SUBJECTIVE STATES: A MORE ROBUST MODEL*

Larry G. Epstein Kyoungwon Seo

March 21, 2007

Abstract

Following Kreps [11], Nehring [15, 16] and Dekel, Lipman and Rustichini [5], we study the demand for flexibility and what it reveals about subjective uncertainty. As in the cited papers, the latter is represented by a *subjective state space* consisting of possible future preferences over actions to be chosen ex post. One contribution is to provide axiomatic foundations for a range of alternative hypotheses about the nature of these ex post preferences. Secondly, we establish a sense in which the subjective state space is uniquely pinned down by the agent's ex ante ranking of (random) menus. For both purposes, we show that it is advantageous to assume that the agent ranks *random menus*, and to think of ex post *upper contour sets* rather than ex post preferences. Finally, we demonstrate the tractability of our representation by showing that it can model the two comparative notions "2 desires more flexibility than 1" and "2 is more averse to flexibility-risk than is 1."

---

*Dept. of Economics, U. Rochester, Rochester, NY 14627, lepn@troi.cc.rochester.edu and kseo@troi.cc.rochester.edu. Epstein gratefully acknowledges the financial support of the NSF (award SES-0611456). We are grateful to Bart Lipman for useful discussions, to Klaus Nehring for informing us about some of his unpublished work, and especially to Massimo Marinacci for showing us how to exploit Choquet's Theorem.

1. INTRODUCTION

Following Kreps [11], Nehring [15, 16] and Dekel, Lipman and Rustichini [5], we study the demand for flexibility and what it reveals about subjective uncertainty. As in the cited papers, the latter is represented by a *subjective state space* consisting of possible future preferences over actions to be chosen ex post. One contribution is to provide axiomatic foundations for a range of alternative hypotheses about the nature of these ex post preferences. Secondly, we establish a sense in which the subjective state space is uniquely pinned down by the agent's ex ante ranking of (random) menus (sets of possible future actions). For both purposes, we show that it is advantageous to assume that the agent ranks *random menus*, and to think of ex post *upper contour sets* rather than ex post preferences. We elaborate now on these contributions and on our value-added relative to the cited literature.

Kreps studies an agent who ranks menus of actions - elements of an abstract set $\mathcal{B}$ - one of which is to be chosen ex post from the menu selected ex ante. When $\mathcal{B}$ is finite, he shows that a simple set of axioms characterizes a representation of preference over menus (all subsets of $\mathcal{B}$) that can be interpreted as reflecting uncertainty about future preferences over $\mathcal{B}$. The representation for preference over menus that he derives has the form:

$$W(x) = \int \max_{\beta \in x} v(\beta)\, d\mu(v), \tag{1.1}$$

where x is a menu (subset of $\mathcal{B}$), and μ is a probability measure over functions $v : \mathcal{B} \to \mathbb{R}$, each of which is a utility function representing an ex post ordering of actions. The support of μ can be thought of as a subjective state space underlying the ex ante ranking of menus. We think of subjective states as describing the agent's conceptualization of the future and thus as being foreseen by her.[1] (Kreps, and also Dekel, Lipman and Rustichini, consider also representations that are not additive over the possible ex post utility functions. However, in this paper we consider only additive models and when referring to the cited papers we have in mind only their additive models.)

Two major extensions of Kreps' analysis have been pursued. One, by Dekel, Lipman and Rustichini (henceforth DLR), is motivated by the desire to derive a *unique* subjective state space for each agent. A finite $\mathcal{B}$ does not afford the

[1] An alternative interpretation, developed by Kreps [13], is that these contingencies are unforeseen.

richness needed to pin down relevant ex post preferences; and, more generally, Kreps (Theorem 2) is able only to provide a hard-to-interpret set of transformations of the subjective state space that preserve the representation of ex ante preference. DLR obtain uniqueness results by assuming that the agent (i) ex ante ranks menus of lotteries, that is, $\mathcal{B} = \Delta(B)$ for some finite set of alternatives B, and (ii) satisfies alternative independence-style axioms that permit a subjective state space consisting only of vNM ex post utilities. In particular, DLR show that the subjective state space is "essentially unique" given the vNM restriction on ex post utilities. Epstein, Marinacci and Seo [7] argue that the latter limitation is unattractive because it precludes subjective states from being ambiguous or coarse. Thus we are led to seek a model that is *more robust* in that it provides axiomatic foundations for a subjective state space defined by less restrictive assumptions about ex post preferences - for example, where they are assumed only to be upper semicontinuous, or alternatively upper semicontinuous and convex.[2]

A second extension of Kreps' model, by Nehring [15, 16], permits the menu from which ex post choice is made to be random. In the published version, that randomness is subjective along the lines of Savage. However, in his working paper [15] he first considers a setting where the randomness of menus is objective - ex ante preference is over lotteries whose outcomes are menus, or over random menus. It is this version of his model that is most pertinent to our work, and thus when we refer to Nehring's contribution, it is to his analysis for the domain of (objective) random menus. We borrow a great deal from it. First, we also adopt the domain of random menus for ex ante preference. Second, our central axiom is adapted from his key axiom, which he calls Indirect Stochastic Dominance (ISD). In addition, Nehring [15, Section 5] points out the importance of ex post upper contour sets, which we also emphasize. (In fact, Kreps [11] was the first to draw attention to ex post (lower) contour sets.)

We add to Nehring's work in several ways. First, we drop his restriction that $\mathcal{B}$ is finite; any compact Polish (complete separable metric) space is permitted, including, in particular, the simplex $\Delta(B)$ as in DLR. However, finite $\mathcal{B}$ is also permitted, and this is noteworthy because we nevertheless provide a uniqueness result for the agent's subjective uncertainty. This is not surprising (in light of Nehring's analysis and) given that our domain is rich because of the presence of lotteries over menus. However, our analysis reveals more: by generalizing

[2] In terms of modeling ambiguous or coarse states, our objective here is only to provide a framework that accommodates them. The framework is applied and specialized in [7] to develop an axiomatic model of preference that is designed explicitly to capture ambiguity or coarseness.

Nehring's analysis to any compact Polish $\mathcal{B}$, we are able to show that richness of $\mathcal{B}$ is neither necessary (given the ranking of random menus) nor sufficient (in the absence of DLR's axioms and given only the ranking of nonrandom menus) for uniqueness. In this sense, the domain of lotteries over menus of alternatives is more powerful than the DLR domain consisting of menus of lotteries over alternatives.

Second, we generalize Nehring's axiom ISD to a parametric *family* of axioms, each of which is shown to characterize a subjective state space where ex post preferences satisfy a specific property (beyond completeness and transitivity) - upper semicontinuity and convexity are two examples of such properties. These results have no counterparts in Nehring's work.[3] Neither does our analysis of comparative behavioral notions (described later in this introduction). Finally, as noted, particularly in his unpublished working paper [15, Section 5], Nehring also shows the usefulness of upper contour sets for describing the uniqueness properties of representations with subjective states. Besides generalizing his results in this regard, we also elaborate and highlight this point which we feel has not been widely recognized and appreciated in the literature.[4]

We must acknowledge at the outset a limitation of our model. Though it is robust in the sense described above, this robustness comes at a cost: we assume certainty that a specified lottery β_* will be worst ex post. This assumption is needed only when $\mathcal{B}$ is infinite, and then its role is purely technical - to show that one may extend a linear functional from a linear subspace to the universal infinite dimensional linear space. Since it has no conceptual role, there is reason to hope that it might be dispensable in the future.

