x \in X}$, one can write $F = [q_1 \text{ on } A_1; q_2 \text{ on } A_2; ...]$ in analogy to the notation for Savage acts. Given a convex-ranged epistemic context $\unrhd$, any AA-act F can be identified with a class $[F]$ of Savage acts by the following stipulation: $f \in [F]$ if, for any i such that

A_i is non-null,

$$f_{|A_i} \text{ is } \Lambda_{A_i}\text{-measurable and } \overline{\pi}(./A_i) \circ f^{-1} = q_i.^{26}$$

Thus $[F]$ consists of all Savage acts that yield the consequence probabilities specified by F as unambiguous conditional probabilities with respect to the given context.

Preferences over Savage acts $\succsim$ induce a preference relation over AA acts $\succsim^{AA}$ on $\mathcal{F}^{AA}$ according

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

If preferences are compatible with the convex-ranged likelihood ordering $\unrhd$, this relation is a well-behaved, uniquely defined weak order order. This claim is entailed by the following generalization of Proposition 1. Say that $\succsim^{AA}$ is monotone (with respect to stochastic dominance) if it is monotone pointwise, i.e. if $F \succsim^{AA} G$ whenever $F(\omega)$ stochastically dominates $G(\omega)$ for all $\omega \in \Omega$, and if $F \succ^{AA} G$ whenever in addition $F(\omega)$ stochastically dominates $G(\omega)$ strictly on some non-null event A. Note that monotonicity implies that $f \sim^{AA} g$ for any $f, g \in [F]$, since such f, g mutually stochastically dominate each other.

Proposition 3 *If $\succsim$ is compatible with the coherent and convex-ranged likelihood relation $\unrhd$, then $\succsim^{AA}$ is a uniquely defined monotone weak order on $\mathcal{F}^{AA}$.*

A key feature of the above construction is its applicability to general convex-ranged belief contexts; the existence of an independent random device as implied by Anscombe-Aumann's (1963) "horse-lottery" interpretation is not assumed.[27] This added generality is of value for example in contexts of social aggregation, when $\unrhd$ is the unanimity relation of individuals' revealed likelihood relations as in Example 1 above. Proposition 3 shows that the AA structure comes for free in such settings.

In the special case in which the context represents the independent random device (i.e. if $\unrhd = \unrhd_{AA}$), all information about the underlying Savage preferences is summarized by preferences over the subset $\mathcal{F}_1^{AA}$ of acts in $\mathcal{F}^{AA}$ that depend on the realization of the "generic uncertainty" Σ_1 (the outcome of the horse race) only[28]. Restricting attention to the AA-acts in $\mathcal{F}_1^{AA}$ has the pragmatic advantage of a more parsimonious representation; it also allows at a formal level a different treatment of the two different sources of uncertainty. For example, Schmeidler's (1989) comonotonic independence

---

[26]More explicityl, the second condition requires that, for any $x \in X$ and any $\pi \in \Pi, \pi(f^{-1}(\{x\}) \cap A_i/A_i) = q_i^x$.

[27]Like Proposition 1, Proposition 3 also applies to superrelations of $\unrhd_{AA}$ defined on the product-algebra $\Sigma_1 \times \Sigma_2$ in the context of External Randomization (Example 3).

[28]Technically, $\mathcal{F}_1^{AA}$ is the subset of $\Sigma_1 \times \{\Omega_2\}$-measurable acts in $\mathcal{F}^{AA}$.

axiom is meaningful only restricted to $\mathcal{F}_1^{AA}$ rather than all of $\mathcal{F}^{AA}$.[29] While we view the case of $\trianglerighteq = \trianglerighteq_{AA}$ as an important special case, the applicability of the general construction to general convex-ranged likelihood relations has the substantial advantage of not having to assume there to be anything special about "randomness" as a distinct kind of uncertainty; for example, the construction applies equally well to superrelations of $\trianglerighteq_{AA}$ that incorporate additional "objective" information or "subjective" beliefs. Thus, our framework is attuned to the traditional Bayesian view that all probability is of one cloth.

Much of the formal power of the AA framework stems from the existence of a mixture operation on acts. This operation has a counterpart in the present epistemically enriched Savage framework as follows. We begin by defining "equal" ("fifty-fifty") mixtures. Consider $f \in [F]$ and $g \in [G]$; we need to define a class of Savage acts $f \oplus g$ that is contained in $[\frac{1}{2}F + \frac{1}{2}G]^{30}$ with a natural mixture interpretation. This can be done by borrowing the notion of an independent even-chance event from section 2.

Definition 4 (Equal Mixture) $h \in f \oplus g$ if $h = [f \text{ on } T, g \text{ on } T^c]$ for some event $T \in \Lambda$ such that $T \equiv T^c$ and such that $f^{-1}(x) \cap T \equiv f^{-1}(x) \cap T^c$ and $g^{-1}(x) \cap T \equiv g^{-1}(x) \cap T^c$ for all $x \in X$.

