Rochester Center for

Economic Research

Coarse Contingencies

Epstein, Larry G., and Massimo Marinacci

Working Paper No. 515
April 2005

UNIVERSITY OF
ROCHESTER

COARSE CONTINGENCIES*

Larry G. Epstein   Massimo Marinacci

April 20, 2005

Abstract

The paper considers an agent who must choose an action today under uncertainty about the consequence of any chosen action but without having in mind a complete list of all the contingencies that could influence outcomes. She conceives of some relevant (subjective) contingencies but she is aware that these contingencies are coarse - they leave out some details that may affect outcomes. Though she may not be able to describe these finer details, she is aware that they exist and this may affect her behavior.

*Epstein is at Department of Economics, University of Rochester, Rochester, NY 14627, lepn@troi.cc.rochester.edu; Marinacci is at Dipartimento di Statistica e Matematica Applicata, Università di Torino, Italy, massimo.marinacci@unito.it. We are grateful to Bart Lipman, Fabio Maccheroni, Mark Machina, Luigi Montrucchio and Todd Sarver for helpful discussions.

1. INTRODUCTION

Consider an agent who must choose an action today under uncertainty about the consequence of any chosen action but without having in mind a complete list of all the contingencies that could influence outcomes. She conceives of some relevant (subjective) contingencies or states of the world but she is aware that these contingencies are coarse - they leave out some details that may affect outcomes. Though she may not be able to describe these finer details, she is aware that they exist and this may affect her behavior. How does one model such an agent?

The standard Savage framework, based on a primitive state space, is inappropriate for two reasons. First, in the Savage model each state is a complete description of the world - it determines a unique outcome for any chosen action. Second, even if we knew how to model a "coarse or incomplete state" and we redefined the Savage state space accordingly, the resulting approach would still be unsatisfactory if, as in Savage, the state space were adopted as a primitive. In that case, the state space is presumed observable by the modeler, an assumption that is all the more problematic when states are coarse. Ideally, the agent's conceptualization of the future should be subjective - it should be derived from preference, that is, from in principle observable behavior.

Kreps [8, 9], and Dekel, Lipman and Rustichini [1] have rendered the state space subjective by positing preference over menus of alternatives, thus addressing the second concern. However, we argue in this paper that their models do not capture coarse perceptions. We focus primarily on the model of Dekel, Lipman and Rustichini (henceforth DLR). They describe (p. 893) the agent they are modeling: "... she sees some relevant considerations, but knows there may be others that she cannot specify. For simplicity, we assume henceforth that the agent conceives of only one situation, 'something happens,' but knows that her conceptualization is incomplete." Though they frequently refer to "unforeseen contingencies", it seems that, at least in part, they have in mind what we prefer to call "coarse contingencies." Later (pp. 919-20), they describe what is needed for a critique of their model: "... just as Ellsberg identified the role of the sure-thing-principle in precluding uncertainty-averse behavior, we believe that one must first find a concrete example of behavior that is a sensible response to unforeseen contingencies but that is precluded by our axioms. An important direction for further research is to see if there is such an Ellsbergian example for this setting and, if so, to explore relaxations of our axioms." This is the direction we pursue here.

Specifically, we argue that the Independence axiom imposed by DLR in their most restrictive model (the EU additive representation) rules out coarse contingencies. We focus also on their weakest Independence-style axiom, called Indifference to Randomization (IR), which they adopt (either explicitly or implicitly) in all of their representation theorems.[1] We argue that the case for IR is not clear-cut, at least given a particular conception of coarse perceptions. Thus we are led to explore two alternative directions for relaxing the DLR axioms - one which continues to assume IR and a second which replaces IR with a new axiom (called Coarseness) that captures our preferred conception of coarse perceptions. In the latter direction, a new functional form for utility is proposed, but our exploration of alternative axioms falls short of providing a complete characterization of the functional form and hence a new axiomatic model.

Our analysis is more conclusive when IR is adopted: the corresponding set of axioms characterizes a multiple-priors functional form for utility functions, paralleling the multiple-priors model of Gilboa and Schmeidler [5]. Their model was developed in order to accommodate aversion to ambiguity such as typified by intuitive behavior in the Ellsberg Paradox. Thus this representation result highlights a feature of retaining IR in attempting to model coarse perceptions: the agent modeled in this way is indistinguishable from one who foresees a complete set of states but is not sure of their likelihoods. Alternatively, one can view our multiple-priors representation result as a contribution to the literature on preferences under ambiguity because it extends the Gilboa and Schmeidler theorem by rendering the state space subjective.

Finally, we mention other connections to the literature on ambiguity. Inspired by Dempster [2] and Shafer [16], Mukerji [12] and Ghirardato [4] argue that an agent who is aware that she has only a coarse perception of the state space can be thought of as using a non-additive probability measure (or capacity). Their approach is much different than ours in that they take the agent's coarse perception as a primitive. However, our modeling approach has in common with theirs the interconnection between coarse perceptions and ambiguity.

The finding that coarse contingencies imply violations of Independence is relevant to the search for theoretical foundations for models with incomplete contracts. One reason for studying decision-making in the absence of Savage's state space is the idea that if some aspects of states are unforeseen or indescribable by the con-

[1] DLR (p. 911) mention the case where ex post utilities are not vNM in the context of establishing a result regarding minimality of the subjective state space. But such violations are not described as germane to the issue of coarse contingencies.

tracting parties, then contractual incompleteness should follow. Maskin and Tirole [11] have shown that this intuition is flawed if possible future payoffs can be *probabilistically* forecast and agents perform dynamic programming *as expected utility maximizers*. However, if agents with coarse perceptions violate Independence, as we suggest, then they do not have probabilistic beliefs or maximize expected utility, which leaves open the possibility of the sought-after foundations. Indeed, we know from Mukerji [13] that inefficient incomplete contracts can emerge if agents hold non-additive beliefs and maximize Choquet expected utility.

The paper proceeds as follows. Next we outline the DLR model and argue that their axioms preclude coarse contingencies. Then we describe two alternative relaxations of their axioms intended to capture coarseness. Proofs are relegated to appendices.

2. THE DLR MODEL

The DLR model has the following primitives:

- B: finite set of actions (these should *not* be thought of as "outcomes"); let $\mid B \mid = B$
- $\Delta(B)$: set of probability measures over B, endowed with the weak convergence topology; generic lotteries are $\beta,\ \gamma, ...$
- $\mathcal{X}$: closed subsets of $\Delta(B)$ generic elements are denoted $x,\ y,$ ... and are called *menus*[2]
- preference $\succeq$ is defined on $\mathcal{X}$

The agent ranks menus at time 0 (ex ante) using $\succeq$ with the understanding that at time 1 (ex post), she will choose a lottery from the previously chosen menu. One can think of a menu as corresponding to an action to be taken ex ante, where the significance of an ex ante action is that it limits options for further action ex post, that is, for the choice of β in $\Delta(B)$. There are no exogenous states of the world, but the agent may envisage some scenarios for time 1. She anticipates learning which scenario is realized before making her choice out of the menu. Thus her subjective conceptualization of the future affects her expected choices

[2] DLR do not restrict menus to be closed but this difference from their model is unimportant and we overlook it throughout.

out of menus and hence also her ex ante evaluation of menus. In other words, her subjective state space underlies the preference $\succeq$ and (under suitable assumptions) is revealed by it.

For example, the ranking

$$\{\beta, \beta'\} \succ \{\beta\} \succ \{\beta'\}$$

reveals that the agent conceives of a circumstance in which she would strictly prefer β over β' and also another circumstance in which she would strictly prefer β' over β. Under DLR's set of axioms, subjective contingencies concern only the possible ex post preference over lotteries. This is natural - payoffs rather than ex post physical states per se are ultimately all that matter.

DLR assume throughout that preference is complete, transitive and suitably continuous. They occasionally, though not universally, adopt also the next axiom.

Monotonicity: For all menus x' and x, $x' \supset x \Longrightarrow x' \succeq x$.

The axiom states that flexibility has non-negative value. For concreteness, we restrict attention here to models satisfying this property.

The first problematic axiom that we consider is Independence.[3] It refers to mixtures of two menus as defined by

$$\alpha x + (1-\alpha)\, y = \{\alpha\beta + (1-\alpha)\,\gamma : \ \beta \in x, \ \gamma \in y\}.$$

Formally, the indicated mixture of x and y is another menu and thus when the agent contemplates that menu ex ante, she anticipates choosing out of $\alpha x + (1-\alpha)\, y$ ex post. It follows that one should think of the randomization corresponding to the α and $(1-\alpha)$ weights as taking place at the end - after she has chosen some mixed lottery $\alpha\beta + (1-\alpha)\,\gamma$ out of the menu.

Independence: For all menus x', x and y and $0 < \alpha < 1$,

$$x' \succeq x \iff \alpha x' + (1-\alpha)\, y \succeq \alpha x + (1-\alpha)\, y.$$

Consider the axiom for an agent who is aware of the incompleteness of her subjective state space. For concreteness, suppose that her subjective conceptualization, like the objective (exogenous) one, is trivial - "something happens".

[3] DLR use the term Independence to refer to a weaker condition than what is stated below. However, the two axioms are equivalent given their continuity axiom.

Suppose further that she is indifferent between the menus $\{\beta\}$ and $\{\beta'\}$. Independence requires that

$$\{\alpha\beta' + (1-\alpha)\beta\} \sim \{\beta'\} \sim \{\beta\}.$$

Is this intuitive? She is aware ex ante that there are unforeseen (finer, or back-of-the-mind) contingencies that could affect the desirability of any action. Though she does not understand these finer details and may not be able to describe them, she is nevertheless aware that they exist, and she may feel that some may make β' more desirable ex post and some may make β more desirable. Randomization may hedge this uncertainty and thus the mixture might be strictly preferable to either lottery, that is,

$$\{\alpha\beta' + (1-\alpha)\beta\} \succ \{\beta'\} \sim \{\beta\},$$

in contradiction to Independence.

DLR show that in conjunction with completeness, transitivity and continuity, Independence implies the following axiom:[4]

Indifference to Randomization (IR): For every menu x, $x \sim co(x)$.

To evaluate this axiom, it is important to understand precisely the meaning of the time line sketched above. It describes the agent's ex ante expectations, for example, that ex post she will be able to choose from the menu that is chosen initially. The critical issue is what information she expects to have at that point. In fact, it may very well be that the true complete (Savage-like) state will be realized before she has to choose out of the menu. But since she does not conceive of them ex ante, she cannot be thinking explicitly in terms of the complete states that might be realized ex post. Rather, given her ex ante conceptualization in terms of coarse contingencies, one natural assumption is that she expects only to know which of these is true before choosing out of the menu. In that case, she expects coarseness to persist even ex post. On the other hand, she need not foresee all the complete states in order to believe that one of them will be realized ex post.[5] Thus an alternative assumption is that the agent anticipates that *some*

[4] $co(x) = \{\alpha\beta + (1-\alpha)\beta' : \beta, \beta' \in x,\ 0 \leq \alpha \leq 1\}$ denotes the convex hull of x. As in the case of Independence, one should think of the randomization as occuring after choice is made out of the menu.

[5] Nevertheless, it seems to us that this assumption presumes a higher degree of awareness of the underlying fine states than does the alternative.

complete state will be realized ex post. The intuitive appeal of IR depends on which of these assumptions is adopted.

If the agent anticipates that some complete state will be realized ex post, then she can be certain that her ex post preference over lotteries will conform to vNM. Thus she anticipates choosing out of the previously chosen menu x in order to maximize a mixture linear utility over lotteries, which means that she will do as well choosing out of x ex post as out of $co\,(x)$. Being certain of this ex ante, she will be indifferent between x and $co\,(x)$. This is the justification for Indifference to Randomization put forth by DLR.

Suppose, however, that coarseness is expected to persist ex post. Then the agent expects to be concerned ex post not only with how any given lottery β will play out, but also with how (payoff-relevant) back-of-the-mind uncertainty will be resolved eventually. This extra layer of uncertainty leads to non-vNM utilities ex post and subsequently to a possibly positive value for randomization. For example, suppose she chooses out of $x = \{\beta', \beta\}$ being aware that there are details that she cannot specify and that could affect the desirability of β' and β. Then she might feel that some may make β' more desirable and some may make β more desirable, and, as in the discussion of Independence, randomization could hedge some of this uncertainty. Thus the ex ante ranking

$$co\,(x) \succeq x,$$

is intuitive, but the indifference assumed in IR is too strong.

Both hypotheses concerning the agent's expectations seem to us to be descriptively plausible. In particular, an example to follow shortly (Savage's omelet) illustrates a case where the hypothesis that coarseness is expected to persist seems natural. Thus in the formal analysis below we consider two alternative sets of axioms - one where IR is imposed and one where it is replaced by an axiom that we feel captures coarse perceptions that are expected to persist.

There is an obvious question: if randomization is valuable, why can't the agent randomize on his own? For example, given the menu $x = \{\beta', \beta\}$ and if she so wished, then she could toss a suitably biased coin ex post to decide on whether β' or β is chosen, thereby providing for herself any desired mixture between β' and β. If ex ante she *believes* that she can commit to tossing such a coin *and* also to choosing according to its realized outcome, then she would presumably view herself as having not only x but indeed all of $co\,(x)$. Since such beliefs are not plausibly observable, we would be left not knowing how to interpret choices between menus - for example, the choice of x over y might reflect the preference for

$co(x)$ over $co(y)$ rather than of x over y. To avoid this problem, we assume that own-randomization is impossible and thus that each menu x indeed describes the complete set of lotteries from which choice is possible ex post - this assumption is at least implicit also in DLR. A possible justification is that any incentive to toss a coin ex post leads also to the incentive to toss it again, and again, thus rendering problematic commitment to choosing according to the outcome of the first toss.