Upper Contour Sets and Uniqueness

As noted, DLR prove that, under their axioms, there is an (essentially) unique representation for preference of the form (1.1) where μ has support on the set of vNM ex post preferences. However, as they are well aware, there may exist also other representations where ex post preferences are not vNM. Figure A.1 illustrates this possibility (dotted areas are regions of indifference). In one representation, the vNM preference corresponding to v is expected with certainty,

[3] However, see [17] for related results characterizing convexity of upper contour sets, albeit formulated in the context of a study of diversity rather than individual decision-making and flexibility.

[4] This may be due in part to the fact that those features of Nehring's analysis that we emphasize appear primarily in his unpublished paper, and that the latter has also other foci - the intrinsic preference for freedom of choice, for example.

where v is normalized to have $[0,1]$ as its range. In the other, she assigns probabilities $a \in (0,1)$ and $(1-a)$ to the payoff functions v_1 and v_2 respectively, where

$$v_1(\beta) = \frac{1}{a}\min\{a, v(\beta)\} \text{ and } v_2(\beta) = \frac{1}{1-a}\max\{0, v(\beta) - a\}.$$

Both specifications imply, via (1.1), the same level of utility $W(x)$ for any menu x; therefore, they imply the same ranking of random menus if, as assumed below, the utility of any lottery over menus is given by the expected value of W. Note that v_1 and v_2 do not conform to expected utility theory, but they do conform to the Betweenness axiom - ex post upper contour sets and lower contour sets are both convex - an axiom studied in risk preference theory [2, 4]. Nonuniqueness above does not rely on any special features of this example - it is the rule - and the underlying intuition for this is clear: when evaluating a given menu x ex ante, and anticipating one of $\{v_i\}_{i=1}^n$ to occur, the implied value of x given v_i, $\max_{\beta \in x} v_i(\beta)$, depends on the highest upper contour set for v_i that intersects x, and not on the entire function $v_i(\cdot)$. This suggests that it might be possible to piece together upper contour sets, or indifference sets, from the various v_i's to construct another set $\{v_i'\}_{i=1}^{n'}$ that would lead to the same evaluation of any menu x.

We see that DLR's proposed remedy for nonuniqueness amounts to the selection of a canonical representation, consisting of vNM preferences ex post, amongst all possible representations, including those where ex post utilities may not conform to vNM. Though seemingly natural, the selection of any particular representation as canonical is invariably ad hoc. DLR offer two convincing justifications for their choice. One is minimality - they prove (Theorem 3B) that the vNM subjective state space is minimal among all subjective state spaces. Secondly, they show that their canonical representation permits an intuitive connection between the size of the state space and the desire for flexibility.

Our concern with DLR's treatment of uniqueness is that it is applicable given only assumptions on preference that are (for some purposes) too strong. Thus we cannot adopt it here. Instead, working within the framework of preferences satisfying our weaker axioms, we propose a canonical representation that deviates from vNM ex post, but that is uniquely pinned down by preference over random menus, and also admits a clear interpretation. We elaborate now on our approach.

The key point is that, while there exist many different representations of the ex ante preference $\succeq$, they all induce the same (suitably defined) distribution m of upper contour sets (see Theorem 3.1 below).[5] This is illustrated by Figure A.1:

[5] As indicated above, uniqueness depends on the assumption that preference is defined over

Recall that the range of v is $[0,1]$ and adopt the uniform (Lebesgue) measure on $[0,1]$. Any upper contour set relevant in this example is indexed by a utility level s in $[0,1]$. Thus we can identify the distribution over upper contour sets induced by v with the uniform distribution on the unit interval. Similarly, the distributions induced by v_1 and v_2 may be identified with the uniform distributions on $[0,a]$ and $[a,1]$ respectively. But the $a:(1-a)$ mixture of these latter distributions equals the uniform distribution on $[0,1]$. Thus the induced distributions over upper contour sets coincide.

Since each upper contour set can be identified with its indicator function, one obtains a representation of the form (1.1) where $\mu = m$ and each utility function v in its support is $0-1$ valued. This is the (unique) canonical representation that we propose.

The subjective state space consisting of (indicator functions of) upper contour sets is large - it is definitely not minimal in any sense. However, in addition to being well-defined given only the weak axioms specified below, the canonical representation we propose has the advantage that it expresses at a glance the (unique) distribution over ex post upper contour sets implied by $\succeq$, and therefore also the nature of the agent's relevant uncertainty about her ex post preferences. As an illustration, suppose that there exists one subjective state space in which all ex post preferences are convex (all upper contour sets are convex). Then uniqueness of the distribution over upper contour sets implies that for every subjective state space every ex post preference is convex; that is, convexity of ex post preferences is a feature of *all* subjective state spaces and thus is a property of the given ex ante preference $\succeq$. Therefore, the latter permits the *unequivocal* (independent of the representation) interpretation that the agent ranks random menus *as if* she is certain that all ex post preferences are convex.[6] Similarly for other properties of ex post preference that can be expressed in the form "every upper contour set satisfies a suitable condition".[7]

Section 4 demonstrates the tractability of our representation, and the intuitive

---

random menus, and not merely menus as in the models of Kreps and DLR. We elaborate on this point below.

[6] In contrast, if preference $\succeq$ satisfies the DLR axioms and thus admits a representation with μ supported by vNM ex post utility functions, the interpretation whereby the agent is certain that she will have vNM preferences ex post is supported by one representation but not by all - this is illustrated by the example in Figure A.1.

[7] The collection of upper contour sets satisfying this condition must be suitably closed. Another example of such a property is Betweenness, where both upper contour sets and their complements are convex.

connection that it affords between subjective uncertainty and the demand for flexibility. In addition, we define the comparative behavioral notion "2 is more averse to flexibity-risk than is 1."[8] We show that in the DLR framework, but not in ours, "2 desires more flexibility than 1" if and only if 2 is more averse to flexibility-risk. Since these two notions seem conceptually distinct, this demonstrates another sense in which our model is more robust.

2. THE MODEL

2.1. Preliminaries

Let $\mathcal{B}$ be a compact Polish space of actions. A *menu* is a (nonempty) closed subset of $\mathcal{B}$; $\mathcal{K}(\mathcal{B})$ denotes the set of all menus.[9] A *random menu* is a lottery over $\mathcal{K}(\mathcal{B})$, that is, an element of $\Delta(\mathcal{K}(\mathcal{B}))$. An ex ante preference $\succeq$ is defined on $\Delta(\mathcal{K}(\mathcal{B}))$.

The agent ranks random menus ex ante *as if* expecting the following time line: a menu x is realized, then some subjective uncertainty is resolved, and finally, at a later ex post stage she chooses an action from x. Though choice at the ex post stage is not explicitly modeled, it underlies intuition for the axioms and for the representation of $\succeq$. In particular, the demand for flexibility (the preference for large menus) is understood as arising from uncertainty about ex post preferences.

A central special case is where $\mathcal{B}$ is a set of lotteries, $\mathcal{B} = \Delta(B)$ for some compact Polish set B. This is the case considered by DLR, though they restrict B to be finite, and consider preference only over (nonrandom) menus. In light of the importance of this special case in the literature, and because it permits more concrete and familiar interpretations, even in the general case we sometimes refer to elements of $\mathcal{B}$ as lotteries.

Generic elements of $\Delta(\mathcal{K}(\mathcal{B}))$ are $P, P', Q, ...$, generic menus are denoted x, x', y ..., and generic lotteries are denoted β, β', γ, ...

We make use of the fact that, by [1, Theorem 3.63], $\{x \in \mathcal{K}(\mathcal{B}) : x \subset z\}$ is open in $\mathcal{K}(\mathcal{B})$ for every open subset $z \subset \mathcal{B}$. Therefore, for any menu y,

$$\{x \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\} = \mathcal{K}(\mathcal{B}) \setminus \{x \in \mathcal{K}(\mathcal{B}) : x \subset \mathcal{B} \setminus y\}$$

[8] Nehring [15, Section 4] defines a related notion of absolute risk aversion, but does not discuss or characterize comparative notions.