The restriction on T asserts that T is stochastically independent of the outcome incidences $f^{-1}(x)$ and $g^{-1}(x)$; this ensures that in fact $h \in [\frac{1}{2}F + \frac{1}{2}G]$, as desired. The definition of equal mixtures of Savage acts can be extended to dyadic mixtures $\alpha f \oplus (1-\alpha)g$ contained in $[\alpha F + (1-\alpha)G]$ (with α a dyadic number in $(0,1)$) in straightforward inductive manner. For example, the mixture $\frac{3}{4}f \oplus \frac{1}{4}g$ can be defined as the composite mixture $f \oplus (f \oplus g)$.[31]

The availability of an interpretable mixture-definition shows that the formal power of the Anscombe-Aumann framework can be replicated in the present explicitly epistemic framework. Indeed, the epistemic framework is more powerful since it allows to treat all uncertainty on par. This is of advantage when further belief-based restrictions on preferences are introduced, as illustrated by the companion papers (Nehring (2001, 2004)). The epistemic framework has also the advantage of greater

---

[29]For if comonotonic independence was required of the latter, it would entail full independence, i.e. SEU; this follows from Sarin-Wakker (1992) and, more directly, Nehring (2004).

[30]The latter is defined pointwise as usual.

[31]This construction parallels Ghirardato et al. (2003). Note, however, in the present setting there is no natural limit construction that would allow one to define arbitrary real-valued mixtures. However, since the dyadic numbers are dense in the reals, this does not seem to be problematic if preferences are appropriately continuous. Alternatively, one could directly define real-valued mistures by considering independent events T with arbitrary real probabilities $\overline{\pi}(T)$.

transparency in that all axioms can be formulated as restrictions of preferences over a standard, conceptually primitive object (i.e. Savage acts), possibly conditioned on epistemic data captured by the context $\unrhd$, rather than on a complex and, from the present point of view, non-primitive object such as Anscombe-Aumman acts.[32]

Earlier representations of the Anscombe-Aumann framework in a Savage setting have been obtained by Pratt-Raiffa-Schleifer (1964) and Klibanoff (2001a); in contrast to ours, the former assumes expected utility maximization, the latter utility sophistication. Our representation is more general also in that it applies to any convex-ranged context, and therefore does not assume the existence of a subjective randomization device as given by the context $\unrhd_{AA}$ defined in section 2.

Machina's (2004) work on "Almost-Objective Uncertainty" also reproduces the power of the Anscombe-Aumann framework in an enriched Savage setting in which the state space has the structure of a Euclidean manifold; like ours, he does not assume expected utility maximatization on unambiguous events. Machina derives the existence of "almost" unambiguous and conditionally unambigiuous events from epistemically motivated restrictions on preferences. However, these assumptions are imposed directly in the form of a smoothness condition, while we model the underlying beliefs directly as a likelihood relation, and obtain analogous preference restrictions Likelihood Consequentialism. One can show that the epistemic substance of Machina's preference restrictions can be modeled via likelihood relations that are "almost" convex-ranged.[33]

An altogether different route to mimicking the Anscombe-Aumann framework in a subjective setting based on a rich set of consequences rather than states is proposed by Ghirardato et al. (2003); since the mixture operation in their proposal is defined in utility terms, the implied interpretation of conditions on AA-preferences in their approach is very different from that in the present epistemically approach.

While these contributions have their distinct merits, none of them shares one key, motivating

---

[32]This complexity is a potential source of intransparency. For example, Eventwise Monotonicity of AA-preferences is much stronger than Eventwise Monotonicity of Savage preferences; moreover, while it does not constrain preferences over lotteries (constant AA-acts) in the standard product set-up (when imposed on $\mathcal{F}_1^{AA}$), it implies the Independence axiom over lotteries when imposed on $\mathcal{F}^{AA}$.

[33]We intend to establish a more precise and fruitful connection to Machina's contribution in future work. While our derivation is behaviorally general, Machina assumes that preferences are "event-smooth" which is behaviorally restrictive; event-smoothness excludes, for example, the minimum expected-utility model due to Gilboa-Schmeidler (1989). Finally, Machina exploits the specific structure of uniformly continuous densities on the real line (or, more generally, on Euclidean manifolds); by contrast, our approach is applicable to arbitrary state spaces.

advantage of the present, epistemic approach, namely its ability to record and derive the implications of *additional* beliefs beyond those giving rise to the AA structure in the first place.[34]

4. DISCUSSION

4.1 Ordinal vs. Cardinal Information about Consequence Valuations

It might be argued that Likelihood Consequentialism is too weak a rationality criterion since it exploits ordinal information about consequences only. Indeed, much stronger normative restrictions can be obtained if one exploits cardinal information about comparisons of utility differences. A theory of "Expected Utility in the Presence of Ambiguity" along such lines is developed in the companion paper Nehring (2004). It is not pursued here, since these stronger constraints on the relation between preferences and beliefs conflict with the desideratum of behavioral generality adopted in this paper.