Finally, for later reference, we describe the most restrictive utility functional form characterized by DLR - the non-negative additive EU representation. To express it, note that each mixture linear $u: \Delta(B) \longrightarrow \mathbb{R}^1$ can be identified with a (unique) vector in $N \subset \mathbb{R}^B$, where the role of the subset N is to normalize vNM utilities so that each u corresponds to a unique ordering of lotteries. (DLR's specification of N is not important here; later we adopt a different specification.) The utility of any menu has the form

$$\mathcal{W}(x) = \int max_{\beta \in x} \; u(\beta) \; d\mu(u), \tag{2.1}$$

where μ is a probability measure on N and $u(\beta) = \Sigma_{b \in B} \beta(b) u_b = u \cdot \beta$. The interpretation is that ex post preference over lotteries, represented by u, conforms to expected utility theory. Given u, then choice out of x will maximize $u(\beta)$, but ex ante, the agent does not know which preference will prevail ex post. The support of μ, corresponding to the set of ex post preferences that she views as possible, constitutes her subjective state space. To evaluate x, she computes its expected payoff assuming an optimal choice of lottery in each subjective state.

Savage's Omelet-Maker: Savage (pp. 13-15) illustrates application of his model through the example of a man who is preparing an omelet.[6] The omelet-maker must decide what to do with a sixth egg given that his wife has prepared a bowl containing five good eggs. There are three possible actions - break the egg into the bowl, break the egg into a saucer for inspection (which necessitates eventually washing the saucer), and discard the egg without inspection. Savage suggests that the choice problem fits into his framework where the state of the world describes the state of the sixth egg, $S = \{good, \; rotten\}$, and outcomes are the natural ones (see his Table 1). Of course, in Savage's approach, states of the world constitute

[6] We borrow also from Shafer's [17] discussion of Savage's example.

complete descriptions of the environment in that once a state is specified, each action leads to a unique outcome.

An obvious concern is that some omelet-makers in the above situation may perceive things differently. For example, as Savage remarks, some may be uncertain also about whether one rotten egg will spoil a six-egg omelet. Savage proposes that to model such a man, one simply expand the state space in the obvious way. However, there may be other uncertainties in the man's mind and the modeler can only guess what they are. The solution that follows from (Kreps and) DLR is to dispense with an exogenous state space and to infer all relevant uncertainties, that is, the subjective state space, from suitable choice behavior.

To describe the requisite choice behavior, assume that when deciding what to do with the sixth egg, which we term the choice of an ex ante action, the man contemplates future decisions (the choice of ex post actions) that are permitted by the ex ante decision. The set of conceivable ex post actions is $\Delta(B)$, where $B = \{b_h, b_5, b_5', b_6, b_6'\}$ and:

- b_h = go hungry
- b_5 = eat 5-egg omelet
- b_5' = eat 5-egg omelet and wash saucer
- b_6 and b_6' are defined similarly for 6-eggs

Then "break into bowl" may be identified with the menu $\{b_6, b_h\}$, "break into saucer" corresponds to the menu $\{b_5', b_6', b_h\}$ and "discard" translates into $\{b_5, b_h\}$. Other menus do not correspond to actions appearing in Savage's description, but they are arguably conceivable objects of choice, even when menus of lotteries over B are included. For example, $\{b_6\}$ corresponds to an irrevocable promise to eat the big omelet regardless of the state of the 6th egg, and $\{b_h\}$ corresponds to discarding all the eggs. More interesting is $\{b_5, b_6'\}$, which corresponds to breaking the 6th egg into the saucer under the agreement that if the 5-egg omelet is eaten, presumably because the 6th egg proves to be rotten, then the wife washes the saucer, while if the 6th egg proves to be good and the 6-egg omelet is eaten, then the man is responsible for washing the saucer. (Implicit is that the man does not care about his wife's dish-washing.) Similarly $\{b_5', b_6\}$ represents the opposite agreement about dish-washing.

How can we tell if the two states proposed by Savage underlie (or can be thought of as underlying) the man's deliberations? Suppose that

$$\{b_6, b_h\} \succ \{b_6\} \succ \{b_h\}.$$

DLR would interpret this ranking as indicating that there are at least two subjective states - one in which he would strictly prefer b_6 to b_h (presumably where the sixth egg is good) and one in which he would strictly prefer b_h (where the sixth egg is rotten). Thus the indicated ranking indicates that the two states "6th egg is good or rotten" are relevant.

Other uncertainties may also be relevant. For example, if

$$\{b_5', b_6', b_h\} \succ \{b_5', b_6'\},$$

then the omelet-maker anticipates the possibility that after breaking the 6th egg into the saucer, he would prefer to go hungry. This may be due to uncertainty about the condition of the first five eggs and the feeling that there may be a connection between the condition of the 6th egg and the freshness of the first five. Thus if he sees that the 6th egg is rotten, he may decide to go without an omelet entirely.

However, one may wish to go further and to ask not only if a particular contingency is relevant (in the man's mind), but also if it is coarse. For example, to this point we have referred to the man being uncertain about whether the 6th egg is "good" or "rotten", but he might view these as coarse descriptions: a "good" egg includes a range of possibilities - the egg may be superlative, in which case the 6-egg omelet would surely be preferable to going hungry, but it may also be stale, in which case, depending on how stale, going hungry may be preferable. Suppose that the man has these finer gradations of "good egg" in the back of his mind. Suppose further that he expects the coarseness to persist even after breaking open the 6th egg - after all, the only way to be certain about the exact quality of the egg is to eat it. Thus he anticipates being aware of the existence of missing details also after seeing the 6th egg and having to choose out of the previously chosen menu. Further, he anticipates also that some of the finer details (a superlative egg) would support choice of b_6 and that others (a very stale egg) would support the choice of b_h. Consequently, he prefers ex ante to have the option ex post of using a coin, presumably biased towards the b_6 outcome, to decide between b_6 and b_h. Then he would exhibit the ranking

$$co\left(\{b_6, b_h\}\right) \succ \{b_6, b_h\}, \tag{2.2}$$

in contradiction to IR.

Note that the ranking (2.2) is inconsistent with the alternative interpretation whereby instead of using the coarse state "good", the man conceives ex ante of the finer contingencies "stale but not rotten" and "superlative" and where he views these as complete descriptions. In that case, he would presumably anticipate knowing all relevant details after the 6th egg is broken, which would leave him indifferent to randomization as in the DLR model.

3. TWO ALTERNATIVE MODELS

3.1. Relaxing Independence

Consider Independence once again. The typical rationale for Independence (DLR, p. 905, for example) relies on the claim that the agent should be indifferent between the mixture $\alpha x + (1-\alpha) x'$ and the two-stage object $\alpha \circ x + (1-\alpha) \circ x'$, which represents a lottery over menus that delivers x with probability α and x' with probability $(1-\alpha)$ and where the lottery is played out immediately, that is, before any subjective uncertainty is resolved. But we argue that her being aware of the incompleteness of her conceptualization renders it intuitive only that

$$\alpha x' + (1-\alpha) x \succeq \alpha \circ x' + (1-\alpha) \circ x. \tag{3.1}$$

The usual intuition for Independence suggests that

$$x' \succeq x \implies \alpha \circ x' + (1-\alpha) \circ x \succeq x.$$

Hence we are led to the following weakening of Independence:

Preference Convexity: $x' \succeq x \implies \alpha x' + (1-\alpha) x \succeq x$.

To illustrate, using the example in our earlier discussion of Independence, where $x' = \{\beta'\} \sim \{\beta\} = x$, then the preceding states that

$$\{\alpha\beta' + (1-\alpha)\beta\} \succeq \alpha \circ \{\beta'\} + (1-\alpha) \circ \{\beta\} \sim \{\beta'\} \sim \{\beta\}.$$

The reason for the indicated weak preference, rather than indifference, is that $\{\alpha\beta' + (1-\alpha)\beta\}$ may hedge back-of-the-mind uncertainty, as suggested above, while $\alpha \circ \{\beta'\} + (1-\alpha) \circ \{\beta\}$ does not. For the latter, because the randomization takes place immediately, one is left ultimately with either $\{\beta'\}$ or $\{\beta\}$ when facing the coarsely perceived future beyond time 0. Thus the prior randomization is arguably of no value.

As another illustration, adapt DLR's illustrative example where the subjective conceptualization is "something will happen", and where $x' = \{\beta_1, \beta_2\}$ and $x = \{\overline{\beta}\}$. When facing x', she conceives of two sets of circumstances (or two events) ex post: those in which she would choose β_1 (event E_1) and those in which she would choose β_2 (event E_2). Thus the two stage object $\alpha \circ x' + (1-\alpha) \circ x$ is expected to lead to $\overline{\beta}$ with probability $1-\alpha$, and to (β_1 if E_1; β_2 if E_2) with probability α. On the other hand, if facing $\alpha x' + (1-\alpha)\, x$, then the randomization corresponding to α is completed only after she observes her (incomplete) subjective state. In particular, before randomization she will already know whether E_1 or E_2 has been realized. Thus it will be feasible for her to choose $\alpha\beta_1 + (1-\alpha)\,\overline{\beta}$ given E_1 and $\alpha\beta_2 + (1-\alpha)\,\overline{\beta}$ given E_2, and thus to receive ultimately $\left(\alpha\beta_1 + (1-\alpha)\,\overline{\beta} \text{ if } E_1; \alpha\beta_2 + (1-\alpha)\,\overline{\beta} \text{ if } E_2\right)$. She may be able to do better but she need not do worse.

How does she rank

$$\begin{array}{l} \left(\alpha\beta_1 + (1-\alpha)\,\overline{\beta} \text{ if } E_1; \alpha\beta_2 + (1-\alpha)\,\overline{\beta} \text{ if } E_2\right) \text{ versus} \\ \left((\beta_1 \text{ if } E_1; \beta_2 \text{ if } E_2)\,, \mathit{prob} = \alpha; \overline{\beta}, \mathit{prob} = (1-\alpha)\right)? \end{array} \tag{3.2}$$

For any assignment of (subjective) probabilities to E_1 and E_2, both prospects imply the identical distribution over B. One is tempted therefore to argue that they should be viewed as indifferent, which leads to a rationale for Independence. However, given the coarseness of contingencies, there is a sense in which the former involves less uncertainty and thus could be preferable ex ante. The agent knows that her conceptualization of the future is incomplete and hence that each E_i includes details, or finer contingencies, that affect the desirability of any action, or lottery β_i. Though she may not be able to specify or describe these finer contingencies, she is aware that they may exist and be relevant. This makes the prospect of choosing even the single β_i on E_i an uncertain prospect (even apart from the randomization involved in the lottery) and creates an incentive to hedge *within* E_i. Only the first prospect in (3.2) possibly affords such hedging, through mixing with $\overline{\beta}$, and therefore may be preferable.

The intuition that 'hedging' motives may render randomization valuable recalls the intuition provided in Schmeidler [15] and in Gilboa and Schmeidler [5] for their relaxations of Independence designed to accommodate ambiguity aversion such as exhibited in the Ellsberg Paradox. In their settings, there is an exogenous set of complete states, objects of choice are acts over these states, and hedging variation across states, which reduces ambiguity, is the rationale for randomization. In the present setting, states or contingencies are subjective and coarse, objects of choice are menus of actions rather than acts, and it is the hedging of variation across "missing details" that is the rationale for randomization. But the basic intuition is similar.

The corresponding axiom in Gilboa and Schmeidler [5] is called "uncertainty (meaning ambiguity) aversion." We have adopted a neutral name for our axiom because it admits rationales other than "coarseness aversion." Indeed, Preference Convexity can be understood also as expressing the gains from hedging for a Gilboa-Schmeidler-type agent who foresees a complete set of states but who is not sure about their likelihoods. To see that Preference Convexity is implied also by ambiguity, suppose the agent ex ante foresees each possible u, an ex post utility function over lotteries. Then she presumably anticipates choosing out of any given menu conditionally on the realization of each u. For example, given x, she anticipates choosing the lottery β_u if u is realized. Thus the menu x is equivalent for her to the (lottery-valued) act given by $u \longmapsto \beta_u$. Similarly, x' can be identified with an act $u \longmapsto \beta'_u$. Then $x' \succeq x$ translates into the weak preference for the primed act over the unprimed one. If states are ambiguous for her, then, as argued by Gilboa and Schmeidler, she may strictly prefer the α-mixture of these two acts to (β_u). But the mixed act is feasible for her by choosing conditionally on each u if she has the menu $\alpha x' + (1-\alpha)\, x$, and thus she can do at least as well with the latter menu as with x, which 'proves' Preference Convexity.

Given the two possible rationales for Preference Convexity, the question is how we might distinguish between them. We suggest that the two stories can be distinguished by the attitudes they imply towards randomization within menus ex post.

3.2. Model 1: Coarseness or Ambiguity?

Assume as in DLR that preference is complete, transitive and suitably continuous, and relax Independence to Preference Convexity. DLR (p. 892) argue that IR is a weak requirement. Though there is room for disagreement on this view (see Section 2 above), this section explores the implications of assuming also IR.

An immediate implication is that the model defined thereby can be interpreted alternatively as the model of an agent who foresees a complete set of states but who finds them to be ambiguous (in terms of likelihoods). As pointed out above, such an agent satisfies Preference Convexity. She also satisfies IR: if she foresees the complete state space, and anticipates choosing out of any given menu x conditionally on seeing the true complete state, then there is no reason for her to value randomization ex post, if also preference over lotteries (in the absence of other background uncertainty) conforms to vNM utility. Thus ex ante she would be indifferent between x and its convex hull, and so would satisfy IR.

We proceed to derive a representation result that will highlight the indistinguishability between coarseness and ambiguity. To do so, we adopt two additional axioms that are admittedly "excess baggage" but are arguably mild - they express ex ante certainty about the payoffs to specific alternatives b_* and b^* in B and certainty also that they will be worst and best lotteries respectively ex post.

Thus fix two alternatives b_* and b^* in B. Any ex post vNM preference over lotteries that ranks b_* worst and b^* best in B can be identified with a unique vector u in N,

$$N = \left\{u \in [0,1]^B : \, u\left(b_*\right) = 0, \, u\left(b^*\right) = 1\right\}. \tag{3.3}$$

With this in mind, define a dominance relation on lotteries by:[7]

$$\beta' \geq_D \beta \text{ if } \beta' \cdot u \geq \beta \cdot u \text{ for all } u \text{ in } N.$$

If $\beta' \geq_D \beta$, then an expected utility maximizer who ranks b_* and b^* as the worst and best alternatives respectively would never choose β alone from the feasible set $\{\beta', \beta\}$. Extend the dominance relation to menus by saying that x' dominates x, written $x' \geq_D x$, if for every β in x there exists β' in x' such that $\beta' \geq_D \beta$. If $x' \geq_D x$ and if the agent is certain ex ante that ex post (her preferences will conform to vNM theory and that) she will rank b_* and b^* as worst and best respectively, then she can be certain of doing as well choosing out of x' as out of $x' \cup x$. This explains the next axiom.