[9] Every metric space X is endowed with the Borel σ-algebra, $\Delta(X)$ denotes the set of all (Borel) probability measures on X endowed with the weak convergence topology, and $\mathcal{K}(X)$ is the set of all nonempty closed subsets of X endowed with the Hausdorff metric topology. Then $\Delta(X)$ and $\mathcal{K}(X)$ are compact Polish if X is compact Polish.

is closed, hence Borel measurable.

2.2. Axioms

We adopt the following axioms for the binary relation $\succeq$ on $\Delta(\mathcal{K}(\mathcal{B}))$.

Axiom 1 (Ex Ante vNM). *There exists $W : \mathcal{K}(\mathcal{B}) \to \mathbb{R}$ bounded and measurable such that $\succeq$ is represented by the expected utility function*

$$\mathcal{W}(P) = \int_{\mathcal{K}(\mathcal{B})} W(x)\, dP(x).$$

The foundations for such a representation are well-known (see Fishburn [9, Theorem 10.3]). The underlying properties of preference are: completeness, transitivity, mixture continuity, independence, and an axiom, denoted A4b by Fishburn, that is similar in spirit to Savage's $P7$. The first four are the axioms used in the Mixture Space Theorem, and the last is needed to ensure the expected utility form. Continuity of preference is not necessary for Ex Ante vNM, though, as shown by Grandmont [10], it is sufficient when combined with completeness, transitivity and independence. (See Kreps [12, pp. 59-67] for a textbook discussion.)

Because we criticized DLR's adoption of independence in the introduction, it is important to distinguish DLR's version of independence from that implied by Ex Ante vNM. The latter version has the following form: For all random menus P, P' and Q and for all $0 < \alpha < 1$,

$$P' \succeq P \iff \alpha P' + (1-\alpha) Q \succeq \alpha P + (1-\alpha) Q.$$

To interpret this condition, note that, since a mixture such as $\alpha P + (1-\alpha) Q$ is a random menu, it follows from the time line described above that a specific menu is realized before the agent sees a subjective state and chooses from the menu. In particular, therefore, all randomization in both component measures P and Q, as well as in the mixing is completed before then. It is this immediacy of the randomization that renders this version of independence intuitive and that distinguishes it from DLR's version, where the coin toss corresponding to the mixing is completed after choice from the menu.[10]

[10] See [7] for elaboration and for an argument that if the coin toss corresponding to the mixing is completed after choice from the menu, then randomization can be valuable and thus DLR's form of independence may not be intuitive.

Though we do not assume that preference is continuous, we do assume that it satisfies the following weaker requirement.[11]

Axiom 2 (Right-Continuity). *If $x_n \searrow x$ and $y \succ x$, then $y \succ x_n$ for some n.*

By $x_n \searrow x$, we mean the set-theoretic conditions $x_{n+1} \subset x_n$ and $\cap_1^\infty x_n = x$. Note, however, that for a declining sequence,[12]

$$\cap_1^\infty x_n = x \text{ if and only if } \lim x_n = x,$$

where the latter indicates convergence in the Hausdorff metric. Consequently, the axiom is weaker than continuity. Another remark is that given also that larger menus are weakly preferable, as implied by our final axiom, it follows that $y \succ x_{n'}$ for all $n' \geq n$.

The next axiom excludes total indifference.

Axiom 3 (Nondegeneracy). *There exist random menus such that $P \succ P'$.*

To introduce our key axiom, we consider first the translation of Nehring's axiom into our setting. For any two random menus P' and P, say that P' *dominates* P if, for all menus y,

$$P'\left(\{x \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}\right) \geq P\left(\{x \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}\right). \tag{2.1}$$

Nehring assumes:

Indirect Stochastic Dominance (ISD): $P' \succeq P$ whenever P' dominates P.

To interpret ISD, think of y as an upper contour set for some conceivable ex post preference over actions. Thus actions in y are "desirable" according to that ex post preference and $x \cap y \neq \emptyset$ indicates that x contains at least one desirable action, in which case we might refer to x as being desirable. Accordingly, P' dominates P if the probability of the realization of a desirable menu is larger under P', **and** if this is true for every set y and hence for every conceivable definition of "desirable."

[11] We adopt the obvious notation, whereby x is identified with δ_x and so on.

[12] Apply the characterization of Hausdorff convergence [1, Theorem 3.65]: let $\cap_1^\infty x_n = x$. Then $x \subset x_n \implies x \subset Lix_n$. Also, $\beta \notin x \implies \beta \notin x_N \implies \beta \in G \subset (x_N)^c \subset (\cup_N^\infty x_n)^c$ for some N and open set $G \implies \beta \notin Lsx_n$. Conclude that $x \subset Lix_n \subset Lsx_n \subset x$, which implies $Lix_n = Lsx_n = x$, and hence $x_n \to x$. The converse is also straightforward.

We strengthen ISD by limiting the menus y for which (2.1) is required to hold. There is already a restriction on the sets y imposed in ISD - they must be closed (by virtue of being menus). This restriction is vacuous in Nehring's setting where $\mathcal{B}$ is finite, but it has content here: since only upper contour sets that are closed enter into the definition of dominance, and hence into ISD, it follows that the agent is certain that other forms of upper contour sets are irrelevant, that is, she is certain that ex post preference will be upper semicontinuous.

Our strengthening of ISD accommodates certainty about other properties of ex post preference. Each such property implies a restriction on upper contour sets, and thus implies that they lie in a particular subset Y of menus. Our general formulation takes Y as parametric; it can be any subset of menus satisfying two restrictions described shortly.

Say that P' Y-*dominates* P, written $P' \geq_Y P$, if (2.1) is satisfied for all menus $y \in Y$. We assume:

Axiom 4 (Y-Dominance). *If $P' \geq_Y P$, then $P' \succeq P$.*

There are two notable implications of the axiom that do not depend on the specification of Y. First, when $P' = \delta_{x'}$ and $P = \delta_x$ are degenerate, then $\delta_{x'}$ dominates δ_x if $x' \supset x$. Therefore, Y-Dominance implies *Monotonicity*:

$$x' \supset x \Longrightarrow x' \succeq x.$$

It also implies (given independence) Kreps' second key axiom [11, condition (1.5)]: given any menus x, x_1 and x_2, let

$$P' = \tfrac{1}{2}\delta_{x \cup x_1} + \tfrac{1}{2}\delta_{x \cup x_2} \text{ and } P = \tfrac{1}{2}\delta_x + \tfrac{1}{2}\delta_{x \cup x_1 \cup x_2}. \tag{2.2}$$

Then P' dominates P, and thus Y-Dominance implies that

$$\tfrac{1}{2}\delta_{x \cup x_1} + \tfrac{1}{2}\delta_{x \cup x_2} \succeq \tfrac{1}{2}\delta_x + \tfrac{1}{2}\delta_{x \cup x_1 \cup x_2}.$$

Deduce (from Independence) that

$$\delta_x \sim \delta_{x \cup x_1} \implies \delta_{x \cup x_2} \succeq \delta_{x \cup x_1 \cup x_2}.$$

Since Monotonicity is also implied, we have finally that (in friendlier notation)

$$x \sim x \cup x_1 \Longrightarrow x \cup x_2 \sim x \cup x_2 \cup x_1,$$

which is Kreps' axiom.

We impose two restrictions on Y:

Y1 There exists β_* such that $\beta_* \notin y$ for every $y \in Y$.