4.2 Non-Exhaustive vs. Exhaustive Interpretation of Belief Context

Stronger rationality restrictions on behavior can also be derived when the likelihood relation is assumed to be an *exhaustive* representation of the decision-maker's beliefs under which incompleteness of the likelihood relation reflects deliberate *suspensions of judgement*. For it may then be argued that symmetries in the likelihood relation should be mirrored isomorphically in symmetries in preferences/choices; see Jaffray (1989), Nehring (1991, 2000) and Walley (1991, p. 227-228) for arguments along these lines. On the more general non-exhaustive interpretation adopted here, these isomorphism conditions are inappropriate since subrelations of a given relation may exhibit symmetries that differ from the original ones.

4.3 State-Dependent Preferences and Intrinsic Event Attitudes

Conversely, Likelihood Consequentialism might be criticized as too strong a rationality requirement, for example for reasons of state-dependence of preference. We will however argue that in many (perhaps in all) cases, such objections can be overcome by a more refined modeling.

For Likelihood Consequentialism to be a normatively compelling axiom, consequences must be

---

[34]Recall the discussion at the end of Example 3 above.

properly individuated, in the sense that a given consequence x is equivalent for the decision-maker in any state *in all valuation-relevant aspects*[35]; we will refer to this as the case of "state-independent preferences". State-dependence of preferences can be modeled by including aspects of the state into the description of the consequences. For example, when states describe the DM's health, gaining 1000\$ when healthy may not be worth as much as gaining 1000\$ when sick; this difference can be accounted for by distinguishing the consequences (1000\$,healthy) and (1000\$,sick). In such cases, one needs to abandon the assumption that the set of possible consequences is identical across states; that is, one would define Savage acts as mappings from $f : \omega \rightarrow \prod_{\omega \in \Omega} X_\omega$. Note that Likelihood Consequentialism remains meaningful but loses some power; it becomes vacuous in the limiting case in which the state-specific consequence sets X_ω are pairwise disjoint.

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

State-dependence of preference is of interest in the present context especially since it allows to capture many – and arguably all – instances of an apparent "intrinsic" attitude towards events. For instance, Chew and Sagi (2003) suggest that decision makers may have a taste for betting on particular types of events that overrides their likelihood assessments.[36] For example, on February 1, 2003, a DM may have attributed equal probability to Saddam Hussein's surviving a US led invasion of Iraq, and to the Iraqui soccer team winning against Brazil. However, the DM may have preferred to bet on the Iraqui soccer team rather than on Saddam Hussein; he may, for example, have expected taking special joy from winning a bet on an underdog team, but regretting having profited from an unjust cause. On their face, such preferences might appear to challenge the normative appeal of Likelihood Consequentialism in that non-likelihood features of events seem to be legitimately valued by the DM. However, this challenge loses its force once one recognizes that the bets do not really involve the same (properly individuated) consequences. For the bets clearly entail different psychological outcomes in various states (the joy, the regret); while these matter to the decision

---

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

[36]Ergin-Gul (2004) make a related argument contrasting objective and subjective probabilities.

maker, they are not captured by a description ("individuation") of consequences in terms of net wealth alone. Again, Likelihood Consequentialism continues to apply once acts are described in terms of properly individuated consequences.[37]

4.4 Epistemic Contexts from a Purely Behavioral Viewpoint

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 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[38]. 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 this position render the notion of decision-making in an epistemic context empirically empty?

---

[37]We have appealed to the informal notion of "proper individuation"/"state-independent preferences" without characterizing it in behavioral terms. This seems unavoidable, since it is unclear what behavioral conditions could take its place. Clearly, Eventwise Monotonicity is not sufficient as a behavioral criterion, since it captures at most the ordinal implications of state independence. One natural candidate for a *sufficient* condition is the Certainty Independence condition of Ghirardato et al. (2002), since Certainty Independence secures the separation of a state-independent utility function from event attitudes in a strong sense; Certainty Independence can be viewed as a cardinal strengthening of Savage's axiom P4 (Nehring 2004). However, Certainty Independence or even P4 are not necessary for state-independence, since P4 may fail due to prize-dependent ambiguity attitudes; see Klibanoff et al. (2005) for a worked-out model with this feature.

Furthermore, adapting a recent argument by Karni (1996), it can be argued that in principle no behavioral condition can yield more than a "prima-facie" criterion. Roughly speaking, Karni argued that under SEU, preferences identify statewise utilities only up to positive affine transformations *state-by-state*; thus, if the "true" consequence utilities are not constant across states, Savage's "revealed likelihood" relation differs from the agent's true likelihood. In the present setting, this amounts to saying that Likelihood Consequentialism will be violated.