Worst-Best: For all menus x' and x, if $x' \geq_D x$ then $x' \cup x \sim x'$; and $\{b^*\} \succ \{b_*\}$.

Though certain that b_* and b^* will be worst and best ex post, the agent may nevertheless be uncertain about the cardinal payoff to one or both of these; moreover, cardinal payoffs are important when the agent evaluates menus ex ante and must weigh payoffs across all possible contingencies. We assume that, in fact,

[7] When $B = \{b_*, b, b^*\}$ consists of only three alternatives, then every u in N^* ranks $b_* \preceq b \preceq b^*$ and $\geq_D$ coincides with the first-order-stochastic dominance relation over lotteries induced by this ranking of alternatives.

expected cardinal payoffs to both b_* and b^* are certain ex ante. To express this certainty, let

$$C = \left\{ \beta_p = (b^*, p; b_*, 1-p) : 0 \le p \le 1 \right\} \subset \Delta(B).$$

If the cardinal payoffs to b_* and b^* are certain (constant across all states), then so are the payoffs to all lotteries in C (for any p in the unit interval and u in N, $u\left(\beta_p\right) = p$). Therefore, mixing with such lotteries provides no hedging gains, which suggests that the invariance required by Independence should be satisfied for such mixtures. This explains:

Certainty Independence: For all menus x' and x, lotteries β_p in C, and for all $0 < \alpha < 1$,

$$x' \succeq x \iff \alpha x' + (1-\alpha)\{\beta_p\} \succeq \alpha x + (1-\alpha)\{\beta_p\}.$$

Our last condition is a continuity axiom à la Herstein and Milnor [7].

Continuity: For all menus x, the sets $\left\{p \in [0,1] : \beta_p \succeq x\right\}$ and $\left\{p \in [0,1] : \beta_p \preceq x\right\}$ are closed.

Denote by $\mathcal{X}^c$ the set of all convex menus. Then we have:

Theorem 3.1. *Preference $\succeq$ on $\mathcal{X}$ satisfies Completeness, Transitivity, Continuity, Monotonicity, Preference Convexity, Worst-Best and Certainty Independence only if it admits a representation by $\mathcal{W}^{MP} : \mathcal{X}^c \to \mathbb{R}$ of the form:*

$$\mathcal{W}^{MP}(x) = min_{\pi \in \Pi} \int \max_{\beta \in x} u(\beta)\, d\pi(u), \tag{3.4}$$

where Π is a convex and weak-compact set of Borel probability measures on N. Moreover, $\succeq$ satisfies all the above axioms and Indifference to Randomization if and only if the representation (3.4) holds for all $x \in \mathcal{X}$.*

In either case, there exists Π that is maximal amongst sets satisfying the above and also the condition: if $x \in \mathcal{X}^c$ is such that, for all $x' \in \mathcal{X}^c$ and all $0 \le \alpha \le 1$,

$$x' \sim x \Longrightarrow \alpha x + (1-\alpha) x' \sim x, \tag{3.5}$$

then, for all π in Π,

$$\mathcal{W}^{MP}(x) = \int \max_{\beta \in x} u(\beta)\, d\pi(u). \tag{3.6}$$

Given the focus of this Section, consider the IR case, where (3.6) holds for all menus and condition (3.5) can be equivalently stated in terms of all menus x and x', not just convex ones.[8] Any menu x consisting solely of lotteries over b_* and b^* satisfies (3.5) - this is a consequence of Certainty Independence. The intuition, given above, is that the agent is certain about the ex post evaluation of such utility-constant lotteries, which renders them useless as a hedge against the uncertainty associated with other menus. More generally, there may also be other menus that cannot serve as hedges and these are captured by condition (3.5). All measures in Π agree on such menus x, in that they imply the same maximum expected utility for x - this is specified by (3.6). As a consequence, if preference satisfies Independence, then every menu satisfies (3.5) and the noted agreement of measures in Π applies to every menu x, which yields (a slight variant of) the DLR representation result for the non-negative additive EU representation (2.1).

Theorem 3.1 is a variation of the multiple-priors representation of Gilboa and Schmeidler [5], though not a trivial one. In fact, our axioms deliver a (superlinear and translation invariant) preference functional defined only on the convex cone of support functions, a meagre subset of the set of all continuous functions on N; in particular, the cone has an empty interior under the supnorm topology. For this reason we have to use different techniques than the ones used in [5], and we exploit the notion of niveloid developed in [10]. The smallness of the domain on which the preference functional is defined results in the non-uniqueness of the set Π of the representation (3.4), the domain not being big enough to pin down a single set of priors, but only a maximal one.

3.3. A Second Model

Section 2 argued that IR is not intuitive if ex ante coarseness is expected to persist. Here we describe an alternative axiom for this case, and then a functional form for utility that seems to capture this conception of coarseness.

Suppose that, contrary to Independence, the lotteries β' and β are such that

$$\{\alpha\beta' + (1-\alpha)\,\beta\} \succ \{\beta'\} \succeq \{\beta\}, \tag{3.7}$$

for a specific α. Suppose further that ex ante the agent has only one contingency

[8] To see the equivalence, recall that

$$co\,(\alpha x + (1-\alpha)\,x') = \alpha co\,(x) + (1-\alpha)\,co\,(x')\,, \qquad \forall \alpha \in [0,1]\,.$$

in mind. Then presumably the indicated value of randomization reveals that she views the contingency as coarse - there is no scope for ambiguity about likelihoods because there is only one contingency. Since the individual anticipates no further information before making her choice out of menus, it follows that she anticipates strictly preferring $\alpha\beta' + (1-\alpha)\beta$ to either of the component lotteries also ex post. Thus she would exhibit the ranking

$$co(\{\beta',\beta\}) \succ \{\beta',\beta\}. \tag{3.8}$$

However, such a sharp connection between (violations of) Independence and IR is not to be expected more generally when the agent conceives of several possible contingencies. For example, it could be that the mixture $\alpha\beta' + (1-\alpha)\beta$ lies strictly between β' and β in preference ex post for every contingency with β' being best for some contingencies and β best for others. The mixed menu $\{\alpha\beta' + (1-\alpha)\beta\}$ might still be best ex ante, as in (3.7), because it yields a higher average payoffs across all contingencies. In this case, neither $\alpha\beta' + (1-\alpha)\beta$, nor any other interior mixture, need be strictly better ex post conditionally on any specific contingency, even if all contingencies are coarse. Hence there need not be strict preference for $co(\{\beta',\beta\})$.

Nevertheless, if ambiguity about likelihoods is not the cause, then violation of Independence in (3.7) reveals the presence of coarse contingencies. Hence randomization should be of value for *some* lotteries, even if not for β' and β. Further, similar intuition applies if the violation of Independence occurs when any (not necessarily singleton) menus are mixed. Thus we suggest the following weaker connection:

Coarseness: For any convex menus x' and x, if $\alpha x' + (1-\alpha)x \succ x \sim x'$ for some $0<\alpha<1$, then $co(\{\gamma',\gamma\}) \succ \{\gamma',\gamma\}$ for some $\gamma,\gamma' \in \Delta(B)$.

A Functional Form: Denote by $\mathcal{K}(N)$ the set of closed subsets of N; U is a generic element. (Endow $\mathcal{K}(N)$ with the Hausdorff metric, which renders it compact metric, and with the corresponding Borel σ-algebra $\sigma(N)$.) Define the utility of a menu by

$$\mathcal{W}(x) = \int max_{\beta\in x}\, min_{u\in U}\, u(\beta)\; d\mu(U)\,, \text{ for all } x \in \mathcal{X}, \tag{3.9}$$

where $\mu : \sigma(N) \rightarrow [0,1]$ is a Borel probability measure. For interpretation, note first that when μ has support on singleton sets $U = \{u\}$, then one obtains the DLR form (2.1). More generally, the functional form suggests the following interpretation: the agent foresees coarse contingencies represented by sets U in the support of μ. Since she does not think explicitly in terms of finer contingencies, she anticipates choosing out of the menu x conditionally on each of these possible coarse contingencies. But she is aware ex ante that each set U leaves out some relevant missing details. Further, she anticipates that she will continue to be aware of the presence of missing details also given the realization of any particular U. This anticipation is captured by the fact that for any lottery β, its anticipated utility conditional on U is given by $min_{u \in U}\, u(\beta)$. (In particular, the ex post utility function over lotteries, given by $\beta \longmapsto min_{u \in U}\, u(\beta)$, is not a vNM functional.) Thus, given U, it is anticipated that a lottery will be chosen from x so as to solve $max_{\beta \in x}\, min_{u \in U}\, u(\beta)$. This leads to the evaluation of x shown in (3.9).

Implicit in the interpretation of the sets U in the support of μ as being coarse contingencies, is the assumption that the support is meaningful, that is, unique given preference. Next we show that such uniqueness obtains in a suitable sense, at least for measures having finite support.

Evidently,

$$\min_{u \in U'} \beta \cdot u = \min_{u \in U} \beta \cdot u \text{ for all } \beta \text{ in } \Delta(B), \tag{3.10}$$

if U' and U have identical convex hulls. Define the comprehensive hull of U by

$$comp(U) \equiv \{u' \in N : \ u' \geq u \text{ for some } u \in U\}.$$

Then (3.10) obtains also if U' and U have identical comprehensive hulls. Thus uniqueness of the support can be expected only if one restricts attention to suitably normalized sets U.

Say that $U \subset N$ is *comprehensive* if $U = comp(U)$ and denote by $\mathcal{K}^c(N)$ the set of closed, convex and comprehensive subsets of N. Say that the probability measure μ on $\mathcal{K}^c(N)$ *represents* $\succeq$ if the latter has utility function given by (3.9). Then we have:

Theorem 3.2. *A preference $\succeq$ can be represented by at most a single finitely supported probability measure μ.*

A distinction: For perspective, consider the alternative functional form obtained by reversing the order of the *max* and *min* appearing inside the integral in (3.9), that is, consider

$$\mathcal{W}^{rev}(x) = \int min_{u \in U}\ max_{\beta \in x}\ u(\beta)\ d\mu(U)\,,\ \text{ for all } x \in \mathcal{X}. \tag{3.11}$$

In general, $\mathcal{W}^{rev}(\cdot)$ is ordinally distinct from $\mathcal{W}(\cdot)$. This is suggested by the fact that the minimax theorem justifying such reversals of order requires that both sets x and U be convex, and neither convexity need hold. More formal confirmation follows shortly.

The obvious interpretation is that the agent conceives ex ante of the finely detailed states u and anticipates choosing out of x after seeing which u is realized. Thus, she anticipates maximizing $u(\cdot)$ ex post. At the same time, these states are ambiguous for her - she is not sufficiently confident of their likelihood to hold probabilistic beliefs, and this leads to the integrand appearing in (3.11). To support the connection to ambiguity, note that preference represented by $\mathcal{W}^{rev}(\cdot)$ satisfies all the assumptions of Theorem 3.1 and so it is a special case of that model. In Appendix B we describe a "concrete" set Π such that

$$\mathcal{W}^{rev}(x) = min_{\pi \in \Pi} \int \max_{\beta \in x} u(\beta)\ d\pi(u)\,. \tag{3.12}$$

The axioms IR and Coarseness distinguish (3.9) from our first model (3.4); both models satisfy Completeness, Transitivity, Continuity, Monotonicity, Worst-Best and Certainty Independence.

Theorem 3.3. *(i) Preference represented by $\mathcal{W}(\cdot)$ defined in (3.9) satisfies Preference Convexity and Coarseness. It satisfies IR iff it can be represented by a utility function having the DLR form (2.1) and (3.3).*

(ii) Preference represented by $\mathcal{W}^{MP}(\cdot)$ defined in (3.4) satisfies Preference Convexity and IR. It satisfies Coarseness iff it can be represented by a utility function having the DLR form (2.1) and (3.3).

In spite of the distinction provided by the theorem, the functional form (3.9) nevertheless admits an interpretation in terms of ambiguity, albeit somewhat different from that provided for the previous model.[9] The functional form might describe an agent who conceives ex ante of the complete states in N, but does not expect to see the true state ex post. Rather, she expects only a "signal" U to

[9]The alternative interpretation can be shown to be consistent also with the intuition for the axioms Preference Convexity and Coarseness.

be realized ex post. There is no prior ambiguity about the likelihoods of signals; however, each signal is "ambiguous" - it will inform the agent that the true state u lies in U, but leave her completely ignorant otherwise. This is a special case of recursive multiple-priors utility studied by Epstein and Schneider [3] (though the information structure is exogenous there and subjective here). In the omelet example, the alternative story is that the man conceives ex ante of all gradations of "good" for the 6th egg, but anticipates that when he breaks that egg he will be able to see only whether or not it is totally rotten and not whether it is superlative or stale.

Dempster-Shafer-style models: Finally, we relate the model (3.9) to the Dempster-Shafer-style models of Mukerji [12] and Ghirardato [4] mentioned in the introduction. They suppose that while there exists a Savage-style state space S, the agent does not conceive of all the complete states in S and has coarse perceptions. These are modeled through an auxiliary epistemic state space Ω and a correspondence Γ from Ω into S. (See the figure below.) There is a probability measure p representing beliefs on Ω.

$$\begin{tikzcd} (\Omega, p) \arrow[r, rightsquigarrow, "\Gamma"] \arrow[dr, "\widehat{f}"'] & (S, \nu) \arrow[d, "f"] \\ & X \end{tikzcd} \tag{3.13}$$

Unlike a Savage agent who would view each physical action as an act from S to the outcome set X, and who would evaluate it via its expected utility (using a probability measure on S), the present agent views each action as a (possibly multi-valued) act on Ω.