Y2 Y is relatively closed in $\mathcal{K}_-$, where $\mathcal{K}_- = \{y \in \mathcal{K}(\mathcal{B}) : \beta_* \notin y\}$.[13]

Y1 implies that $\mathcal{B} \notin Y$ - intuitively, in order to define a meaningful notion of 'desirable', an upper contour set should exclude something. Any action β_* satisfying Y1 is not desirable ex post regardless of how 'desirable' is defined. Therefore, Y-Dominance expresses ex ante certainty that β_* will be worst ex post. The assumption was discussed in the introduction, where it was pointed out that it is needed only in the case where $\mathcal{B}$ is infinite.

Secondly, we would like to assume that Y is closed, but that is too strong given Y1 - it is possible that $\beta_* \notin y_n \to y$, $\beta_* \in y$, and hence $y \notin Y$ even if $y_n \in Y$ for all n. Thus we adopt the weaker assumption that Y is relatively closed in $\mathcal{K}_-$. Though the latter rules out some cases of interest - the set of all upper contour sets generated by a single (upper semi-)continuous utility function v is not closed if v has thick indifference sets - it admits a range of natural specifications. One example is

$$Y = \{y \in \mathcal{K}(\mathcal{B}) : \beta_* \notin y\},$$

the set of all menus not containing β_*. Another important example is

$$Y = \{y \in \mathcal{K}(\mathcal{B}) : y \text{ is convex and } \beta_* \notin y\}.$$

Finally, let V^0 be any family of (upper semi-)continuous functions $v : \mathcal{B} \to [0,1]$, such that $v(\beta_*) = 0$, and let Y be the set of all upper contour sets generated by V^0, in the sense that

$$Y = cl\left(\left\{\{\beta : v(\beta) \geq \kappa\} : v \in V^0 \text{ and } 0 \leq \kappa \leq 1\right\}\right) \cap \mathcal{K}_-. \tag{2.3}$$

The above two restrictions imply that Y is measurable: $\mathcal{K}_-$ is open and $Y = cl(Y) \cap \mathcal{K}_-$.

2.3. Utility

We wish to adopt minimal assumptions on the nature of ex post preferences over $\mathcal{B}$. Completeness and transitivity are relatively innocuous. In order to ensure the existence of optimal elements in every menu ex post (though ex post choice exists only in the mind of the agent), we assume that ex post preferences are upper

[13] That is, $Y = cl(Y) \cap \mathcal{K}_-$.

semicontinuous. Since $\mathcal{B}$ is compact Polish, every such ex post preference can be represented by an upper semicontinuous payoff function (in fact, this is true much more generally - see Rader [19], for example). It follows that any upper semicontinuous ex post preference that is not total indifference and that ranks the specified lottery β_* as worst has a utility representation by some (nonunique) $v : \mathcal{B} \to \mathbb{R}$ lying in V - the set of all upper semicontinuous (ex post) payoff functions satisfying

$$0 = v(\beta_*) \leq v(\cdot) \leq \max_{\beta \in \mathcal{B}} v(\beta) = 1. \tag{2.4}$$

To focus on preferences whose upper contour sets lie in the subset of menus Y, consider also

$$V^Y = \{v \in V : \{\beta : v(\beta) \geq \kappa\} \in Y \text{ for all } 0 < \kappa \leq 1\}. \tag{2.5}$$

We need a topology (and corresponding Borel σ-algebra) for V that we now describe. Denote by $USC(\mathcal{B})$ the set of upper semicontinuous (usc) functions from $\mathcal{B}$ into $[0,1]$. Adopt the topology τ generated by the subbasis:

$$\{v : \sup_{\beta \in z} v(\beta) > \kappa\} \text{ and } \{v : \sup_{\beta \in x} v(\beta) < \kappa\} \tag{2.6}$$

where z and x vary over open and compact (or equivalently, closed) sets respectively. This is the weakest topology such that the mapping $v \longmapsto \sup_{\beta \in y} v(\beta)$ is lower semicontinuous (lsc) for each open y and usc for each closed y. Then τ renders $USC(\mathcal{B})$ compact Polish. See [18] for details about τ supporting assertions made here; for an application of this topology in economics, and for other properties, see Epstein and Peters [8].[14]

A critical property of τ is that it is consistent with the Hausdorff metric topology on $\mathcal{K}(\mathcal{B})$. Each closed subset y can be identified with the usc function $\mathbf{1}_y(\cdot)$. Under this identification, $\mathcal{K}(\mathcal{B}) \subset USC(\mathcal{B})$ and the restriction of τ coincides with the Hausdorff metric topology.

With this topology in place, we can now assert that V is closed in $USC(\mathcal{B})$ and (see Lemma A.1) that V^Y is a Borel-measurable subset of $USC(\mathcal{B})$.

[14] One such property, used below, is that the mapping $(v, \beta) \longmapsto v(\beta)$ is usc on $USC(\mathcal{B}) \times \mathcal{B}$. An implication is that $(x, v) \longmapsto \max_{\beta \in x} v(\beta)$ is usc; this follows from a form of the Maximum Theorem [1, Lemma 14.29].

Any Borel probability measure $\mu \in \Delta(V)$ generates a utility function $\mathcal{W}$ on $\Delta(\mathcal{K}(\mathcal{B}))$ of the form:[15]

$$\mathcal{W}(P) = \int \left(\int \max_{\beta \in x} v(\beta)\, d\mu(v) \right) dP(x). \tag{2.7}$$

Refer to μ as a *representation* (of the preference corresponding to $\mathcal{W}$), and as a *Y-representation* if, for the given Y, μ is carried by V^Y, that is,

$$\mu\left(V^Y\right) = 1. \tag{2.8}$$

The next theorem is our first main result.

Theorem 2.1. *Let the set Y satisfy conditions Y1 and Y2.*

(a) Then $\succeq$ satisfies Ex Ante vNM, Right-Continuity, Nondegeneracy and Y-Dominance if and only if it admits a Y-representation.

(b) Moreover, if μ' is any representation for a preference $\succeq$ satisfying the conditions in (a), then

$$\mu'\left(V^Y\right) = 1. \tag{2.9}$$

Part (a) describes the foundations for our model of 'preference for flexibility'. Consider first its place in the literature. The implied utility for (nonrandom) menus is $W : \mathcal{K}(\mathcal{B}) \to \mathbb{R}$, where W has the form described in the introduction:

$$W(x) = \int_V \max_{\beta \in x} v(\beta)\, d\mu(v).$$

As described earlier, Kreps [11] derives such a representation when $\mathcal{B}$ is finite, and DLR characterize the special case where $\mathcal{B}$ is the simplex and μ has support on the set of vNM utility functions. Roughly, and ignoring technical differences, our result fits between theirs: in contrast to Kreps, it imposes structure on ex post utility via the condition $\mu\left(V^Y\right) = 1$, and it differs from DLR in that the latter condition is more general (or robust) than the DLR restriction to vNM ex post utilities. Nehring [15, Theorem 1] proves a counterpart of Theorem 2.1 for the particular case $Y = \mathcal{K}(\mathcal{B})$, without requiring a worst action as in Y1, but under the assumption that $\mathcal{B}$ is finite.[16] In our more general setting, upper semicontinuity of ex post preference (or utility) is of interest, and it is characterized in our theorem.

[15] The integral is well-defined by the Fubini Theorem because $(x, v) \longmapsto \max_{\beta \in x} v(\beta)$ is usc, hence product measurable.

[16] Nehring [16, p. 108] asserts that his representation theorem generalizes to the infinite case, and gives a very brief and incomplete sketch of how to achieve it. Our approach is different.

To illustrate the flexibility of part (a), consider the special case where Y consists of all convex menus (not containing β_*). Then the representing charge μ assigns probability 1 to the set of (usc and) quasiconcave utilities. The model thus suggests perfect certainty that ex post preference will be convex, but uncertainty about which convex preference will apply ex post. This special case is of particular importance given the argument by Epstein, Marinacci and Seo [7] that randomization may be valuable ex post. In fact, the cited paper *applies* Theorem 2.1, for a particular specification in which Y consists of all convex menus satisfying some added conditions, in order to axiomatize their (second) model of decision-making with coarse contingencies (or ambiguity, depending on which interpretation is preferred).