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

Evidently Likelihood Consequentialism 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 Consequentialism is absorbed into a behavioral definition of compatibility with an epistemic context: simply say that a preference relation $\succsim$ satisfy *Compatibility-with-*$\trianglerighteq$ if the pair $(\succsim, \trianglerighteq)$ satisfies Likelihood Consequentialism. 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 condition is therefore clear-cut.[39]

One can summarize this behaviorist use of likelihood relations by saying that the context $\trianglerighteq$ represents an "epistemic *constraint* on preferences" rather than an independent "epistemic primitive". From this point of view, results 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.

---

[39]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 $\trianglerighteq$.

A.1   Unambiguous Events

Comparison to Zhang (1999).—

Consider the restriction $\unrhd_{|\Lambda}$ of any likelihood relation $\unrhd$ satisfying the assumptions of Theorem 2 on Σ to the family of unambiguous events Λ. By construction, $\unrhd_{|\Lambda}$ is complete (on Λ); Theorem 2 implies that $\unrhd_{|\Lambda}$ can be represented by a (dyadically) convex-ranged, finitely additive set function $\overline{\pi}$ on Λ; moreover, $\overline{\pi}$ can be extended to some finitely additive probability measure π on all of Σ.

Zhang (1999) considered likelihood relations defined on arbitrary given lambda-systems as primitives and characterized those relations that are representable by a convex-ranged, countably additive set function on Λ. Zhang's result is a key ingredient in Epstein-Zhang's (2001) characterization of revealed unambiguous events discussed below. His result is not directly comparable to the corollary to Theorem 2 described in the preceding paragraph, as it derives a weaker conclusion from weaker premises. Zhang's assumptions on the likelihood relation are weaker in that they apply only to Λ and not to (an incomplete relation defined on) some super-algebra Σ; on the other hand, his conclusion is weaker as well in that it does not imply representation by an additive set-function that can be extended to all of Σ. It is not known under which conditions such an extension exists; Epstein (1999) and Nehring (1999) provide counterexamples in finite state-spaces. In cases in which such an extension does not exist, the likelihoood relation $\unrhd_{\Lambda}$ (viewed as an incomplete relation on Σ) is *inconsistent* in the terminology of section 2.5; by consequence, it stands to reason that such likelihood relations do not represent a well-defined set of probabilistic beliefs.

Comparison to Epstein-Zhang (2001).—

It is also of interest to compare the present definition of unambiguous events based on explicit beliefs Λ ("explicitly unambiguous events") to the preference-based definition proposed by Epstein-Zhang (2001) Λ^{EZ}. A central issue arising from any behavioral definition is the extent to and sense in which preferences over unambiguous acts reveal the decision-maker's unconditional probabilistic beliefs. Epstein-Zhang (2001) provide a partial answer by showing under certain assumptions on preferences and a richness assumption on the endogenously defined family Λ^{EZ} that preferences over unambiguous acts are probabilistically sophisticated with respect to an additive set-function on Λ^{EZ}.

For simplicity, we shall confine the following discussion to the case of two consequences ($\#X =$

$\{x, y\}$ with $x \succ y$). In this setting, the act "betting on A" [x on A, y on A^c] will be denoted simply by the set A. Using this notation, an event T is *EZ-unambiguous* ($T \in \Lambda^{EZ}$) if, for all A, B disjoint from T, $A \succsim B$ if and only $A + T \succsim B + T$, and if the same holds for T^c instead of T.

Two questions arise naturally. First, if preferences are compatible with explict beliefs, are the explicitly unambiguous events also EZ-unambiguous; in other words: will "truly" unambiguous events be revealed as such by the EZ definition? Not necessarily; indeed, this happens only if betting preferences and beliefs are related by the following "Union Invariance" condition: for all $T \in \Lambda^0$ and all A, B disjoint from T, $A \succsim B$ if and only $A + T \succsim B + T$. While this condition has intuitive appeal, it is clearly a substantive restriction on ambiguity attitudes.[40]

Conversely, given a preference relation $\succsim$ with EZ-unambiguous events Λ^{EZ}, does there exist necessarily exist a coherent epistemic context $\unrhd$ compatible with $\succsim$ such that $\Lambda^{EZ} \subseteq \Lambda$, where Λ is the family of unambiguous events associated with that context? Again the answer appears to be negative in general, and is definitively negative in finite settings.[41] First, in view of the discussion of Zhang (1999) above, the likelihood relation revealed on Λ^{EZ} may be inconsistent, hence there simply may not exist any coherent context with associated Λ such that $\Lambda^{EZ} \subseteq \Lambda$ and such that preferences over EZ-unambiguous acts $\succsim_{|\Lambda^{EZ}}$ are compatible with $\unrhd$. Second, in cases in which such a context exists, it will contain substantial further likelihood comparisons over ambiguous events, entailing substantial further restrictions on preferences over bets on ambiguous events. While the EZ definition of unambiguous events entails substantial restrictions on these preferences of its own through the Union Invariance condition, it is not clear and seems a priori doubtful that these would be strong enough to encompass all of the restrictions entailed by $\unrhd$ in general. Thus, it remains an interesting question for future research under which conditions a DM's preferences over EZ-unambiguous events reflect genuine (coherent) probabilistic beliefs over these events in the sense of this paper.