Ghirardato assumes that each $\widehat{f}$ is multi-valued where the nonsingleton nature of $\widehat{f}(\omega)$ reflects her awareness that ω is a coarse contingency. Its utility is given by

$$V^G\left(\widehat{f}\right) = \int_\Omega \left(\min_{x \in \widehat{f}(\omega)} u(x) \right) dp.$$

In this formulation, both Ω and the acts $\widehat{f}$ are taken to be objective and hence observable to the analyst. One can view our model (3.9) as one possible way to render them subjective: take $X = \mathbb{R}^1$, $\Omega = Supp(\mu) \subset \mathcal{K}(N)$, and $p = \mu$, where μ is the measure appearing in our representation; and identify each lottery β in

$\Delta(B)$ with the multi-valued act $\widehat{\beta}$,

$$\widehat{\beta} : U \longrightarrow \{u(\beta) : u \in U\}.$$

Then

$$V^G\left(\widehat{\beta}\right) = \mathcal{W}(\{\beta\}).$$

Turn to the rest of the triangle (3.13). It is commutative if $\widehat{f}(\omega) = f(\Gamma(\omega))$. This is satisfied in our model if we take $S = N$ and $\Gamma(U) = U \subset N$.

Finally, we can write

$$V^G\left(\widehat{f}\right) = \int_S u(f)\ d\nu(s),$$

where ν is the non-additive measure or capacity

$$\nu(Y) = \mu(\{\omega : \Gamma(\omega) \subset Y\}),$$

and the integral on the right is in the sense of Choquet (see Schmeidler [15]).[10] Since Schmeidler's Choquet expected utility model was devised in order to accommodate ambiguity, this demonstrates once again the close connection between coarse perceptions and ambiguity.

Though there are differences in detail, similar remarks apply to Mukerji [12]; in particular, our model can be viewed as a way to endogenize the state spaces Ω and S, as well as the correspondence Γ, all of which are taken as primitives by Mukerji.

A. Appendix: Proof of Theorem 3.2

Theorem 3.2 is a consequence of a more general uniqueness result on maxitive functions defined on mixture lattices, which may be useful in other contexts. We therefore first prove such a result (Theorem A.5), and then show how Theorem 3.2 follows.

A.1. Maxitive Functions

Say that $\mathcal{M}$ is a *mixture lattice* if it is a mixture space endowed with a binary relation that makes it a lattice. Given $x, y \in \mathcal{W}$, denote by $[x, y]$ the segment $\{\alpha x + (1-\alpha) y : \alpha \in (0,1)\}$.

[10] More precisely, it corresponds to the special case where ν is a belief function.

The *domain of additivity* $\mathcal{E}(f)$ of a function $f : \mathcal{M} \to \mathbb{R}$ is the set

$$\{y \in \mathcal{M} : f(\alpha x + (1-\alpha) y) = \alpha f(x) + (1-\alpha) f(y) \ \ \forall x \in \mathcal{M}, \forall \alpha \in [0,1]\}.$$

Given $x_*, x^* \in \mathcal{M}$, the segment $[x_*, x^*]$ is a *diagonal* for f if $[x_*, x^*] \subseteq \mathcal{E}(f)$ and $f(x_*) \leq f(x) \leq f(x^*)$ for all $x \in \mathcal{M}$. Observe that this implies that, for each $x \in \mathcal{M}$,

$$f(\alpha x + (1-\alpha) y) = \alpha f(x) + (1-\alpha) f(y), \quad \forall y \in [x_*, x^*], \forall \alpha \in [0,1].$$

Two functions $f, g : \mathcal{M} \to \mathbb{R}$ are *comonotonic* if

$$[f(x) - f(y)][g(x) - g(y)] \geq 0, \quad \forall x, y \in \mathcal{M}.$$

A function $f : \mathcal{M} \to \mathbb{R}$ is *maxitive* if $f(x \vee y) = \max\{f(x), f(y)\}$ for all $x, y \in \mathcal{M}$.

Lemma A.1. *Let $\{f_i\}_{i=1}^n$ and $\{g_i\}_{j=1}^m$ be maxitive functions such that, for some $\{\alpha_i\}_{i=1}^n \subseteq \mathbb{R}_+$ and $\{\beta_j\}_{j=1}^m \subseteq \mathbb{R}_+$,*

$$\sum_{i=1}^n \alpha_i f_i(x) = \sum_{j=1}^m \beta_j g_j(x), \qquad \forall x \in \mathcal{M}. \tag{A.1}$$

Then there is a non-decreasing function $W = (W_1, ..., W_n) : \Gamma \subseteq \mathbb{R}^m \to \mathbb{R}^n$ such that, for each $x \in \mathcal{X}$,

$$(f_1(x), ..., f_n(x)) = W(g_1(x), ..., g_m(x)),$$

where $\Gamma = \{(g_1(x), ..., g_m(x)) : x \in \mathcal{M}\}$.

Proof. Suppose x and y are such that $g_j(x) = g_j(y)$ for each $j = 1, ..., m$. Then $g_j(x) = \max\{g_j(x), g_j(y)\} = g_j(x \vee y)$ for each j.

By (A.1), we can then write:

$$\sum_{i=1}^n \alpha_i f_i(x) = \sum_{j=1}^m \beta_j g_j(x) = \sum_{j=1}^m \beta_j g_j(x \vee y) = \sum_{i=1}^n \alpha_i f_i(x \vee y).$$

As $f_i(x) \leq f_i(x \vee y)$ for each i, conclude that $f_i(x) = f_i(x \vee y)$ for each i. A similar argument shows that $f_i(y) = f_i(x \vee y)$ for each i, and so $f_i(x) = f_i(y)$ for each i, as desired.

To show that W is non-decreasing, suppose that x and y are such that $g_j(x) \geq g_j(y)$ for each $j = 1, ..., m$. Then, $g_j(x) = g_j(x \vee y)$ for each j, and so (A.1) implies:

$$\sum_{i=1}^{n} \alpha_i f_i(x) = \sum_{j=1}^{m} \beta_j g_j(x) = \sum_{j=1}^{m} \beta_j g_j(x \vee y) = \sum_{i=1}^{n} \alpha_i f_i(x \vee y).$$

Hence $f_i(x) = f_i(x \vee y)$ for each i , so that $f_i(x) = f_i(x \vee y) \geq f_i(y)$. ∎

Lemma A.2. *Suppose that some W_i in Lemma A.1 is strictly increasing. Then the functions $\{g_i\}_{j=1}^{m}$ are pairwise comonotonic.*

Proof. Wlog suppose W_1 is strictly increasing. Consider $\{x, y\} \in \mathcal{M}$ and suppose $f_1(x) \geq f_2(y)$. Then,

$$\begin{aligned} f_1(x) &= W_1(g_1(x), ..., g_m(x)), \\ f_1(x \vee y) &= W_1(g_1(x \vee y), ..., g_m(x \vee y)). \end{aligned}$$

As f_1 is maxitive,

$$W_1(g_1(x), ..., g_m(x)) = W_1(g_1(x \vee y), ..., g_m(x \vee y)).$$

As each g_j is maxitive, we have $g_j(x) \leq g_j(x \vee y)$ for each $j = 1, ..., m$. Since W_1 is strictly increasing, we then have $g_j(x) = g_j(x \vee y)$ for each $j = 1, ..., m$, which in turn implies the desired result. ∎

Corollary A.3. *Let $\{g_i\}_{j=1}^{m}$ be maxitive functions. Then the following conditions are equivalent:*

(i) there is a maxitive function f such that, for each $x \in \mathcal{M}$,

$$f(x) = \sum_{j=1}^{m} \beta_j g_j(x),$$

where $\beta_j \geq 0$ for each j;

(ii) the sum function $\sum_{j=1}^{m} g_j$ is maxitive;

(iii) the functions $\{g_i\}_{j=1}^{m}$ are pairwise comonotonic.

Proof. (i) implies (iii): Here $W(x) = \sum_{j=1}^{m} x_j$ for each $x \in \Gamma$. As W is strictly increasing, by Lemma A.2 the functions $\{g_i\}_{j=1}^{m}$ are pairwise comonotonic, and so (iii) holds.

(iii) implies (ii): Given $x, y \in \mathcal{M}$, assume $\sum_{j=1}^{m} g_j(y) \leq \sum_{j=1}^{m} g_j(x)$. Then there is at least one $j = 1, ..., m$, say $j = 1$, such that $g_1(y) \leq g_1(x)$. By (iii), this implies $g_j(y) \leq g_j(x)$ for all $j = 2, ..., m$, and so $g_j(x) = g_j(x \vee y)$ for all $j = 1, ..., m$. Hence, $\sum_{j=1}^{m} g_j(x \vee y) = \sum_{j=1}^{m} g_j(x)$, and so $\sum_{j=1}^{m} g_j$ is maxitive.

As (ii) trivially implies (i), this completes the proof. ■

Lemma A.4. *Let $f, g : \mathcal{M} \to \mathbb{R}$ be two functions sharing the same diagonal $[x_*, x^*]$, with $f(x_*) = g(x_*)$ and $f(x^*) = g(x^*)$. Then f and g are comonotonic if and only if they are equal.*

Proof. Suppose f and g are comonotonic. If $f \neq g$, there exists x such that, say, $f(x) > g(x)$. Hence, there exists $p \in \mathbb{R}$ such that $f(x) > p > g(x)$, and there exists $\alpha \in [0, 1]$ such that

$$f(\alpha x_* + (1 - \alpha) x^*) = g(\alpha x_* + (1 - \alpha) x^*) = p.$$

Thus

$$f(x) > f(\alpha x_* + (1 - \alpha) x^*) = g(\alpha x_* + (1 - \alpha) x^*) > g(x),$$

which contradicts comonotonicity. Conclude that $f = g$. ■

A.2. A Uniqueness Result

Theorem A.5. *Let $\{f_i\}_{i=1}^{n}$ and $\{g_j\}_{j=1}^{n}$ be maxitive real-valued functions defined on a mixture lattice $\mathcal{M}$ sharing the same diagonal $[x_*, x^*]$, with $f_i(x_*) = g_j(x_*) = k_*$ and $f_i(x^*) = g_j(x^*) = k^*$ for all i and j. Suppose that, for some $\{\alpha_i\}_{i=1}^{n} \subseteq \mathbb{R}_+$ and $\{\beta_j\}_{j=1}^{m} \subseteq \mathbb{R}_+$,*

$$\sum_{i=1}^{n} \alpha_i f_i(x) = \sum_{i=1}^{n} \alpha_i f_i(y) \Longrightarrow \sum_{j=1}^{m} \beta_j g_j(x) = \sum_{j=1}^{m} \beta_j g_j(y), \quad \forall x, y \in \mathcal{M}. \quad \text{(A.2)}$$

Then

$$\{f_1, ..., f_n\} = \{g_1, ..., g_m\},$$

and $\alpha_i = \beta_j$ whenever $f_i = g_j$.

Proof. It is enough to prove

$$\{\alpha_1 f_1, ..., \alpha_n f_n\} = \{\beta_1 g_1, ..., \beta_m g_m\}. \tag{A.3}$$

In fact, if $\alpha_i f_i = \beta_j g_j$, then $\alpha_i f_i(x) = \beta_j g_j(x)$ for all $x \in [x_*, x^*]$, and so $\alpha_i = \beta_j$.

By (A.2), there exists $\Phi : [k_*, k^*] \to \mathbb{R}$ such that

$$\sum_{i=1}^{n} \alpha_i f_i(x) = \Phi\left(\sum_{j=1}^{m} \beta_j g_j(x)\right), \qquad \forall x \in \mathcal{M}. \tag{A.4}$$

Given $p \in [k_*, k^*]$, let $x_p \in [x_*, x^*]$ be such that $f_i(x_p) = g_j(x_p) = p$ for all i and j. By considering for each $p \in [0,1]$ the corresponding element $x_p \in \mathcal{M}$, we deduce that $\Phi(p) = p$ for all $p \in [0,1]$. Therefore, (A.4) reduces to:

$$\sum_{i=1}^{n} \alpha_i f_i(x) = \sum_{j=1}^{m} \beta_j g_j(x), \qquad \forall x \in \mathcal{M}, \tag{A.5}$$

By Lemma A.1, there exists a non-decreasing function $W = (W_1, ..., W_n) : [k_*, k^*]^m \to [k_*, k^*]^n$ such that

$$(f_1(x), ..., f_n(x)) = W(g_1(x), ..., g_m(x)), \qquad \forall x \in \mathcal{X}. \tag{A.6}$$

The proof is now divided into a few steps.

Step 1. We show that (A.3) holds when either $n = 1$ or $m = 1$. Consider the former case; the other may be treated similarly. Hence (A.5) becomes

$$f(x) = \sum_{j=1}^{m} \beta_j g_j(x),$$

and so Corollary A.3 implies that the functions $\{g_i\}_{j=1}^{m}$ are pairwise comonotonic. By Lemma (A.4), they are all equal, and this proves (A.3).

Step 2. Given (n, m), suppose (A.3) holds for all (n', m') with $n' < n$ and $m' < m$. In view of Step 1, to complete the proof it is enough to show that (A.3) holds for (n, m).

If all functions in the collection $\{f_i\}_{i=1}^{n}$ are equal, then (A.3) holds since we are back to Step 1. Suppose, therefore, that at least two such functions are distinct, say f_1 and f_2. Then, there is $\overline{x} \in \mathcal{X}$ and $p \in (k_*, k^*)$ such that, say,

$f_1(\overline{x}) > p > f_2(\overline{x})$. Wlog assume (if needed, replace $\overline{x}$ by $\alpha\overline{x} + (1-\alpha)x^*$ for α close to 1),

$$f_i(\overline{x}) \neq p,\;\; g_j(\overline{x}) \neq p \;\text{ for all } i \text{ and } j.$$

We now show that this implies

$$\Sigma_{i\in I}\alpha_i f_i(x) = \Sigma_{j\in J}\beta_j g_j(x), \qquad \forall x \in \mathcal{X}, \tag{A.7}$$

where

$$I = \{i: \; f_i(\overline{x}) > p\} \text{ and } J = \{j : g_j(\overline{x}) > p\}.$$

By hypothesis, 1 lies in I, and 2 lies in its complement; thus $\emptyset \neq I \neq \{1,...,n\}$. By (A.6), $\emptyset \neq J \neq \{1,...,n'\}$. Therefore, (A.7) has fewer elements in the sums than does (A.5), and so the proof of the theorem is complete once we prove that (A.7) holds.