Turn to part (b) of the theorem. By definition, any Y-representation μ assigns probability 1 to V^Y, that is, to payoff functions whose upper contour sets all lie in Y. We have interpreted this condition as reflecting agent's certainty about an aspect of her future preferences; for example, that ex post preferences will be convex. However, this begs the question whether a similar interpretation is justified for all other representations. Thus part (b) begins with any representation, not necessarily carried by V^Y. The conclusion is that indeed, μ' is also carried by V^Y, and thus certainty about upper contour sets lying in Y is a property of preference and not just of a particular representation.

A final observation is that the utility function in (2.7) is upper semicontinuous on $\Delta(\mathcal{K}(\mathcal{B}))$.[17] Two implications follow. First, not only is (hypothetical or notional) ex post choice out of menus well-defined, but so also is the ('real' or part of the formal model) ex ante choice out of any compact feasible set of random menus. Secondly, since our axioms characterize the functional form, they necessarily imply upper semicontinuity of preference.

3. WHAT IS REVEALED BY THE RANKING OF RANDOM MENUS?

We have seen that the representing measure μ provided by Theorem 2.1 is not unique. Before describing what is uniquely determined by preference, it is useful to consider first the reasons for the nonuniqueness of μ. One reason that comes to

[17] By [14, Proposition D.7], $W(\cdot)$ is usc because it is monotone and right-continuous. Secondly, if $P_n \to P$, then $\limsup \int W(x)\, dP_n \leq \int W(x)\, dP$ by the nature of the weak convergence topology [1, Theorem 12.4]. Therefore, $P \longmapsto \int W(x)\, dP$ is usc.

mind is the state-dependence problem - one can always rescale each $v(\cdot)$ by a positive multiplicative constant a_v and then use the modified measure $d\mu' = d\mu/a_v$. However, such rescaling is ruled out by the normalizations of payoff functions in (2.4). Nonuniqueness arises here for another reason. We are conditioned to feel that preference 'should' reveal beliefs by Savage's celebrated theorem. However, think of the functional form $W(x) = \int_V \max_{\beta \in x} v(\beta)\, d\mu(v)$ as the subjective expected utility of the (real-valued) act f_x, $f_x(v) = \max_{\beta \in x} v(\beta)$, where the state space is V. Savage is able to determine a unique probability measure only by assuming that the agent ranks *all* acts over the state space. But here, the relevant set of acts $\{f_x : x \in \mathcal{K}(\mathcal{B})\}$ is only a 'small' proper subset (in fact, every f_x is usc and decreasing in the pointwise ordering on V). Thus one should not expect observable choice to determine a unique probability measure over states, and the focus should be rather to identify what is in fact pinned down by observable choice.

As illustrated by Figure A.1 and the surrounding discussion, intuition suggests that only the upper contour sets associated with ex post preferences, and not the latter per se, matter for ex ante choice. Accordingly, we will show that *the ranking $\succeq$ of random menus pins down, and is in turn completely determined by, the (suitably defined) distribution over ex post upper contour sets* that is implied by any representing μ. This will be shown to provide another perspective on and rationale for our choice of a canonical representation (see the discussion in the introduction).

For each $\mu \in \Delta(V)$, define a Borel probability measure $m_\mu \in \Delta(\mathcal{K}(\mathcal{B}))$, viewed as a measure over upper contour sets. Let $T_s : V \to \mathcal{K}(\mathcal{B})$ be given by

$$T_s(v) = \{\beta : v(\beta) \geq s\}, \text{ for each } s \in [0,1], \tag{3.1}$$

and define $m_\mu(\cdot)$, on the Borel σ-algebra, by

$$m_\mu(\cdot) = \int_0^1 \mu \circ T_s^{-1}(\cdot)\, dL(s), \tag{3.2}$$

where dL denotes the Lebesgue measure on $[0,1]$.

It is straightforward to see that m_μ is a well-defined measure: define Σ to be the collection of all Borel sets A such that $s \mapsto \mu \circ T_s^{-1}(A)$ is Lebesgue integrable. Because μ is countably additive, Σ is a σ-algebra. Moreover, Σ contains all sets of the form

$$\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\} \text{ for } x \in \mathcal{K}(\mathcal{B}). \tag{3.3}$$

This is because

$$\begin{aligned} T_s^{-1}\left(\{y : x \cap y \neq \varnothing\}\right) &= \{v \in V : x \cap \{\beta : v(\beta) \geq s\} \neq \varnothing\} \\ &= \left\{v \in V : \max_{\beta \in x} v(\beta) \geq s\right\}, \end{aligned}$$

which is Borel measurable because it is τ-closed; and because $\mu \circ T_s^{-1}\left(\{y : z \cap y \neq \varnothing\}\right)$ is nonincreasing on $[0,1]$, and thus (Riemann) integrable. But the Borel σ-algebra is the smallest σ-algebra containing all sets of the form (3.3) - see [1, Theorem 14.69] - and thus Σ contains all Borel sets.

One can interpret m_μ as summarizing the probability distribution of upper contour sets generated by 2–stage process. First, a utility level s is drawn from a uniform distribution over $[0,1]$, and then an ex post utility function v, (and thus also the upper contour set $T_s(v)$), is drawn according to μ. In short, m_μ is the "expected distribution" over upper contour sets induced by μ.

A simple example may be useful. Let $y_1 \subset y_2 \subset \mathcal{B}$, where $\beta_* \notin y_2$, and define

$$v(\cdot) = s_1 \mathbf{1}_{y_1}(\cdot) + (1 - s_1)\mathbf{1}_{y_2}(\cdot) = \begin{cases} 1 & \text{if } \beta \in y_1 \\ 1 - s_1 & \beta \in y_2 \backslash y_1 \\ 0 & \text{otherwise.} \end{cases}$$

Then v lies in V and has upper contour sets y_1, y_2 and $\mathcal{B}$. Setting $\mu = \delta_v$ gives a representation for the preference with utility function

$$\mathcal{W}(P) = \int_{\mathcal{K}(\mathcal{B})} \max_{\beta \in x} \left(s_1 \mathbf{1}_{y_1}(\beta) + (1 - s_1)\mathbf{1}_{y_2}(\beta)\right) dP(x);$$

the 'natural' interpretation is that of certainty that the ex post payoff function will be v. Compute that m_μ has two points of support:

$$m_\mu(\{y_1\}) = s_1, \text{ and } m_\mu(\{y_2\}) = (1 - s_1). \tag{3.4}$$

The third upper contour set $\mathcal{B}$ receives no weight, reflecting the fact that it is common to all payoff functions in V and thus is not relevant to distinguishing our particular v.

But there is another representation μ' for the same preference: let μ' assign probability s_1 to v_1 and $(1 - s_1)$ to v_2, where $v_i(\cdot) = \mathbf{1}_{y_i}(\cdot)$, $i = 1, 2$. Then μ' defines, via (2.7), the same utility function $\mathcal{W}$ given above, because

$$\max_{\beta \in x} \left(s_1 \mathbf{1}_{y_1}(\beta) + (1 - s_1)\mathbf{1}_{y_2}(\beta)\right) = s_1 \max_{\beta \in x} \mathbf{1}_{y_1}(\beta) + (1 - s_1) \max_{\beta \in x} \mathbf{1}_{y_2}(\beta),$$

though it suggests a different interpretation - uncertainty about whether payoffs will be given by v_1 or by v_2. Note, however, that μ' and μ have in common the induced distribution over upper contour sets, that is, $m_{\mu'} = m_\mu$. The uniqueness of this induced distribution across all representations is established more generally in the next theorem.