---

[40]In a similar vein, Klibanoff et al. (2005) have pointed out that the Epstein-Zhang definition makes substantive implicit assumptions about the decision-maker's ambiguity attitudes.

[41]Note that, the Epstein-Zhang (2001) definition applies (and is meant to apply by Epstein and Zhang) whether or not the resulting family Λ^{EZ} is rich; in particular, it applies to finite state-spaces.

A2.   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, $B_2 \triangleright A_1$, and thus by transitivity, $B_2 \triangleright A_2$ and $B_1 \triangleright A_1$. Thus by Strong Additivity (Lemma 1), $B_1 + B_2 \triangleright A_1 + A_2$, the desired contradiction.

Lemma 1 *(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,* $A + C \triangleright B + D$ *if in addition* $A \triangleright B$ *or* $C \triangleright D$.

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. $\square$∎

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 $C \unrhd D$ imply $A \unrhd B$.

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 $A' \unrhd B'$, hence by Splitting $A \unrhd B$.

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

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 $A_i \unrhd B_i$ 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 $\widehat{\unrhd}$ of $\unrhd$. Let $\mathcal{D}$ denote the set of dyadic-valued random-variables, $\mathcal{D} := \{Z : \Omega \to \mathbf{D}, Z \text{ is } \Sigma\text{-measurable and has finite range}\}$. Any finite-ranged Z can be canonically written as $\sum_i z_i 1_{E_i}$, where $E_i = Z^{-1}(\{z_i\})$. For 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 $\widehat{\unrhd}$ on $\mathcal{D}$ as follows,

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

To establish various properties of $\widehat{\unrhd}$, 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 $i \in I$, $\sum_{i \in I} A_i \in \alpha \left(\sum_{i \in I} E_i \right)$.

36

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 $x_i = \frac{\ell_i}{2^{k_i}}, y_i = \frac{\ell_i'}{2^{k_i}}$, and $z_i = \frac{\ell_i''}{2^{k_i}}$. Split D_i into 2^{k_i} equally likely events $\{D_{i1}, ..., D_{i2^{k_i}}\}$, and set $C_i := \sum_{j \leq \ell_i} D_{ij} \in z_i D_i$, $A_i = \sum_{j=\ell_i+1}^{\ell_i+\ell_i'} D_{ij} \in x_i D_i$, and $B_i = \sum_{j=\ell_i+1}^{\ell_i+\ell_i''} D_{ij} \in y_i D_i$. 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 $C \backslash (\sum_i B_i) = \sum_i (C_i \backslash B_i) \in [X]$, as desired. $\square$

Lemma 9 *The relation $\widehat{\unrhd}$ on $\mathcal{D}$ is transitive, reflexive and satisfies the following conditions*

1. *(Extension)* $\quad 1_A \widehat{\unrhd} 1_B$ *if and only if* $A \unrhd B$.

2. *(Positivity)* $\quad X \widehat{\unrhd} \mathbf{0}$ *for all* X.

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

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

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

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

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

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

To verify <u>Transitivity</u>, 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.

<u>Weak Homogeneity</u> is an immediate consequence of Lemmas 3 and 5.

To verify <u>Additivity</u>, consider any X, Y, Z such that $X + Z, Y + Z \in \mathcal{D}$. According 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 $X + Z \widehat{\trianglerighteq} Y + Z$.

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

To verify this claim, take any X', Y' such that $\| X' - X \| \leq \frac{1}{K}$ and $\| Y' - Y \| \leq \frac{1}{K}$. By the Positivity and Strong Additivity of $\trianglerighteq$, it is without loss of generality to assume that X' (respectively Y') is measurable with respect to the partition generated by X (respectively Y), and that $X' \geq X$ and $Y' \leq Y$. Then there exist by Lemma 8ii) $A' \in [X' - X]$ such that $A + A' \in [X']$; likewise, by Lemma 8iii), there exist and $B' \in [Y - Y']$ and $B'' \in [Y']$ such that $B' + B'' = B$. Clearly, A' and B' are $\frac{1}{K}$−events, and therefore it is not the case that $A + A' \trianglerighteq B \backslash B' = B''$. Therefore, in view of Lemma 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 $\mathcal{B}$ is the completion of $\mathcal{D}$, and thus $\mathcal{B} \times \mathcal{B}$ of $\mathcal{D} \times \mathcal{D}$, the closure $cl\widehat{\trianglerighteq}$ of $\widehat{\trianglerighteq}$ in $\mathcal{B} \times \mathcal{B}$ restricted to $\mathcal{D} \times \mathcal{D}$ is simply $\widehat{\trianglerighteq}$, since $\widehat{\trianglerighteq}$ is closed in $\mathcal{D} \times \mathcal{D}$. Thus, $cl\widehat{\trianglerighteq}$ is an extension of $\widehat{\trianglerighteq}$, and will be referred to as "$\widehat{\trianglerighteq}$ on $\mathcal{B}$", or simply also as "$\widehat{\trianglerighteq}$" if no misunderstanding is possible. Clearly $X\widehat{\trianglerighteq}Y$ if and only if there exist sequences $\{X_n\}$ and $\{Y_n\}$ in $\mathcal{D}$ converging to X and Y, respectively, such that $X_n\widehat{\trianglerighteq}Y_n$ for all n.