Take the menu $\overline{x} \vee x_p$ in (A.5), to deduce that

$$\Sigma_{i\in I}\alpha_i f_i(\overline{x}) + p\Sigma_{i\notin I}\alpha_i = \Sigma_{j\in J}\beta g_j(\overline{x}) + p\Sigma_{j\notin J}\beta_j. \tag{A.8}$$

Replace $\overline{x}$ by $\alpha\overline{x}+(1-\alpha)x^*$ with $0<\alpha<1$ small enough so that all of the above strict inequalities are preserved. Then

$$\begin{aligned} & \alpha\Sigma_{i\in I}\alpha_i f_i(\overline{x}) + (1-\alpha)\Sigma_{i\in I}\alpha_i + p\Sigma_{i\notin I}\alpha_i \\ = \; & \alpha\Sigma_{j\in J}\beta g_j(\overline{x}) + (1-\alpha)\Sigma_{j\in J}\beta_j + p\Sigma_{j\notin J}\beta_j, \end{aligned}$$

and therefore,

$$\Sigma_{i\in I}\alpha_i + p\Sigma_{i\notin I}\alpha_i = \Sigma_{j\in J}\beta_j + p\Sigma_{j\notin J}\beta_j \tag{A.9}$$

For small changes in p, all of above is unchanged. Thus (A.9) holds for an open set of p's and we can conclude that

$$\begin{aligned} M_1 &= \Sigma_{i\in I}\alpha_i = \Sigma_{j\in J}\beta_j = M_1' > 0, & \text{(A.10)} \\ M_2 &= \Sigma_{i\notin I}\alpha_i = \Sigma_{j\notin J}\beta_j = M_2' > 0. & \text{(A.11)} \end{aligned}$$

It is convenient to define

$$m_i = \begin{cases} \alpha_i/M_1 & i \in I \\ \alpha_i/M_2 & i \notin I \end{cases}, \quad m_j' = \begin{cases} \beta_j/M_1' & j \in J \\ \beta_j/M_2' & j \notin J \end{cases},$$

and, for all $x \in \mathcal{X}$,

$$H_1(x) = \Sigma_{i\in I} m_i f_i(x),\; H_2(x) = \Sigma_{i\notin I} m_i f_i(x),$$

$$H_1'(x) = \Sigma_{j\in J} m_j' g_j(x) \text{ and } H_2'(x) = \Sigma_{j\notin J} m_j' g_j(x).$$

Then, by (A.8), (A.10), and (A.11),

$$H_1(\overline{x}) = H_1'(\overline{x}). \tag{A.12}$$

In order to show that (A.7) holds we have to prove that $H_1(x) = H_1'(x)$ for all $x \in \mathcal{X}$. Consider two cases:

Case 1: There exists $x \in \mathcal{M}$ such that, for each $n \geq 1$,

$$H_1\left(\frac{1}{n}x + \left(1-\frac{1}{n}\right)x_p\right) \neq H_1'\left(\frac{1}{n}x + \left(1-\frac{1}{n}\right)x_p\right).$$

Set $x_n = \frac{1}{n}x + \left(1-\frac{1}{n}\right)x_p$, for all n. Then

$$\begin{aligned} f_i(x_n) &= \frac{1}{n}f_i(x) + \left(1-\frac{1}{n}\right)p, \quad \forall i = 1, ..., n, \\ g_j(x_n) &= \frac{1}{n}g_j(x) + \left(1-\frac{1}{n}\right)p, \quad \forall j = 1, ..., m, \end{aligned}$$

and so, for n large enough,

$$\begin{aligned} f_i(\overline{x}) &> f_i(x_n) \ \forall i \in I,\ f_i(\overline{x}) < f_i(x_n) \ \forall i \notin I, \\ g_j(\overline{x}) &> g_j(x_n) \ \forall j \in J \text{ and } g_j(\overline{x}) < g_j(x_n) \ \forall j \notin J. \end{aligned}$$

If we take $\overline{x} \vee x_n$ in (A.5), for all n large enough, then

$$M_1H_1(\overline{x}) + M_2H_2(x_n) = M_1'H_1'(\overline{x}) + M_2'H_2'(x_n).$$

Since $H_1(\overline{x}) = H_1'(\overline{x})$, (A.10) and (A.11) imply that $H_2(x_n) = H_2'(x_n)$. But (A.5) applied to x_n implies that

$$M_1H_1(x_n) + M_2H_2(x_n) = M_1'H_1'(x_n) + M_2'H_2'(x_n).$$

Therefore, $H_1(x_n) = H_1'(x_n)$, a contradiction.

Case 2: For each $x \in \mathcal{M}$, there exists $n \geq 1$ such that $H_1(x_n) = H_1'(x_n)$. Hence

$$\frac{1}{n}H_1(x) + \left(1-\frac{1}{n}\right)H_1(x_p) = \frac{1}{n}H_1'(x) + \left(1-\frac{1}{n}\right)H_1'(x_p),$$

and so $H_1(x) = H_1'(x)$, as desired. ■

A.3. Proof of Theorem 3.2

For each compact and convex $U \subset N$, define $h_U : \Delta(B) \to \mathbb{R}$ by

$$h_U(\beta) = \min_{u \in U} \ \beta \cdot u, \tag{A.13}$$

and define $E_U : \mathcal{X} \to \mathbb{R}$ by $E_U(x) = \max_{\beta \in x} h_U(\beta)$. The function $E_U : \mathcal{X} \to \mathbb{R}$ is maxitive over the mixture lattice $\mathcal{X}$.

Lemma A.6. *For any (closed and convex) comprehensive subsets U and U',*

$$E_U = E_{U'} \iff h_U = h_{U'} \iff U = U'.$$

Proof. The only non trivial implication is that $h_U = h_{U'} \Rightarrow U = U'$. Suppose $\exists u^* \in U' \backslash U$. It is enough to show that any $u^* \in N \backslash U$ can be separated from U by some $\beta \in \Delta(B)$. Identify any u in N with a point $v \in \mathbb{R}^{B-2}$, $v_b = u_b$, $b \neq b_*, b^*$. Thus identify U with the set $V \subset \mathbb{R}^{B-2}$. Now take the comprehensive hull in $\mathbb{R}^{B-2}$, that is, define

$$\widehat{V} = \{v' \in \mathbb{R}^{B-2} : v' \geq v, \, v \in V\}.$$

Then v^*, the image of u^*, does not lie in $\widehat{V}$, and the latter is convex. Therefore, they can be separated by some $0 \neq p \in \mathbb{R}^{B-2}$:

$$p \cdot v^* < min_{v \in \widehat{V}} \, p \cdot v.$$

In the usual way, the comprehensiveness (and unboundedness) above of $\widehat{V}$ implies that $p \geq 0$. Moreover, $\Sigma_{b \neq b_*, b^*} \, p_b u_b^* < min_{u \in U} \Sigma_{b \neq b_*, b^*} \, p_b u_b$. Define p_{b^*} and p_{b^*} arbitrarily non-negative. Then, $p \cdot u^* < min_{u \in U} \, p \cdot u$, because terms for b_* and b^* cancel. Finally, normalize p to obtain the desired separating lottery β. ■

To prove Theorem 3.2, suppose that μ and μ' have finite support and that they both represent $\succeq$. Then,

$$\Sigma_{i=1}^{n} \mu_i \max_{\beta \in x} h_{U_i}(\beta) = \Phi\left(\Sigma_{j=1}^{m} \mu'_j \max_{\beta \in x} h_{U'_j}(\beta)\right), \qquad \forall x \in \mathcal{X}, \tag{A.14}$$

for some strictly increasing $\Phi : [0,1] \to \mathbb{R}$, where $\mu_i > 0$ and $\mu'_j > 0$ for all i and j. Observe that the domain of Φ is $[0,1]$ because, for each $p \in [0,1]$, we have $u \cdot \beta_p = p$ for all $u \in N$.

We can write (A.14) as

$$\Sigma_{i=1}^{n}\mu_i E_{U_i}(x) = \Phi\left(\Sigma_{j=1}^{m}\mu'_j E_{U'_j}(x)\right), \qquad \forall x \in \mathcal{X},$$

and so Theorem A.5 implies that

$$\{\mu_1 E_{U_1}, ..., \mu_n E_{U_n}\} = \left\{\mu'_1 E_{U'_1}, ..., \mu'_m E_{U'_m}\right\}.$$

By Lemma A.6,

$$\{U_1, ..., U_n\} = \{U'_1, ..., U'_{n'}\} \text{ and } \mu = \mu' \text{ as measures.}$$

B. Appendix: Proof of Equation (3.12)

To derive (3.12), let

$$\mathcal{U} = supp\mu, \text{ and}$$

$$\sigma_x(u) = \max_{\beta \in x} u(\beta), \text{ for all } u \text{ in } N.$$

Say that the collection $\{p_U(\cdot) : U \in \mathcal{U}\}$ is a *$\mathcal{U}$-conditional probability system* ($\mathcal{U}$-cps) if $\mu - a.s.$: (i) $p_U(\cdot) \in \Delta(U)$, and (ii) $A \longmapsto p_U(A)$ is suitably measurable. Define Π by

$$\Pi = \left\{\pi : \pi(\cdot) = \int p_U(\cdot)\, d\mu(U) \text{ for some } \mathcal{U}\text{-cps } \{p_U(\cdot)\}\right\}. \tag{B.1}$$

Then

$$min_{\pi\in\Pi} \int \sigma_x(u)\, d\pi(u) = min_{\{p_U\}} \int \left(\int \sigma_x(u)\, dp_U(u)\right) d\mu(U) =$$

$$\int \left(min_{p_U} \int \sigma_x(u)\, dp_U(u)\right) d\mu(U) = \int \left(min_{u\in U}\, \sigma_x(u)\right) d\mu(U) = \mathcal{W}^{rev}(x). \tag{B.2}$$

Roughly, the structure (B.1) for Π permits the minimum appearing outside the integral in (3.12) to be taken inside.

C. Appendix: Proof of Theorem 3.3

C.1. Preliminaries

Let C be a convex subset of some normed vector space. A function $h : C \to \mathbb{R}$ is quasi-convex if $h(t\beta' + (1-t)\beta'') \leq \max\{h(\beta'), h(\beta'')\}$ for each $\beta', \beta'' \in C$. It is quasi-concave if $-f$ is quasi-convex, and it is quasi-monotone if it is both quasi-convex and quasi-concave.

The following result is due to [6, p. 1559].

Lemma C.1. *Let $h : C \to \mathbb{R}$ be continuous. Then h is quasi-monotone if and only if the sets $\{\beta : h(\beta) = c\}$ are convex for all $c \in \mathbb{R}$.*

Suppose K is a compact set in some topological space, and let $h : C \to \mathbb{R}$ be given by

$$h(\beta) = \min_{y \in K} T(\beta, y),$$

where $T : C \times K \to \mathbb{R}$ is continuous on K and concave on C, i.e., for each $\beta', \beta'' \in C$,

$$T(t\beta' + (1-t)\beta'', y) \geq tT(\beta', y) + (1-t)T(\beta'', y), \quad \forall y \in K. \tag{C.1}$$

Lemma C.2. *The function $h : C \to \mathbb{R}$ is concave.*

Define

$$\Theta(\beta) = \arg\min{}_{y \in K} T(\beta, y), \quad \forall \beta \in C.$$

Say that h is affine on some convex subset $Q \subseteq C$ if $h(\lambda\beta_1 + (1-\lambda)\beta_2) = \lambda h(\beta_1) + (1-\lambda)h(\beta_2)$ for all $\lambda \in [0,1]$ and all $\beta_1, \beta_2 \in Q$.

Lemma C.3. *(i) For any finite collection $\{\beta_i\}_{i \in I} \subseteq C$,*

$$h\left(\sum_{i=1}^{n} \lambda_i \beta_i\right) = \sum_{i=1}^{n} \lambda_i h(\beta_i) \tag{C.2}$$

for some collection $\{\lambda_i\}_{i \in I}$ with $\lambda_i \in (0,1)$ and $\sum_{i \in I} \lambda_i = 1$, if and only if $\bigcap_{i \in I} \Theta(\beta_i) \neq \emptyset$. In this case, $\bigcap_{i \in I} \Theta(\beta_i) = \Theta\left(\sum_{i=1}^{n} \lambda_i \beta_i\right)$.

(ii) Let $Q \subset C$ be a convex set. Then h is affine on Q if and only if

$$\bigcap_{\beta \in Q} \Theta(\beta) \neq \emptyset.$$

(iii) Given $c \in \mathbb{R}$, a nonempty set $A \subseteq \{\beta \in C : h(\beta) = c\}$ is convex if and only if

$$\bigcap\nolimits_{\beta \in A} \Theta(\beta) \neq \emptyset.$$

Proof. In all points we prove the "only if", the converse being trivial. (i) Let $\widehat{y} \in \Theta\left(\sum_{i=1}^{n} \lambda_i \beta_i\right)$ and $\widehat{y}_i \in \Theta(\beta_i)$ for $i \in I$. By (C.1) and (C.2),

$$\sum_{i=1}^{n} \lambda_i T(\beta_i, \widehat{y}) \leq T\left(\sum_{i=1}^{n} \lambda_i \beta_i, \widehat{y}\right) = \sum_{i=1}^{n} \lambda_i T(\beta_i, \widehat{y}_i). \tag{C.3}$$

On the other hand, $\widehat{y}_i \in \Theta(\beta_i)$ implies:

$$T(\beta_i, \widehat{y}_i) \leq T(\beta_i, \widehat{y}), \qquad \forall i \in I.$$

Hence, by (C.3) we have

$$\sum_{i=1}^{n} \lambda_i T(\beta_i, \widehat{y}) = \sum_{i=1}^{n} \lambda_i T(\beta_i, \widehat{y}_i),$$

which in turn implies $T(\beta_i, \widehat{y}) = T(\beta_i, \widehat{y}_i)$ for each $i \in I$. This shows that $\Theta\left(\sum_{i=1}^{n} \lambda_i \beta_i\right) \subseteq \bigcap_{i \in I} \Theta(\beta_i)$. The converse inclusion is trivial, and we conclude that $\Theta\left(\sum_{i=1}^{n} \lambda_i \beta_i\right) = \bigcap_{i \in I} \Theta(\beta_i) \neq \emptyset$.