A possibly puzzling feature of the example is that the measure m_μ defined in (3.4) involves the utility levels s_1 and s_2. It is important to keep in mind, however, that these are not ordinal values, but rather have unambiguous meaning in terms of the given preference over random menus. For example, for the preference order in the example, s_1 is the unique probability p such that the random menu $(\mathcal{B}, (1-p); \{\beta_*\}, p)$ is indifferent to receiving the menu $\{\beta\}$ with certainty, where β is any lottery in $y_2 \backslash y_1$. This illustrates that the adoption of a domain of *random* menus is crucial for our analysis (another illustration is given below).

The main result of this section is that *any* two representations for $\succeq$ (satisfying our axioms) generate the identical measure over upper contour sets.

Theorem 3.1. *Let the set Y satisfy conditions Y1 and Y2. Suppose that $\succeq$ satisfies our axioms and that μ is a representation. Then:*

(a) For all $x \in \mathcal{K}(\mathcal{B})$,

$$m_\mu\left(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}\right) = \int_V \max_{\beta \in x} v(\beta)\, d\mu(v). \tag{3.5}$$

(b) Let μ' be any other representation for $\succeq$. Then $m_{\mu'} = m_\mu$.

(c) For any representation μ', $m_{\mu'}(Y) = 1$.

Note that the ranking of (nonrandom) menus alone is not sufficient to pin down a unique measure on upper contour sets. For example, let y, y' be closed subsets, and let

$$\begin{aligned} m_2 &= \frac{1}{2}\delta_y + \frac{1}{2}\delta_{y'} \text{ and} \\ m_1 &= \frac{1}{3}\delta_y + \frac{1}{3}\delta_{y'} + \frac{1}{3}\delta_{y \cup y'}. \end{aligned} \tag{3.6}$$

Then m_1 and m_2 represent the same preference on $\mathcal{K}(\mathcal{B})$, via (3.5), but they imply different rankings on $\Delta(\mathcal{K}(\mathcal{B}))$. This is easy to see: let $W_i(x) =$
$m_i(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\})$, $i = 1, 2$. On the domain of nonrandom menus, W_2

assumes the values $0, \frac{1}{2}$ or 1 depending on whether the menu in question intersects none, one or both of y and y', while W_1 assumes $0, \frac{2}{3}$ or 1. Therefore, they are ordinally, but not cardinally, equivalent on $\mathcal{K}(\mathcal{B})$.

Proof: (a) Compute that

$$\begin{aligned} m_\mu(\{y : x \cap y \neq \varnothing\}) &= \int_0^1 \mu \circ T_s^{-1}(\{y : x \cap y \neq \varnothing\})\, ds \\ &= \int_0^1 \mu(\{v : x \cap \{\beta : v(\beta) \geq s\} \neq \varnothing\})\, ds \\ &= \int_0^1 \mu\left(\left\{v : \max_{\beta \in x} v(\beta) \geq s\right\}\right) ds \\ &= 1 - \int_0^1 \mu\left(\left\{v : \max_{\beta \in x} v(\beta) < s\right\}\right) ds \\ &= \int_0^1 s dF(s), \quad \left(F(s) = \mu\left(\left\{v : \max_{\beta \in x} v(\beta) < s\right\}\right)\right), \\ &= \int_V \max_{\beta \in x} v(\beta)\, d\mu(v), \end{aligned}$$

where the next to last equality follows from integration by parts, and the last by a change of variables.

(b) If μ and μ' are any two representations, then $\succeq$ can be represented as an expected utility function with vNM index W, $W(x) = \int \max_{\beta \in x} v(\beta)\, d\mu(v)$, and also with vNM index W', $W'(x) = \int \max_{\beta \in x} v(\beta)\, d\mu'(v)$. By the uniqueness properties of vNM utility, it follows that, for some $a > 0$ and $b \in \mathbb{R}$, and for all $x \in \mathcal{K}(\mathcal{B})$,

$$\int \max_{\beta \in x} v(\beta)\, d\mu'(v) = a \int \max_{\beta \in x} v(\beta)\, d\mu(v) + b. \tag{3.7}$$

Letting $x = \{\beta_*\}$ and $x = \mathcal{B}$ yields $0 = b$ and $1 = a + b$, which implies

$$\int \max_{\beta \in x} v(\beta)\, d\mu'(v) = \int \max_{\beta \in x} v(\beta)\, d\mu(v) \text{ for all } x \in \mathcal{K}(\mathcal{B}).$$

Hence, by (a),

$$m_{\mu'}(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) = m_\mu(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) \text{ for all } x \in \mathcal{K}(\mathcal{B}).$$

But these sets generate the Borel σ-algebra [1, Theorem 14.69] - hence $m_{\mu'} = m_\mu$.

(c) By Theorem 2.1(a), $\succeq$ admits a Y-representation μ''. Then, by (b), $m_{\mu'}(Y) = m_{\mu''}(Y) = \int_0^1 \mu'' \circ T_s^{-1}(Y)\, dL(s) = \int_0^1 \mu''(\{v \in V : T_s(v) \in Y\})\, dL(s) = \int_0^1 \mu''(V^Y)\, dL(s) = 1$. ■

The theorem proves not only that the implied distribution over ex post upper contour sets is unique (part (b)), but also that it contains all relevant information about preference - indeed, by (a), we can rewrite the utility function $\mathcal{W}$ from (2.7) in the form

$$\mathcal{W}(P) = \int m_\mu(\{y : x \cap y \neq \varnothing\})\ dP(x). \tag{3.8}$$

Evidently,

$$m_\mu(\{y : x \cap y \neq \varnothing\}) = \int \max_{\beta \in x} \mathbf{1}_y(\beta)\, dm_\mu(y). \tag{3.9}$$

Since each indicator function $\mathbf{1}_y(\cdot)$ is usc, indeed an element of V^Y, and since $\mathcal{K}(\mathcal{B})$ is homeomorphic to a subspace of $USC(\mathcal{B})$, m_μ can be viewed as a Y-representation.

This perspective on the theorem relates more explicitly to the discussion in the opening paragraph of this section concerning a unique canonical representation. The expression (3.9) suggests that uncertainty about ex post preferences is confined to binary utility functions. Though binary payoff functions are clearly very special, we feel that nevertheless the representation (3.9) is useful as a reduced form: the agent may view more general (nonbinary) payoff functions as possible, but, as we have seen, it is only the implied uncertainty about upper contour sets that matter for the ranking of random menus. Thus ultimately all that matters for observable behavior are these expectations regarding upper contour sets, or equivalently their indicator functions, which are captured by the representation (3.9).

Another implication of the theorem is worth emphasizing: *equation (3.2) describes all the representations μ corresponding to a fixed preference.* Let $\succeq$ satisfy our axioms. Then there exists a unique canonical representation m as in (3.8). Now let μ be any measure satisfying, on the Borel σ-algebra,

$$m(\cdot) = \int_0^1 \mu \circ T_s^{-1}(\cdot)\, dL(s).$$

Then μ represents, via (2.7), some preference $\succeq'$ over random menus. But m_μ defined by (3.2) also represents $\succeq'$. However, $m_\mu = m$ and thus $\succeq' = \succeq$. In other words, μ represents $\succeq$ *if and only if* it satisfies (3.2).