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

Lemma 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 $\alpha_n \leq \min\left(\frac{\|X\|}{\|X_n\|}, \frac{\|Y\|}{\|Y_n\|}\right)$. This ensures that, for all $n, \ell\alpha_n X_n \in \mathcal{D}$ and $\ell\alpha_n Y_n \in \mathcal{D}$. By Weak Homogeneity of $\widehat{\trianglerighteq}$ on $\mathcal{D}$, $\alpha_n X_n\widehat{\trianglerighteq}\alpha_n Y_n$ for all n. Hence by $(\ell-1)$-fold application of Strong Additivity of $\widehat{\trianglerighteq}$ on $\mathcal{D}$, also $\ell\alpha_n X_n\widehat{\trianglerighteq}\ell\alpha_n Y_n$ for all n. By Continuity on $\mathcal{B}$, $\ell\alpha X\widehat{\trianglerighteq}\ell\alpha Y$, as desired.

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

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

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

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

39

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

In a final step, extend $\widehat{\unrhd}$ on $\mathcal{B}$ to the set of all bounded random-variables $\mathcal{R} := B(\Sigma, \mathbf{R})$ by defining $\widehat{\unrhd}$ on $B(\Sigma, \mathbf{R})$ as the unique relation $\widetilde{\unrhd}$ on $B(\Sigma, \mathbf{R})$ that coincides on $\mathcal{B}$ with $\widehat{\unrhd}$ 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 $\widehat{\unrhd}$ on $\mathcal{B}$). As in section 2.2, say that a relation $\widehat{\unrhd}$ 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 $\widehat{\unrhd}$ 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 3 *A relation $\widetilde{\unrhd}$ 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 $X, Y \in \mathcal{R}$,*

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

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

To complete the proof, apply Theorem 3 to the relation $\widehat{\unrhd}$ on $\mathcal{R}$ obtained in Lemma 11. By Extension, for all $A, B \in \Sigma$,

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

Thus Π is indeed a multi-prior representation of $\unrhd$. 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 $\widehat{\unrhd}_{\Pi'}$. From the uniqueness part of Theorem 3, there exist $X, Y \in \mathcal{R}$ such that $X \widehat{\unrhd} Y$ and not $X \widehat{\unrhd}_{\Pi'} Y$, or such that $X \widehat{\unrhd}_{\Pi'} Y$ and not $X \widehat{\unrhd} Y$. Consider the former case; the latter is dealt with symmetrically. By Additivity and Homogeneity, it can be assumed that $X, Y \in \mathcal{B}$. By continuity,

monotonicity, and the density of $\mathbf{D}$ in $[0,1]$ it can in fact be assumed that $X, Y \in \mathcal{D}$. Take any $A \in [X]$ and $B \in [Y]$. By Extension, $1_A \widehat{\cong} X$ and $1_B \widehat{\cong} Y$, hence $A \widehat{\unrhd} B$. 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 $\unrhd$.

Lemma 12 *For any $\pi \in \Pi'$ such that $\widehat{\unrhd}_{\Pi'} = \widehat{\unrhd}$, 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 $\widehat{\unrhd}_{\Pi'} = \widehat{\unrhd}$, $\pi(E_{ij}) = \pi(E_{ij'})$ for all i, j, j', hence $\pi(A_i) = \frac{\ell_i}{2^{k_i}} \pi(E_i)$ by additivity of π. Hence $\pi(A) = \sum_i \frac{\ell_i}{2^{k_i}} \pi(E_i) = E_\pi X$. $\square$

Proof of Fact 3.

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 $A \backslash A_i \unrhd B \cup B_j$ for all $i \in I$, $j \in J$. By consistency, $\Pi_\unrhd \neq \emptyset$. For all $\pi \in \Pi_\unrhd$, $\pi(A \backslash A_i) \geq \pi(B)$ for all $i \in I$, hence

$$\pi(A) = \frac{1}{\#I - 1} \Sigma_{i \in I} \pi(A \backslash A_i) \geq \frac{\#I}{\#I - 1} \pi(B). \tag{8}$$

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

$$\min_{\pi \in \Pi_\unrhd} \pi(A) \geq \frac{1}{\#J}. \tag{9}$$

By (8), $\pi(B) \leq \frac{\#I - 1}{\#I} \pi(A)$ for all $\pi \in \Pi_\unrhd$, and thus by (9)

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

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 $A \backslash A_i \unrhd B \cup B_j$ by coherence. $\square$

The following lemma is used in the proof of Proposition 1 below.