(ii) Suppose h is affine on Q. As Q is convex, this implies that $h\left(\sum_{i=1}^{n} \lambda_i \beta_i\right) = \sum_{i=1}^{n} \lambda_i h(\beta_i)$ for any finite collection $\{\beta_i\}_{i \in I} \subseteq Q$, and any $\{\lambda_i\}_{i \in I}$ with $\lambda_i \in [0,1]$ and $\sum_{i \in I} \lambda_i = 1$. By the previous point, $\bigcap_{i \in I} \Theta(\beta_i) \neq \emptyset$. As all sets $\Theta(\beta)$ are compact, the Finite Intersection Property implies that $\bigcap_{\beta \in Q} \Theta(\beta) \neq \emptyset$.

(iii) Let $\beta_1, \beta_2 \in A$. As A is convex,

$$h(\lambda\beta_1 + (1-\lambda)\beta_2) = \lambda h(\beta_1) + (1-\lambda) h(\beta_2), \qquad \forall \lambda \in [0,1].$$

By the previous point, $\bigcap_{\beta \in A} \Theta(\beta) \neq \emptyset$. ■

Lemma C.4. *Let $Q \subset C$ be a convex set and suppose h is continuous. Then h is affine on Q if and only if h is quasi-convex on Q and there exists $\gamma \in Q$ such that, for all $\lambda \in [0,1]$ and all $\beta \in Q$,*

$$h(\beta) \geq h(\gamma) \quad \text{and} \quad h(\lambda\beta + (1-\lambda)\gamma) = \lambda h(\beta) + (1-\lambda) h(\gamma).$$

In this case, there exists $y \in K$ such that

$$h(\beta) = T(\beta, y), \qquad \forall \beta \in Q. \tag{C.4}$$

Proof. For the "only if" part, just take $\gamma \in \arg\min_{\beta\in\Delta(B)} h(\beta)$. Consider now the "if" part. Set $\widetilde{Q} = \{\beta \in Q : h(\beta) > h(\gamma)\}$. Let $\beta_1, \beta_2 \in \widetilde{Q}$, and suppose $h(\beta_1) \geq h(\beta_2)$. If $h(\beta_1) = h(\beta_2)$, then

$$h(\lambda\beta_1 + (1-\lambda)\beta_2) = \lambda h(\beta_1) + (1-\lambda) h(\beta_2), \qquad \forall \lambda \in [0,1]$$

since h is quasi-monotone. Suppose $h(\beta_1) > h(\beta_2)$. There exists $\lambda \in (0,1)$ such that $h(\beta_2) = \lambda h(\beta_1) + (1-\lambda) h(\gamma) = h(\lambda\beta_1 + (1-\lambda)\gamma)$. By Lemma C.3(iii), there is $\widehat{y} \in \Theta(\beta_2)$ such that

$$\begin{aligned}
h(\lambda\beta_1 + (1-\lambda)\gamma) &= T(\lambda\beta_1 + (1-\lambda)\gamma, \widehat{y}) \\
&\geq \lambda T(\beta_1, \widehat{y}) + (1-\lambda) T(\gamma, \widehat{y}) \\
&\geq \lambda h(\beta_1) + (1-\lambda) h(\gamma) \\
&= h(\lambda\beta_1 + (1-\lambda)\gamma),
\end{aligned}$$

so that $T(\beta_1, \widehat{y}) = h(\beta_1)$. Hence, $\widehat{y} \in \Theta(\beta_1)$, and we conclude that $\Theta(\beta_1) \cap \Theta(\beta_2) \neq \emptyset$, i.e., $h(\lambda\beta_1 + (1-\lambda)\beta_2) = \lambda h(\beta_1) + (1-\lambda) h(\beta_2)$ for all $\lambda \in [0,1]$. The function h is therefore affine on $\widetilde{Q}$.

Let $\overline{\beta} \in Q$ be such that $h(\overline{\beta}) = h(\gamma)$. If $\beta^* \in Q$ is also such that $h(\beta^*) = h(\gamma)$, then $h(t\beta^* + (1-t)\overline{\beta}) = th(\beta^*) + (1-t) h(\overline{\beta})$ since h is quasi-monotone. Suppose that $\beta^* \in \widetilde{Q}$. Given $\lambda \in [0,1]$, set $\beta_\lambda = \lambda\beta^* + (1-\lambda)\overline{\beta}$. Since h is concave and $h(\beta^*) > h(\overline{\beta})$, we have

$$h(\beta_\lambda) \geq \lambda h(\beta^*) + (1-\lambda) h(\overline{\beta}) > h(\overline{\beta}), \qquad \forall \lambda \in (0,1].$$

and so $\beta_\lambda \in \widetilde{Q}$ for each $\lambda \in (0,1]$. By the continuity of h, we then have:

$$\begin{aligned}
th(\beta^*) + (1-t) h(\overline{\beta}) &= \lim_{\lambda\to 0} th(\beta^*) + (1-t) h(\beta_\lambda) \\
&= \lim_{\lambda\to 0} h(t\beta^* + (1-t)\beta_\lambda) \\
&= h(t\beta^* + (1-t)\overline{\beta}),
\end{aligned}$$

and so we can conclude that h is affine on Q. By Lemma C.3(ii), $\bigcap_{\beta\in Q} \Theta(\beta) \neq \emptyset$. Any $y \in \bigcap_{\beta\in Q} \Theta(\beta)$ satisfies (C.4). ■

Given a subset $D \subseteq C$, say that $h : C \to \mathbb{R}$ is D-quasi-convex if, for all $\beta', \beta'' \in D$, $h(t\beta' + (1-t)\beta'') \leq \max\{h(\beta'), h(\beta'')\}$ for all $t \in [0,1]$.

Lemma C.5. *Let $Q \subset \Delta(B)$ be a convex set and let D be a dense subset of Q. Then h is D-quasi-convex iff it is convex.*

C.2. Proof of Theorem 3.3(i)

For each $U \in \mathcal{K}(N)$, define $h_U : \gneqq (B) \longrightarrow \mathbb{R}$ by

$$h_U(\beta) = \min_{u \in U} u \cdot \beta, \quad \forall \beta \in \mathbb{R}^n.$$

It is easy to check that h_U is continuous on $\gneqq (B)$.

Preference Convexity: Using the non-negativity of μ,

$$\begin{aligned}
\mathcal{W}(\alpha x + (1-\alpha) x') &= \int \max_{\beta \in x, \beta \in x'} \min_{u \in U} (\alpha u(\beta) + (1-\alpha) u(\beta')) \, d\mu(U) \\
&\geq \int \max_{\beta \in x, \beta' \in x'} \left(\alpha \min_{u \in U} u(\beta) + (1-\alpha) \min_{u \in U} u(\beta') \right) d\mu(U) \\
&= \int \left(\alpha \max_{\beta \in x} \min_{u \in U} u(\beta) + (1-\alpha) \max_{\beta' \in x'} \min_{u \in U} u(\beta') \right) d\mu(U) \\
&= \alpha \mathcal{W}(x) + (1-\alpha) \mathcal{W}(x').
\end{aligned}$$

Coarseness: suppose $\mathcal{W}$ does not satisfy it. Then there exists a violation of Independence as in (3.7) and yet $co(\{\gamma', \gamma\}) \sim \{\gamma', \gamma\}$ for all γ', $\gamma \in \Delta(B)$. Let D be a countable dense subset of $\Delta(B)$ and $\widetilde{D} = \{\{\beta', \beta''\} : \beta', \beta'' \in D\}$. For each $x \in \widetilde{D}$, there exists $A_x \in \sigma(\mathcal{K}(N))$ with $\mu(A_x) = 1$ such that $\max_{\gamma \in x} h_U = \max_{\gamma \in co(x)} h_U$ for all $U \in A_x$. The set $\widetilde{D}$ is countable and so $\mu\left(\bigcap_{x \in \widetilde{D}} A_x\right) = 1$. Fix $U \in \bigcap_{x \in \widetilde{D}} A_x$. Then $\max_{\gamma \in x} h_U = \max_{\gamma \in co(x)} h_U$ for each $x \in \widetilde{D}$, and so

$$\begin{aligned}
\max\{h_U(\gamma'), h_U(\gamma'')\} &= \max_{\gamma \in x} h_U(\gamma) = \max_{\gamma \in co(x)} h_U(\gamma) \\
&\geq t h_U(\gamma') + (1-t) h_U(\gamma''),
\end{aligned}$$

for all $t \in [0,1]$ and $\gamma, \gamma' \in D$. Hence, h_U is D-quasi-convex on $\Delta(B)$. By Lemma C.5, h_U is convex on $\Delta(B)$. Finally, for each $\beta \in \Delta(B)$,

$$h_U\left(t\beta + (1-t)\beta_{b_*}\right) = t h_U(\beta) + (1-t) h_U\left(\beta_{b_*}\right) \text{ and } h_U(\beta) \geq h_U\left(\beta_{b_*}\right)$$

and so Lemma C.4 implies that h_U is affine on $\Delta(B)$ whenever $U \in \bigcap_{x\in\widetilde{D}} A_x$. As $\mu\left(\bigcap_{x\in\widetilde{D}} A_x\right) = 1$, for any $x', x'' \in \mathcal{X}$ we have

$$\begin{aligned}
\mathcal{W}(tx' + (1-t)x'') &= \int_{\bigcap_{x\in\widetilde{D}} A_x} \max_{\beta\in tx'+(1-t)x''} h_U(\beta)\, d\mu \\
&= \int_{\bigcap_{x\in\widetilde{D}} A_x} \left(\max_{\beta'\in x', \beta''\in x''} h_U(t\beta' + (1-t)\beta'')\right) d\mu \\
&= \int_{\bigcap_{x\in\widetilde{D}} A_x} \left(t \max_{\beta\in x'} h_U(\beta) + (1-t)\max_{\beta\in x''} h_U(\beta)\right) d\mu \\
&= t\mathcal{W}(x') + (1-t)\mathcal{W}(x'').
\end{aligned}$$

Hence, $\mathcal{W}$ is affine, which contradicts the hypothesis (3.7).

IR: We show that $\mathcal{W}(x) = \mathcal{W}(co(x))$ for all $x \in \mathcal{X}$ implies that $\mathcal{W}$ is affine. For each $\beta \in \Delta(B)$, we have $h_U\left(t\beta + (1-t)\beta_p\right) = th_U(\beta) + (1-t)p$. Suppose $\mathcal{W}(x) = \mathcal{W}(co(x))$ for all $x \in \mathcal{X}$, and let D be a countable dense subset of $\Delta(B)$. For each $x \in \mathcal{X}$ there exists $A_x \in \sigma(\mathcal{K}(N))$ with $\mu(A_x) = 1$ such that $\max_{\beta\in x} h_U = \max_{\beta\in co(x)} h_U$ for all $U \in A_x$. Let $\widetilde{D} = \{\{\beta', \beta''\} : \beta', \beta'' \in D\}$. The set $\widetilde{D}$ is countable and so $\mu\left(\bigcap_{x\in\widetilde{D}} A_x\right) = 1$.

Fix $U \in \bigcap_{x\in\widetilde{D}} A_x$. We have $\max_{\beta\in x} h_U = \max_{\beta\in co(x)} h_U$ for each $x \in \widetilde{D}$, and so

$$\begin{aligned}
\max\{h_U(\beta'), h_U(\beta'')\} &= \max_{\beta\in x} h_U(\beta) = \max_{\beta\in co(x)} h_U(\beta) \\
&\geq th_U(\beta') + (1-t)h_U(\beta''),
\end{aligned}$$

for all $t \in [0,1]$ and $\beta, \beta' \in D$. Hence h_U is D-quasi-convex. By Lemma C.5, h_U is convex, and so it is affine by Lemma C.4.

We have therefore shown that h_U is affine whenever $U \in \bigcap_{x\in\widetilde{D}} A_x$. As $\mu\left(\bigcap_{x\in\widetilde{D}} A_x\right) = 1$, by proceeding as before this implies that $\mathcal{W}$ is affine. ■

C.3. Proof of Theorem 3.3(ii)

(ii) Preference Convexity is as in (i). IR is obvious.

It remains only to derive the implications of Coarseness. Since IR is satisfied, then for all convex x' and x,

$$x' \sim x \Rightarrow \alpha x' + (1-\alpha)x \preceq x, \qquad \forall \alpha \in (0,1).$$

On the other hand, Preference Convexity implies

$$x' \sim x \Rightarrow \alpha x' + (1-\alpha) x \succeq x, \qquad \forall \alpha \in [0,1].$$

Hence, for all convex x' and x,

$$x' \sim x \Rightarrow \alpha x' + (1-\alpha) x \sim x, \qquad \forall \alpha \in (0,1),$$

and so every $x \in \mathcal{X}^c$ satisfies condition (3.5). The desired conclusion then follows from (3.6) and IR. ■

D. Appendix: Proof of Theorem 3.1

This appendix proves Theorem 3.1. Necessity is immediate; for example, IR is satisfied because

$$\max_{\beta \in x} u \cdot \beta = \max_{\beta \in co(x)} u \cdot \beta, \text{ for any } x.$$

The proof of sufficiency is quite long. We provide the complete argument here, including preliminary results on niveloids that are formulated for a more abstract setting and that extend some results of [10].

D.1. Niveloids

Let $(E, \geq, \|\cdot\|)$ be a normed Riesz space and let H be a convex cone in E containing an order unit e.[11] Say that $\|\cdot\|$ is an *e-norm* if there exists $k > 0$ such that $|f| \leq k \|f\| e$. Throughout we consider only e-norms.

Example D.1. *Each normed Riesz space has a natural e-norm, called the e-uniform Riesz norm, given by*

$$\|f\|_e = \inf \{k \geq 0 : |f| \leq ke\}.$$

In this case, $|f| \leq \|f\|_e e$ for all $f \in E$. For example, if E is a function space and e is 1_Ω, then $\|\cdot\|_e$ is the supnorm. ▲

[11] That is, E is a lattice under the order $\geq$ and the norm $\|\cdot\|$ is such that, for all $f, g \in E$, $\|f\| \leq \|g\|$ whenever $|f| \leq |g|$. Recall that

$$|f| = f^+ + f^- = f \vee 0 + (-f) \vee 0,$$

and that $e \in E_+$ is an order unit if for each $f \in E$ there is $\alpha > 0$ such that $|f| \leq \alpha e$.