Finally, there is a sense in which the uniqueness proven in the theorem may seem not completely satisfactory. The theorem shows that every representation generates the same measure m over upper contour sets, but the definition of "representation" imposes *a priori* that β_* is worst according to all ex post payoff functions. While this interpretation has been adopted throughout, that does not justify restricting attention only to such representations. Thus, for the moment, broaden "representation" to include any probability measure μ' on $\widehat{V} = \{v \in USC(\mathcal{B}) : \max_{\beta \in \mathcal{B}} v(\beta) = 1\}$ such that (2.7) is a utility function for $\succeq$. Define $m_{\mu'}$ by the counterpart of (3.2),

$$m_{\mu'}(A) = \int_0^1 \mu\left(\{v \in \widehat{V} : \{\beta : v(\beta) \geq s\} \in A\right) dL(s).$$

Then the argument used in the proof of part (b) of the theorem leads, in the absence of the normalization involving β_*, to: for some $a > 0$ and $b \in \mathbb{R}$, and for all $x \in \mathcal{K}(\mathcal{B})$,

$$\begin{aligned} m_{\mu'}(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) &= am_\mu(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) + (1-a) \\ &= (am_\mu + (1-a)\delta_{\mathcal{B}})(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}). \end{aligned}$$

Since these sets generate the Borel σ-algebra [1, Theorem 14.69],

$$m_{\mu'} = am_\mu + (1-a)\,\delta_{\mathcal{B}}.$$

Thus while two induced measures may be distinct, they agree once conditioned on the complement of $\{\mathcal{B}\}$; and, while $m_{\mu'}(Y) \neq 1$ in general, all measures satisfy $m_{\mu'}(Y \cup \{\mathcal{B}\}) = 1$. These differences call for obvious and minor changes in our interpretations and discussions.

4. FLEXIBILITY AND FLEXIBILITY-RISK

Here we show that our model is useful for capturing intuitive forms of behavior having to do with flexibility. A limitation of the DLR model in this regard is pointed out, thereby establishing another sense in which our model is more robust.

Adopt the following notation: for the vNM index W provided by Ex Ante vNM, define

$$\Delta_{x_1} W(x) = W(x) - W(x \cup x_1), \text{ and}$$

$$\Delta_{x_2}\Delta_{x_1} W(x) = \Delta_{x_1} W(x) - \Delta_{x_1} W(x \cup x_2) =$$

$$W(x) - W(x \cup x_1) - (W(x \cup x_2) - W(x \cup x_1 \cup x_2)).$$

For later reference, define also, for every $n > 1$,

$$\Delta_{x_n} ... \Delta_{x_1} W(x) = \Delta_{x_{n-1}} W(x) - \Delta_{x_{n-1}} W(x \cup x_n). \tag{4.1}$$

Given the canonical representation m provided by Theorem 3.1, we have:

$$W(x) = m(\{y : y \cap x \neq \varnothing\}),$$

$$-\Delta_{x_1} W(x) = m(\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing\}), \text{ and}$$

$$-\Delta_{x_2} \Delta_{x_1} W(x) = m(\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing, y \cap x_2 \neq \varnothing\}). \tag{4.2}$$

For any agent satisfying our axioms, larger menus are weakly preferred, which we describe in terms of a demand for (or value of) flexibility. This demand is simply characterized, since $x \cup x_1 \succ x \iff -\Delta_{x_1} W(x) > 0$ and so

$$x \cup x_1 \succ x \iff m(\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing\}) > 0. \tag{4.3}$$

Thus the value of flexibility is summarized by properties of m on Σ_1, the collection of all sets of the form $\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing\}$ as x and x_1 vary over all menus.

Now compare the desire for flexibility of two agents. Let $\succeq_1$ and $\succeq_2$ be two preferences satisfying our axioms (with representing measures m_1 and m_2). Say that 2 *desires flexibility more than* 1 if

$$x \cup x_1 \succ_1 x \Longrightarrow x \cup x_1 \succ_2 x, \tag{4.4}$$

that is, if whenever 1 strictly values the flexibility afforded by $x_1 \backslash x$, then so does 2. Then it follows from (4.3) that 2 desires more flexibility than 1 *if and only if* m_1 is absolutely continuous with respect to m_2 on Σ_1 (abbreviated $m_1 << m_2$ on Σ_1).[18]

The desire for flexibility captures only one way in which flexibility enters into preference. Since preference is defined over random menus, we can also consider the attitude towards "risk in flexibility". Say that the agent is *averse to flexibility-risk* if, for all menus x, x_1 and x_2, with $x_1 \cap x_2 = \varnothing$,

$$\tfrac{1}{2}\delta_x + \tfrac{1}{2}\delta_{x \cup x_1 \cup x_2} \preceq \tfrac{1}{2}\delta_{x \cup x_1} + \tfrac{1}{2}\delta_{x \cup x_2}. \tag{4.5}$$

[18] For all $A \in \Sigma_1$, $m_2(A) = 0 \Longrightarrow m_1(A) = 0$.

Consider an agent with menu x and two alternative ways to provide more flexibility. In one, corresponding to the LHS lottery, she receives either no added flexibility or the large supplement $x_1 \cup x_2$, each with probability $\frac{1}{2}$. In the alternative on the RHS, the large supplement $x_1 \cup x_2$ is partitioned into the two pieces x_1 and x_2, so that the added flexibility is less variable across the two probability $\frac{1}{2}$ events, and hence less risky. Thus the indicated ranking reflects a preference for the less risky way to receive added flexibility.

The restriction $x_1 \cap x_2 = \varnothing$ is important for the intuition that (4.5) concerns risk alone. For example, if $x_1 = x_2$, and if the independence axiom is satisfied (as implied by Ex Ante vNM), then (4.5) implies (for all x and x_1)

$$\delta_x \preceq \delta_{x \cup x_1},$$

which is Monotonicity, or the value of flexibility, and has nothing to do with flexibility being risky. More generally, if x_1 and x_2 are not necessarily disjoint, then the ranking (4.5) reflects both the value of flexibility and the attitude towards flexibility-risk. To combine both properties, say that preference is 2-*alternating* if (4.5) is satisfied for *all* x, x_1 and x_2.[19] Thus a 2-alternating preference both values flexibility and is averse to flexibility-risk. The converse is also valid given Ex Ante vNM:

$$\delta_{x \cup x_1} \succeq \delta_{x \cup (x_1 \setminus x_2)} \Longrightarrow$$
$$\tfrac{1}{2}\delta_{x \cup x_1} + \tfrac{1}{2}\delta_{x \cup x_2} \succeq \tfrac{1}{2}\delta_{x \cup (x_1 \setminus x_2)} + \tfrac{1}{2}\delta_{x \cup x_2} \succeq \tfrac{1}{2}\delta_x + \tfrac{1}{2}\delta_{x \cup x_1 \cup x_2},$$

where we have applied in turn Monotonicity, independence and flexibility-risk aversion.

Every preference satisfying our axioms is 2-alternating (hence averse to flexibility-risk) - in fact, since (recall (2.2))

$$\tfrac{1}{2}\delta_{x \cup x_1} + \tfrac{1}{2}\delta_{x \cup x_2} \geq_Y \tfrac{1}{2}\delta_x + \tfrac{1}{2}\delta_{x \cup x_1 \cup x_2}$$

for any Y, then Y-Dominance alone implies (4.5). In terms of the functional form for utility, there is aversion to flexibility-risk iff $-\Delta_{x_2}\Delta_{x_1} W(x) \geq 0$ whenever x_1 and x_2 are disjoint, which by (4.2) is always the case. (Similarly for 2-alternating.)

Turn to comparative aversion. Given two preferences $\succeq_1$ and $\succeq_2$, say that 2 *is more averse to flexibility-risk than* 1 if whenever (4.5) holds strictly for 1, then it also holds strictly for 2. Therefore, 2 *is more averse to flexibility-risk than* 1 if and only if $m_1 << m_2$ on Σ_2, the collection of all sets of the form

[19] We are adapting terminology from the theory of capacities - see [14, p. 7], for example.

$\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing,\ y \cap x_2 \neq \varnothing\}$ as x, x_1 and x_2 vary over all menus such that x_1 and x_2 are disjoint.