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

41

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

$$sup_k\overline{\pi}(A_{[k]}) + sup_k\overline{\pi}(A^c_{[k]}) = 1. \tag{10}$$

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 (10), 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

$$\frac{\ell_j+1}{2^j} \geq \alpha. \tag{11}$$

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, $A_{[k]} = \sum_{j \leq k} A'_{[k]}$; 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} 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

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

42

Therefore, using step 1,

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

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

$$\sup_{j \to \infty} \overline{\pi}\left(B_{[j]}\right) \geq \sup_{j \to \infty} \overline{\pi}\left(B_j\right) = \alpha \overline{\pi}\left(A\right). \tag{12}$$

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

$$B_{[2j]}^c \supseteq C_j + A_{[2j]}^c.$$

Hence

$$\sup_{j \to \infty} \overline{\pi}\left(B_{[j]}^c\right) \geq \sup_{j \to \infty} \overline{\pi}\left(C_j\right) + \sup_{j \to \infty} \overline{\pi}\left(A_{[j]}^c\right) = (1 - \alpha)\overline{\pi}\left(A\right) + (1 - \overline{\pi}\left(A\right)) = 1 - \alpha \overline{\pi}\left(A\right). \tag{13}$$

Combining (12) and (13), it follows that $B \in \mathcal{A}$ and $\overline{\pi}\left(B\right) = \alpha \overline{\pi}\left(A\right)$, demonstrating convex-rangedness. $\square$

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 applied to the λ-system Λ_A, there exists an event $B \in \Lambda_A$ such that $\overline{\pi}\left(B/A\right) = \alpha$, verifying convex-rangedness.

Proof of Proposition 1.

If Λ is a σ-algebra, or if more generally Λ is an algebra with $\overline{\pi}$ convex-ranged, then Likelihood Consequentialism restricted to betting preferences implies that the revealed likelihood relation $\succsim_\ell$ agrees with $\unrhd$ on Λ, and Likelihood Consequentialism 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-valuedness 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 $\mathcal{A}_f^\perp \supseteq \mathcal{B}_f \cup \mathcal{A}^\perp$ and $\mathcal{A}_g^\perp \supseteq \mathcal{B}_g \cup \mathcal{A}^\perp$.

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 Proposition 2.

"Only if". Suppose $\Psi \subseteq \Pi$. Let ψ^- denote the lower probability with respect to Ψ. Evidently, $\psi_{|\Lambda}^- = \overline{\pi}$, hence $\succsim$ maximizes expected utility on unambiguous acts and therefore satisfies the sure-thing principle on those acts.

To verify Likelihood Consequentialism, take any $f \in \mathcal{F}$, $x, y \in X$ such that $x \succsim y$ and events $A, B \in \Sigma$ such that $A \trianglerighteq B$. Let $g = [x \text{ on } A \backslash B; y \text{ on } B \backslash A; f(\omega) \text{ elsewhere}]$ and $h = [x \text{ on } B \backslash A; y \text{ on } A \backslash B; f(\omega) \text{ elsewhere}]$, and let $\pi' \in \arg \min_{\pi \in \Psi} \sum_{x \in X} u(x) \pi(\{\omega : g(\omega) = x\})$. Since $A \trianglerighteq B$ and $\Psi \subseteq \Pi$, $\pi'(A) \geq \pi'(B)$, and thus

$$
\begin{aligned}
\min_{\pi \in \Psi} \sum_{x \in X} u(x) \pi(\{\omega : g(\omega) = x\}) &= \sum_{x \in X} u(x) \pi'(\{\omega : g(\omega) = x\}) \geq \\
\sum_{x \in X} u(x) \pi'(\{\omega : h(\omega) = x\}) &\geq \min_{\pi \in \Psi} \sum_{x \in X} u(x) \pi(\{\omega : h(\omega) = x\}),
\end{aligned}
$$

and thus $g \succsim h$; the strict part of Likelihood Consequentialism is shown similarly.

"If". Conversely, suppose that $\succsim$ is compatible with $\trianglerighteq$ and satisfies Savage's Sure-Thing Principle P2 on the set of unambiguous acts $\mathcal{F}^{ua}$. By Savage's Theorem and Proposition 2, $\succsim$ maximizes expected utility on unambiguous acts, with $\psi_{|\Lambda}^- = \overline{\pi}$.

If not $\Psi \subseteq \Pi$, then there must exist $A \in \Sigma$ such that $\psi^-(A) < \pi^-(A)$ by the uniqueness part of Theorem 2. By the convex-rangedness of $\overline{\pi}$, there exists $T \in \Lambda$ such that $\overline{\pi}(T) = \pi^-(A)$. Since T is unambiguous, $\pi^+(T) \leq \pi^-(A)$ and thus $T \trianglelefteq A$; on the other hand, since $\psi^-(T) = \overline{\pi}(T) =$

$\pi^-(A) > \psi^-(A)$, for any $x \succ y$ one has $[x$ on A, y on $A^c] \prec [x$ on T, y on $T^c]$, violating Likelihood Consequentialism. $\square$

Proof of Fact 4.