Lemma D.2. *If $h_1, h_2 \in H$, then $h_1 + h_2 \in H$.*

A functional $I : H \to \mathbb{R}$ with $I(0) = 0$ is an *e-niveloid* if it is monotone, $I(e) = 1$, and satisfies

$$I(h + \alpha e) = I(h) + \alpha \text{ for all } h \in H \text{ and } \alpha \geq 0.$$

If the preceding is true also for all $\alpha < 0$ such that $h + \alpha e \in H$, say that I is e-translation invariant.

Lemma D.3. *Any e-niveloid $I : H \to \mathbb{R}$ is Lipschitz continuous and e-translation invariant.*

Remark. Observe that Lemma D.3 applies to any e-niveloid on a convex cone.

Given an e-niveloid $I : H \to \mathbb{R}$, let

$$\mathcal{E} = \{h' \in H : I(h + h') = I(h) + I(h') \ \ \forall h \in H\}.$$

The set $\mathcal{E}$ is the *domain of additivity* of I. It contains both e and 0, and it is closed under addition. When I is positively homogeneous, $\mathcal{E}$ is a convex cone and $\mathcal{E} - \mathcal{E}$ is the vector subspace that it generates.

For any functional $I' : H' \longrightarrow \mathbb{R}^1$ where H' is a convex cone satisfying $H + \mathcal{E} - \mathcal{E} \subset H' \subset E$, define its domain of additivity in the obvious way paralleling the definition given for I. We omit the proof of the next Theorem as it is a variation on results in [10].

Theorem D.4. *Let $I : H \to \mathbb{R}$ be an e-niveloid with domain of additivity $\mathcal{E}$. Define*

$$\widehat{I}(f) = \sup_{h \in H \text{ and } \xi_1, \xi_2 \in \mathcal{E}} \{I(h) + I(\xi_1) - I(\xi_2) : h + \xi_1 - \xi_2 \leq f\}, \ f \in E.$$

Then $\widehat{I} : E \to \mathbb{R}$ is an e-niveloid that extends I and whose domain of additivity includes $\mathcal{E}$. Moreover, $\widehat{I}$ is the minimal such extension in that $\widehat{I}(\cdot) \leq J(\cdot)$ for any other extension J having the preceding properties. Finally, $\widehat{I}$ is positively homogeneous if I is, it is concave if I is quasi-concave, and it is additive if I is additive and $E = \overline{H - H}$.[12]

[12] I is quasi-concave if $I(th_1 + (1-t)h_2) \geq \min\{I(h_1), I(h_2)\}$ for all $h_1, h_2 \in H$ and all $t \in [0,1]$; it is additive if $I(h_1 + h_2) = I(h_1) + I(h_2)$ for all $h_1, h_2 \in H$.

Let $(E, \geq, \|\cdot\|)$ be a Banach lattice with topological dual E^*. Denote by E_+^* the set of all monotone elements in E^* and let $\Delta = \{L \in E_+^* : L(e) = 1\}$.

Theorem D.5. *Let $I : H \to \mathbb{R}$ be a quasi-concave and positively homogeneous e-niveloid. Then there exists a maximal convex and weak*-compact set $\Gamma \subseteq \Delta$ such that*

$$I(f) = \min_{L \in \Gamma} L(f) \quad \textit{for every } f \in H, \ \textit{ and}$$
$$L'(\xi) = L(\xi) \ \textit{ for every } L', L \in \Gamma \textit{ and } \xi \in \mathcal{E} - \mathcal{E}.$$

The set Γ is a singleton if I is additive and $E = \overline{H - H}$.

Proof. Consider the extension $\widehat{I} : E \to \mathbb{R}$ given by Theorem D.4. The superdifferential $\partial \widehat{I}(f)$ at $f \in E$ is given by

$$\partial \widehat{I}(f) = \left\{ L \in E^* : \widehat{I}(g) \leq \widehat{I}(f) + L(g - f) \ \text{ for every } g \in E \right\}. \tag{D.1}$$

Since $\widehat{I}$ is concave and, by Lemma D.3, Lipschitz continuous, the set $\partial \widehat{I}(f)$ is nonempty, convex and weak*-compact for each $f \in E$.

Show that $\widehat{I}(f) = \min_{L \in \partial \widehat{I}(0)} L(f)$: Let $L \in \partial \widehat{I}(f)$. If we take $g = 0$ in (D.1), we get $\widehat{I}(f) \leq L(f)$, while if we take $g = 2f$, then we get $\widehat{I}(f) \geq L(f)$. Conclude that $\widehat{I}(f) = L(f)$. This implies that $\partial \widehat{I}(f) = \left\{ L \in \partial \widehat{I}(0) : L(f) = \widehat{I}(f) \right\}$, and so $\widehat{I}(f) = \min_{L \in \partial \widehat{I}(0)} L(f)$.

It remains to show that $\partial \widehat{I}(0) \subseteq \Delta$. We prove first that $L(\xi) = \widehat{I}(\xi)$ for all $L \in \partial \widehat{I}(0)$ and $\xi \in \mathcal{E} - \mathcal{E}$. If not, then there exist $L \in \partial \widehat{I}(0)$ and $\xi \in \mathcal{E} - \mathcal{E}$ such that $L(\xi) \neq \widehat{I}(\xi)$. Fix $f \in E$. We have $\widehat{I}(f + \alpha\xi) \leq L(f + \alpha\xi)$ for all $\alpha \in \mathbb{R}$, and so $\widehat{I}(f) \leq L(f) + \alpha \left(L(\xi) - \widehat{I}(\xi) \right)$ for all $\alpha \in \mathbb{R}$, which contradicts $\widehat{I}(f) > -\infty$.

Show now that $\partial \widehat{I}(0) \subseteq E_+^*$. Otherwise, there exist $L \in \partial \widehat{I}(0)$ and $f \in E_+$ such that $L(f) < 0$. Then $\widehat{I}(f) \leq L(f) < 0$, which contradicts the monotonicity of $\widehat{I}$.

Let $\Gamma' \subseteq \Delta$ be a convex and weak*-compact set such that $I(h) = \min_{L \in \Gamma'} L(h)$ and $L'(\xi) = L(\xi)$ for each $L', L \in \Gamma$ and each $\xi \in \mathcal{E} - \mathcal{E}$. Consider the functional $\widehat{J} : E \to \mathbb{R}$ given by $\widehat{J}(f) = \min_{L \in \Gamma'} L(f)$. Then $\widehat{J}$ is well defined since Γ' is weak*-compact. It is a concave and positively homogeneous e-niveloid that extends I and whose domain of additivity includes $\mathcal{E}$. By Theorem D.4, $\widehat{J} \geq \widehat{I}$, and this in turn implies $\Gamma' \subseteq \partial \widehat{I}(0)$. Conclude that $\partial \widehat{I}(0)$ is the subset of Δ we are seeking.

Finally, suppose that I additive on H and that $E = \overline{H - H}$. By Theorem D.4, $\widehat{I}$ is a monotone linear functional on E, and so, by standard results, Γ is a singleton. ■

Let $W : G \to \mathbb{R}$, where G is a convex subset of E containing both 0 and e. Denote by $\mathcal{G}$ the subset of G consisting of all $h' \in G$ such that

$$W\left(\frac{h}{2} + \frac{h'}{2}\right) = \frac{W(h)}{2} + \frac{W(h')}{2} \tag{D.2}$$

for all $t \in [0,1]$ and all $h \in G$ with $W(h) = W(h')$.

Corollary D.6. *Let G be a convex subset of E containing both 0 and e. Let $W : G \to \mathbb{R}$ be quasi-concave and monotone, satisfying $W(0) = 0$ and*

$$W(tf + (1-t)\gamma e) = tW(f) + (1-t)\gamma, \quad \forall f \in G, \forall t, \gamma \in [0,1]. \tag{D.3}$$

Then there exists a maximal convex and weak-compact set $\Gamma \subseteq \Delta$ such that*

$$W(f) = \min_{L \in \Gamma} L(f), \quad \forall f \in G, \text{ and}$$

$$L'(h) = L(h) \quad \forall L', L \in \Gamma, \forall h \in \mathcal{G}.$$

The set Γ is a singleton if W is affine and $E = \overline{\langle G \rangle}$.

Proof. Let $H = \bigcup_{\alpha \geq 0} \alpha G$ be the cone generated by G. For each $h \in H$, there exists $\alpha > 0$ such that $h/\alpha \in H$. Define $I : H \to \mathbb{R}$ by $I(h) = \alpha W(h/\alpha)$.

The functional I is well defined: suppose that, for a given $h \in H$, there exist $\alpha, \beta > 0$ such that $h/\alpha, h/\beta \in G$. Wlog suppose $\beta \leq \alpha$. Then

$$W\left(\frac{h}{\alpha}\right) = W\left(\frac{\beta}{\alpha}\frac{h}{\beta}\right) = W\left(\frac{\beta}{\alpha}\frac{h}{\beta} + \left(1 - \frac{\beta}{\alpha}\right)0\right) = \frac{\beta}{\alpha}W\left(\frac{h}{\beta}\right),$$

as desired. Observe that $h/\beta \in G$ and $\beta \leq \alpha$ imply $h/\alpha \in G$. Hence, given any $h_1, h_2 \in H$, there exists $\alpha \geq 1$ such that $h_i/\alpha \in G$ and $I(h_i) = \alpha W(h_i/\alpha)$ for each $i = 1, 2$. This property will be tacitly used in the sequel.

The functional I is clearly positively homogeneous, and its domain of additivity $\mathcal{E}$ is given here by $\mathcal{E} = \{h' \in H : I(h+h') = I(h) + I(h')\ \ \forall h \in H\}$. It can be shown that

$$\mathcal{E} = \{h' \in H : I(h+h') = I(h) + I(h')\ \ \forall h \in H \text{ s.t. } I(h) = I(h')\}.$$

Now we show that $\mathcal{E} = \bigcup_{\alpha \geq 1} \alpha \mathcal{G}$. First show that $\bigcup_{\alpha \geq 1} \alpha \mathcal{G} \subseteq \mathcal{E}$. Given $h' \in \mathcal{G}$ and $\alpha \geq 1$, we want to show that $\alpha h' \in \mathcal{E}$. Let $h \in H$ and $I(h) = I(\alpha h')$. There exists $\beta \geq 1$ such that $h'/\beta, h/\alpha\beta \in G$, and

$$\begin{aligned} I(\alpha h' + h) &= \alpha I\left(h' + \frac{h}{\alpha}\right) = \alpha I\left(\beta \frac{h'}{\beta} + \beta \frac{h}{\alpha\beta}\right) \\ &= 2\alpha\beta W\left(\frac{1}{2}\frac{h'}{\beta} + \frac{1}{2}\frac{h}{\alpha\beta}\right) \\ &= \alpha\beta W\left(\frac{h'}{\beta}\right) + \alpha\beta W\left(\frac{h}{\alpha\beta}\right) = I(\alpha h') + I(h). \end{aligned}$$

This proves that $\alpha h' \in \mathcal{E}$, as desired.

Conversely, let $h' \in \mathcal{E}$. There is $\beta \geq 1$ such that $h'/\beta \in G$. As $\mathcal{E}$ is a cone, $h'/2\beta \in \mathcal{E}$. Hence, for each $h \in G$ we have:

$$W\left(\frac{h}{2} + \frac{h'}{2\beta}\right) = I\left(\frac{h}{2} + \frac{h'}{2\beta}\right) =$$

$$I\left(\frac{h}{2}\right) + I\left(\frac{h'}{2\beta}\right) = \frac{W(h)}{2} + \frac{1}{2}W\left(\frac{h'}{\beta}\right),$$

and so $h'/\beta \in \mathcal{G}$, $h' \in \beta\mathcal{G}$, and we conclude that $\mathcal{E} \subseteq \bigcup_{\alpha \geq 1} \alpha \mathcal{G}$.

In sum, we have proved that $\mathcal{E} = \bigcup_{\alpha \geq 0} \alpha \mathcal{G}$. As $e \in \mathcal{G}$, in particular this implies that I is an e-niveloid.

For monotonicity, let $h_1, h_2 \in H$ be such that $h_1 \geq h_2$. There exists $\alpha > 0$ such that $h_1/\alpha, h_2/\alpha \in G$. Then

$$I(h_1) = \alpha W\left(\frac{h_1}{\alpha}\right) \geq \alpha W\left(\frac{h_2}{\alpha}\right) = I(h_2),$$

and we can conclude that I is an e-niveloid on H.

Show that I is quasi-concave on H: given any $h_1, h_2 \in H$, there exist $\alpha > 0$ such that $h_1/\alpha, h_2/\alpha \in G$. For any $t \in [0,1]$,

$$
\begin{aligned}
I(th_1 + (1-t)h_2) &= I\left(t\alpha \frac{h_1}{\alpha} + (1-t)\alpha \frac{h_2}{\alpha}\right) \\
&= \alpha I\left(t\frac{h_1}{\alpha} + (1-t)\frac{h_2}{\alpha}\right) = \alpha W\left(t\frac{h_1}{\alpha} + (1-t)\frac{h_2}{\alpha}\right) \\
&\geq \alpha \min\left\{W\left(\frac{h_1}{\alpha}\right), W\left(\frac{h_2}{\alpha}\right)\right\} = \min\{I(h_1), I(h_2)\}.
\end{aligned}
$$

Thus I is quasi-concave. Application of Theorem D.5 completes the proof. ■

D.2. Application

Define $\succeq^*$ on $\mathcal{X}$ as follows: $x \succeq^* x'$ iff $co(x) \succeq co(x')$. Then $\succeq^*$ satisfies Completeness, Transitivity, Continuity, Monotonicity, Worst-Best, and Indifference to Randomization. Moreover, since for all $\alpha \in [0,1]$ we have

$$co(\alpha x + (1-\alpha)x') = \alpha co(x) + (1-\alpha) co(x'),$$

then $\succeq^*$ also satisfies Preference Convexity and Certainty Independence. The rest of the proof is devoted to showing that $\succeq^*$ admits the representation (3.4) on $\mathcal{X}^c$. This is enough to complete the proof since (i) $\succeq$ and $\succeq^*$ agree on $\mathcal{X}^c$, and (ii) they agree on all of $\mathcal{X}$ iff $\succeq$ satisfies IR.