Let $u : X \longrightarrow \mathbf{R}$ a non-constant function, ϕ be a strictly increasing mapping from $[0,1]$ onto itself, and $\upsilon : \Sigma_1 \to [0,1]$ the capacity $\phi \circ \overline{\pi}_1$. Any act f can be written as $[x_{i,j}$ on $A_i \times T_j]_{i,j}$. Define $Z_f : \Sigma_1 \to R$ by setting $Z_f(\omega) = \left(\Sigma_j u(x_{i(\omega),j})\overline{\pi}_1(\omega)\right)$ where $i(\omega)$ is given as the unique i such that $A_i \ni \omega$. Now define $\succsim$ as follows: by setting

$$f \succsim g \text{ if and only if } \int Z_f d\upsilon \geq \int Z_g d\upsilon, \tag{14}$$

where the integral is the Choquet integral. The preference relation $\succsim$ can be viewed as a special case of the CEU model due to Schmeidler (1989) adapted to a Savage framework. In fact, (14) defines exactly the class of "second-order probabilistically sophisticated CEU" preferences studied in Ergin-Gul (2004).

It is clear that $\succsim$ is probabilistically sophisticated on $\mathcal{F}_1 \cup \mathcal{F}_2$ if and only if $\phi = id$, i.e. $\upsilon = \overline{\pi}_1$, in which case it is SEU. Likelihood Consequentialism with respect to $\unrhd_{AA}$ is immediate from the construction. To verify Likelihood Consequentialism with respect to $\unrhd_1$, consider any $A', B' \in \Sigma_1$ such that $A := A' \times \Omega_2 \unrhd B := B' \times \Omega_2$, $f \in \mathcal{F}$ and $x, y \in X$ such that $x \succsim y$. Setting $g = [x$ on $A \backslash B$; y on $B \backslash A$; f elsewhere] and $h = [x$ on $B \backslash A$; y on $A \backslash B$; f elsewhere], note that, in act notation, $Z_g = [u(x)$ on $A' \backslash B'$; $u(y)$ on $B' \backslash A'$; Z_g elsewhere] and $Z_h = [u(x)$ on $B' \backslash A'$; $u(y)$ on $A' \backslash B'$; Z_g elsewhere]. Since $\overline{\pi}_1(A') \geq \overline{\pi}_1(B')$ by assumption, one has

$$\upsilon(\{\omega\} : Z_g(\omega) \geq z) \geq \upsilon(\{\omega\} : Z_h(\omega) \geq z)$$

for all $z \in \mathbf{R}$, and thus $\int Z_g d\upsilon \geq \int Z_h d\upsilon$ by the definition of the Choquet integral, verifying the weak part of Likelihood Consequentialism with respect to $\unrhd_1$; the strict part is verified analogously. $\square$

Proof of Proposition 3.

The relation $\succsim^{AA}$ is uniquely defined since convex-rangedness of the context ensures that $[F]$ is non-empty.

Consider any F and G such that $F(\omega)$ stochastically dominates $G(\omega)$ for all $\omega \in \Omega$. Take any partition $\{A_i\}_{1,...,n}$ such that both F and G are measurable with respect to this partition. For

$i = 0, .., n$, define AA acts $F_i = [F(\omega)$ on $\sum_{j \leq i} A_j$, $G(\omega)$ on $\sum_{j > i} A_j]$, and take Savage acts $f_i \in [F_i]$. By Proposition 1, for $i \geq 1$, $f_i \succsim f_{i-1}$. Since by construction $f_0 \in [G]$ and $f_n \in [F]$, by transitivity of $\succsim$ one infers that $F \succsim^{AA} G$, demonstrating the "weak" part of monotonicity. The strict part follows from an analogous argument.

Completeness of $\succsim^{AA}$ is an immediate consequence of the completeness of $\succsim$. To verify transitivity, assume that $[F] \succsim^{AA} [G]$ and $[G] \succsim^{AA} [H]$. By definition of $[.]$, there exist $f \in [F], g, g' \in [G], h \in [H]$ such that $f \succsim g$ and $g' \succsim h$. Note that since $g, g' \in [G]$, they stochastically dominate each other, and therefore $g \sim g'$, whence by transitivity of $\succsim$, $f \succsim h$, as needed to be shown. $\square$

REFERENCES

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

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

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

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

[5] Blume, L., Brandenburger, A. and E. Dekel (1991): "Lexicographic probabilities and choice under uncertainty," *Econometrica* 59 61–79.

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

[7] 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.

[8] 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.

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

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

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

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

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

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

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

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

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

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

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

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

[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), "Certainty Independence and the Separation of Utility and Beliefs," *Journal of Economic Theory* 120, 129-136.

[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, pp. 107-112.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[47] Nehring, K. (2004): "Bernoulli without Bayes: A Theory of Utility-Sophisticated Preference", mimeo.

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

[49] 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.

[50] 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.

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

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

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

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

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

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

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

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

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