Define B, $\Delta(B)$, b_*, b^* and $N = \left\{u \in [0,1]^B : u_{b_*} = 0 \text{ and } u_{b^*} = 1\right\}$ as in the text. Let $\Delta(N)$ be the set of all Borel probability measures on N and $\mathcal{C}(N)$ be the set of all continuous functions on N. Endow $\mathcal{C}(N)$ with the supnorm $\|\cdot\|_s$ and define an order $\geq$ on $\mathcal{C}(N)$ by $f \geq g$ if $f(u) \geq g(u)$ for all $u \in N$.

The triple $(\mathcal{C}(N), \geq, \|\cdot\|_s)$ forms a Banach lattice; its dual is given by the set of all bounded Borel measures and the set Δ of Theorem D.5 is given by $\Delta(N)$. Moreover, 1_N is an order unit e that makes $\|\cdot\|_s$ an e-norm, while the 0 is the function on N that is identically zero.

Denote by Σ the set of all support functions $\sigma_x : N \to \mathbb{R}$, given by $\sigma_x(u) = \max_{\beta \in x} \beta \cdot u$ for each $u \in N$ and $x \in \mathcal{X}^c$. Then Σ is a convex subset of $\mathcal{C}(N)$ containing both e and 0. For the latter, note that $\sigma_{\{b^*\}} = 1_N$ and $\sigma_{\{b_*\}} = 0$.

The next result is an immediate consequence of Corollary D.6. Observe that, by standard results, a weak*-compact subsets of $\Delta(N)$ is weakly compact.

Corollary D.7. *Let $W : \Sigma \to \mathbb{R}$ be monotone, quasi-concave, satisfying $W(0) = 0$, $W(e) = 1$, and*

$$W(\alpha\sigma_x + (1-\alpha)\gamma e) = \alpha W(\sigma_x) + (1-\alpha)\gamma, \quad \forall \sigma_x \in \Sigma, \forall \alpha, \gamma \in [0,1].$$

Define $\mathcal{G}$ as in (D.2). Then there exists a maximal convex and weakly compact subset $\Pi \subseteq \Delta(N)$ such that

$$I(\sigma_x) = \min_{\pi \in \Pi} \int_N \sigma_x(u)\, d\pi, \text{ for all } \sigma_x \in \Sigma, \text{ and}$$

$$\int_N \sigma_x(u)\, d\pi' = \int_N \sigma_x(u)\, d\pi, \text{ for every } \pi', \pi \in \Pi \text{ and } \sigma_x \in \mathcal{G}.$$

The set Π is a singleton if W is affine and $E = \overline{H - H}$.

Turn finally to the proof of Theorem 3.1. Adopt the hypotheses stated there.

Lemma D.8. *There exists $\mathcal{W} : \mathcal{X}^c \to \mathbb{R}$ that represents $\succeq^*$ and such that, for each $x, x' \in \mathcal{X}^c$, $\alpha \in [0,1]$ and $\beta_p \in C$,*

$$\begin{aligned} \mathcal{W}\left(\alpha x + (1-\alpha)\beta_p\right) &= \alpha \mathcal{W}(x) + p, \text{ and} \\ \mathcal{W}(\alpha x + (1-\alpha) x') &\geq \min\{\mathcal{W}(x), \mathcal{W}(x')\}. \end{aligned}$$

The functional $\mathcal{W}$ is unique up to positive affine transformations.

Proof. The set $C = \left\{\beta_p \equiv pb_* + (1-p)b^* : p \in [0,1]\right\}$ is a convex subset of the vector space $\{\alpha b_* + \beta b^* : \alpha, \beta \in \mathbb{R}\}$. Because the preference $\succeq^*$ satisfies the vNM axioms on C, there exists an affine function $u : C \to \mathbb{R}$, unique up to positive affine transformations, such that $c \succeq^* c'$ iff $u(c) \geq u(c')$. Normalize u so that $u(b_*) = 0$ and $u(b^*) = 1$. Hence,

$$u\left(\beta_p\right) = u(pb^* + (1-p)b_*) = pu(b^*) + (1-p)u(b_*) = p.$$

For any $x \in \mathcal{X}^c$, Worst-Best and Monotonicity imply

$$\{b^*\} \sim^* \{b^*\} \cup x \succeq^* x \sim^* \{b_*\} \cup x \succeq^* \{b_*\}.$$

By Continuity there exists a $p \in [0,1]$ such that $\beta_p \equiv pb^* + (1-p)b_* \sim^* x$. Uniqueness obtains by Independence on C because $\{b^*\} \succ \{b_*\}$.

Set $\mathcal{W}(x) = u(\beta_p) = p$. Clearly, $x \succeq^* x'$ iff $\mathcal{W}(x) \geq \mathcal{W}(x')$, and $\mathcal{W}$ is the unique functional on $\mathcal{X}^c$ representing $\succeq^*$ that reduces to u on C.

Consider $x \in \mathcal{X}^c$ and $\beta_p \in C$. There exists $\beta_q \in C$ such that $x \sim^* \beta_q$. By Certainty Independence,

$$x \sim^* \beta_q \Longleftrightarrow \alpha x + (1-\alpha)\beta_p \sim^* \alpha\beta_q + (1-\alpha)\beta_p,$$

for all $\alpha \in [0,1]$, and so

$$\begin{aligned}\mathcal{W}\left(\alpha x + (1-\alpha)\beta_p\right) &= \mathcal{W}\left(\alpha\beta_q + (1-\alpha)\beta_p\right) = \alpha\mathcal{W}\left(\beta_q\right) + (1-\alpha)\mathcal{W}\left(\beta_p\right) \\ &= \alpha\mathcal{W}(x) + (1-\alpha)\mathcal{W}\left(\beta_p\right).\end{aligned}$$

Finally, the quasi-concavity of $\mathcal{W}$ is a direct consequence of Preference Convexity. ■

For any menu x, define its $\geq_D$-hull by

$$hull(x) = \{\beta \in \Delta(B) : \beta' \geq_D \beta \text{ for some } \beta' \in x\}.$$

If x is convex, then so is $hull(x)$.

Lemma D.9. *(i) For any $x \in \mathcal{X}^c$ and $\beta^0 \notin hull(x)$, there exists u in N such that*

$$\sigma_x(u) < u\left(\beta^0\right). \tag{D.4}$$

(ii) For any x and y in $\mathcal{X}^c$,

$$\begin{aligned}\sigma_y(\cdot) &= \sigma_x(\cdot) \Longrightarrow y \sim^* x, \\ \sigma_y &\geq \sigma_x \Longrightarrow y \succeq^* x\end{aligned}$$

Proof. (i) Let $y = \{\beta \in \Delta(B) : \beta \geq_D \beta^0\}$. Then y and $hull(x)$ are disjoint closed convex sets. Therefore, there exists v in $\mathbb{R}^B$ such that

$$sup_{\beta \in hull(x)}\, v \cdot \beta < v \cdot \beta^0 \leq v \cdot \beta'$$

for all β' in y. Note that $b^* \in y$ and $b_* \in hull(x)$. It follows that

$$v(b_*) \leq v(b) \leq v(b^*) \text{ for all } b.$$

Thus we can renormalize v into u in N satisfying (D.4).

(ii) We have

$$x \geq_D hull\left(x\right) \text{ and } y \geq_D hull\left(y\right).$$

By Worst-Best,

$$x \sim^* hull\left(x\right) \text{ and } y \sim^* hull\left(y\right). \tag{D.5}$$

Thus it suffices to show that

$$\sigma_x\left(\cdot\right) = \sigma_y\left(\cdot\right) \Longrightarrow hull\left(x\right) = hull\left(y\right).$$

Suppose that $\beta^0 \in hull\left(x\right) \backslash hull\left(y\right)$. Then by (i) there exists u in N such that

$$\sigma_y\left(u\right) < \beta^0 \cdot u \leq \sigma_x\left(u\right), \tag{D.6}$$

a contradiction. Conclude that $hull\left(x\right) = hull\left(y\right)$.

Finally, let $\sigma_y \geq \sigma_x$ on N. By (D.5), it is enough to show that $hull\left(x\right) \subseteq hull\left(y\right)$. Otherwise, there exists $\beta^0 \in hull\left(x\right) \backslash hull\left(y\right)$, which implies (D.6), contradicting our hypothesis. ■

Define $W : \Sigma \to \mathbb{R}$ by

$$W\left(\sigma_x\right) = \mathcal{W}\left(x\right) \text{ for each } x \in \mathcal{X}^c.$$

By Lemma D.9, W is well defined and monotone, with $W\left(0\right) = 0$ and $W\left(e\right) = 1$. Moreover,

$$\begin{aligned} W\left(\alpha\sigma_x + \left(1-\alpha\right)\gamma\sigma_{\{\beta^*\}}\right) &= W\left(\sigma_{\alpha x + (1-\alpha)\beta_\gamma}\right) = \mathcal{W}\left(\alpha x + \left(1-\alpha\right)\beta_\gamma\right) \\ &= \alpha\mathcal{W}\left(x\right) + \left(1-\alpha\right)\mathcal{W}\left(\beta_\gamma\right) \\ &= \alpha W\left(\sigma_x\right) + \left(1-\alpha\right)\gamma W\left(\sigma_{\{b^*\}}\right), \quad \text{and} \end{aligned}$$

$$\begin{aligned} W\left(\alpha\sigma_x + \left(1-\alpha\right)\sigma_{x'}\right) &= W\left(\sigma_{\alpha x + (1-\alpha)x'}\right) = \mathcal{W}\left(\alpha x + \left(1-\alpha\right)x'\right) \\ &\geq \min\left\{\mathcal{W}\left(x\right), \mathcal{W}\left(x'\right)\right\} = \min\left\{W\left(x\right), W\left(x'\right)\right\}. \end{aligned}$$

Hence W satisfies the hypotheses of Corollary D.7, and so there exists a maximal convex and weakly compact subset $\Pi \subseteq \Delta\left(N\right)$ such that, for each $\sigma \in \Sigma$,

$$\mathcal{W}\left(x\right) = W\left(\sigma_x\right) = \min_{\pi\in\Pi}\int_N \sigma_x\left(u\right)d\pi = \min_{\pi\in\Pi}\int_N \max_{\beta\in x} u\left(\beta\right)\,d\pi,$$

and, for each $\pi', \pi \in \Pi$ and each $\sigma_x \in \mathcal{G}$,

$$\int_N \sigma_x(u)\, d\pi' = \int_N \sigma_x(u)\, d\pi. \tag{D.7}$$

Here $H = \bigcup_{\alpha \geq 0} \alpha\Sigma$ and $\mathcal{G}$ is the subset of Σ consisting of all $\sigma_{x'}$ such that

$$W\left(\frac{1}{2}\sigma_x + \frac{1}{2}\sigma_{x'}\right) = \frac{1}{2}W(\sigma_x) + \frac{1}{2}W(\sigma_{x'})$$

for all $\alpha \in [0,1]$ and all $\sigma_x \in \Sigma$ such that $W(\sigma_x) = W(\sigma_{x'})$.

Observe that $\sigma_{x'} \in \mathcal{G}$ iff $x' \in \mathcal{G}_{\succeq^*}$, where

$$\mathcal{G}_{\succeq^*} = \left\{x' \in \mathcal{X}^c : \frac{1}{2}x + \frac{1}{2}x' \sim^* x, \quad \forall x \in \mathcal{X}^c \text{ with } x \sim^* x'\right\}.$$

Therefore (D.7) implies that

$$\min_{\pi \in \Pi} \int_N \max_{\beta \in x'} u(\beta)\ d\pi' = \min_{\pi \in \Pi} \int_N \max_{\beta \in x'} u(\beta)\, d\pi$$

for all $\pi', \pi \in \Pi$ and for all $x' \in \mathcal{G}_{\succeq^*}$.

Finally, the desired result follows from Corollary D.7 because W satisfies the hypotheses adopted there.

References

[1] E. Dekel, B. Lipman and A. Rustichini, A unique subjective state space for unforeseen contingencies, *Econometrica* 69 (2001), 891-934.

[2] A.P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, *Ann. Math. Statist.* 38 (1967), 325-339.

[3] L.G. Epstein and M. Schneider, Recursive multiple-priors, *J. Econ. Theory* 113 (2003), 1-31.

[4] P. Ghirardato, Coping with ignorance, unforeseen contingencies and nonadditive uncertainty, *Econ. Theory* 17 (2001), 247-276.

[5] I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique priors, *J. Math. Econ.* 18 (1989), 141-153.

[6] H. J. Greenberg and W. P. Pierskalla, A review of quasi-convex functions, *Oper. Res.* 19 (1971), 1553-1570.

[7] I. N Herstein and J. Milnor, An axiomatic approach to measurable utility, *Econometrica* 21 (1953), 291–297.

[8] D. Kreps, A representation theorem for 'preference for flexibility', *Econometrica* 47 (1979), 565-577.

[9] D. Kreps, Static choice in the presence of unforeseen contingencies, in *Essays in Honour of F. Hahn*, P. Dasgupta et al eds., MIT Press, 1992.

[10] F. Maccheroni, M. Marinacci, and A. Rustichini, Ambiguity aversion, malevolent nature, and the variational representation of preferences, 2004.

[11] E. Maskin and J. Tirole, Unforeseen contingencies and incomplete contracts, *Rev. Ec. Studies* 66 (1999), 83-114.

[12] S. Mukerji, Understanding the nonadditive probability decision model, *Econ. Theory* 9 (1997), 23-46.

[13] S. Mukerji, Ambiguity aversion and incompleteness of contractual form, *Amer. Econ. Rev.* 88 (1998), 1207-1231.

[14] L. J. Savage, *The Foundations of Statistics*, 2nd edition, Dover, 1972.

[15] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* 57, 571-587, 1989.

[16] G. Shafer, *A Mathematical Theory of Evidence*, Princeton U. Press, Princeton, 1976.

[17] G. Shafer, Savage revisited, *Statistical Science* 1 (1986), 463-501.