Conceptually, the desire for flexibility and aversion to flexibility-risk seem distinct. Since these properties are satisfied universally in our model, we cannot accommodate, for example, an agent who values flexibility and is not averse to flexibility-risk. However, our model does permit a distinction between the corresponding comparative notions. For example, it is possible for 2 to desire flexibility more than 1 and not be more averse to flexibility-risk than 1. Take the example in (3.6), where

$$m_2 = \frac{1}{2}\delta_y + \frac{1}{2}\delta_{y'} \text{ and}$$
$$m_1 = \frac{1}{3}\delta_y + \frac{1}{3}\delta_{y'} + \frac{1}{3}\delta_{y \cup y'}.$$

Then $m_1 << m_2$ on Σ_1, that is,

$$m_2\left(\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing\}\right) = 0 \Longrightarrow$$

$$m_1\left(\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing\}\right) = 0,$$

but $m_1 \not<\not< m_2$ on Σ_2: let

$$y \cap x = \varnothing,\ y \cap x_1 = \varnothing,\ y \cap x_2 \neq \varnothing, \text{ and}$$

$$y' \cap x = \varnothing,\ y' \cap x_1 \neq \varnothing,\ y' \cap x_2 = \varnothing.$$

However, a distinction is *not* possible within the DLR model (with a finite subjective state space), where, if 2 desires flexibility more than 1, then 2 is necessarily more averse to flexibility-risk than 1. To accommodate the DLR model within our framework, we assume here that $\mathcal{B} = \Delta(B)$ is the set of lotteries over a finite outcome set B.

Theorem 4.1. *Assume that $\mathcal{B} = \Delta(B)$ is the set of lotteries over a finite outcome set B. Let $\mu_1, \mu_2 \in \Delta(V)$ be two representations, such that each support $supp(\mu_i)$ is finite and consists of linear functions. Let $m_1, m_2 \in \Delta(\mathcal{K}(\mathcal{B}))$ be the corresponding canonical representations provided by Theorem 3.1. Then $m_1 << m_2$ on the Borel σ-algebra if $m_1 << m_2$ on either Σ_1 or Σ_2. Consequently, $m_1 << m_2$ on Σ_1 if and only if $m_1 << m_2$ on Σ_2.*

Proof : First, we show that

$$supp\,(\mu_1) \subset supp\,(\mu_2) \tag{4.6}$$

implies $m_1 << m_2$ on the Borel σ-algebra. (This is true even without linearity of ex post payoff functions.) By (3.2), for any Borel set A,

$$m_2\,(A) = 0 \Longrightarrow \int_0^1 \mu_2 \circ T_s^{-1}\,(A)\,dL\,(s) = 0 \Longrightarrow$$

$$\sum_{v \in supp(\mu_2)} \mu_2\,(\{v\})\,\mathbf{1}_{T_s^{-1}(A)}(v) = \mu_2 \circ T_s^{-1}\,(A) = 0,$$

for all $s \in E$, where $L\,(E) = 1$. Thus $\mathbf{1}_{T_s^{-1}(A)}(v) = 0$ for all $v \in supp\,(\mu_2)$ and $s \in E$. Therefore,

$$\mu_1 \circ T_s^{-1}\,(A) = \sum_{v \in supp(\mu_1)} \mu_1\,(\{v\})\,\mathbf{1}_{T_s^{-1}(A)}(v) = 0$$

for all $s \in E$, which implies that $m_1\,(A) = \int_0^1 \mu_1 \circ T_s^{-1}\,(A)\,dL\,(s) = 0$.

Suppose that $m_1 << m_2$ on Σ_1. Then, as argued above, 2 desires more flexibility than 1. But then DLR's Theorem 2 implies (4.6), and hence absolute continuity $m_1 << m_2$ on the entire Borel σ-algebra.

Finally, assume that there exists $v^* \in supp\,(\mu_1) \backslash supp\,(\mu_2)$ and show that m_1 is not absolutely continuous with respect to m_2 on Σ_2. Take interior points $\beta' \neq \beta'' \in \mathcal{B}$ such that

$$v^*(\beta') = v^*(\beta'') > 0.$$

(Since v^* is linear and not constant, this is possible.) For $\epsilon > 0$, let

$$x^\epsilon = \{\beta : v^*\,(\beta) \leq v^*\,(\beta') - \epsilon\}\,.$$

Then x^ϵ is nonempty for small enough $\epsilon > 0$. Let

$$D^\epsilon \equiv \{y \in supp\,(m_1) \cup supp\,(m_2) : y \cap x^\epsilon = \varnothing, \beta' \in y, \beta'' \in y\}\,.$$

We claim that, for each $v \neq v^*$ in $supp\,(\mu_1) \backslash supp\,(\mu_2)$, there exists $\epsilon^v > 0$ such that

$$T_s\,(v) \notin D^\epsilon \text{ for all } s \in (0,1] \text{ and } \epsilon < \epsilon^v.$$

To construct ϵ^v, note that, because $v \neq v^*$, there exists $\gamma \in \mathcal{B}$ such that $v(\gamma) = v(\beta')$ and $v^*(\gamma) < v^*(\beta')$. Set $\epsilon^v = v^*(\beta') - v^*(\gamma)$. Now it suffices to show that for all $s \in (0,1]$ and $\epsilon < \epsilon^v$,

$$\beta' \in T_s(v) \Rightarrow T_s(v) \cap x^\epsilon \neq \varnothing.$$

Since $v^*(\gamma) = v^*(\beta') - \epsilon^v < v^*(\beta') - \epsilon$ for all $\epsilon < \epsilon^v$, we have $\gamma \in x^\epsilon$ for all $\epsilon < \epsilon^v$. Thus

$$\beta' \in T_s(v) \Rightarrow v(\beta') \geq s \Rightarrow v(\gamma) \geq s \Rightarrow \gamma \in T_s(v),$$

and $\gamma \in T_s(v) \cap x^\epsilon \neq \varnothing$. This proves the claim.

To proceed, if $supp(\mu_1) \backslash supp(\mu_2) = \{v^*\}$, let $\bar{\epsilon}$ be any positive number. If not, let $\bar{\epsilon} = \min\{\epsilon^v : v \in supp(\mu_1) \backslash supp(\mu_2), v \neq v^*\}$. Since $supp(\mu_1) \backslash supp(\mu_2)$ is finite, $\bar{\epsilon} > 0$. Fix positive $\epsilon < \bar{\epsilon}$. Then $D^\epsilon \subset \{T_s(v^*) : s \in (0,1]\}$, by the above claim. More specifically,

$$D^\epsilon = \{T_s(v^*) : s \in (v^*(\beta') - \epsilon, v^*(\beta')]\}, \text{ and}$$

$$\begin{aligned} & m_1(\{y \in \mathcal{K}(\mathcal{B}) : y \cap x^\epsilon = \varnothing, y \cap \{\beta'\} \neq \varnothing, y \cap \{\beta''\} \neq \varnothing\}) \\ = \; & m_1(D^\epsilon) = \int_0^1 \mu_1 \circ T_s^{-1}(D^\epsilon)\, dL(s) = \int_{v^*(\beta')-\epsilon}^{v^*(\beta')} \mu_1(\{v^*\})\, dL(s) > 0, \end{aligned}$$

where the first equality is due to $D^\epsilon \subset supp(m_1)$. But since $T_s(v^*) \in \mathcal{K}(\mathcal{B}) \backslash supp(m_2)$, we have $D^\epsilon \subset \mathcal{K}(\mathcal{B}) \backslash supp(m_2)$ and

$$\begin{aligned} & m_2(\{y : y \cap x^\epsilon = \varnothing, y \cap \{\beta'\} \neq \varnothing, y \cap \{\beta''\} \neq \varnothing\}) \\ = \; & m_2(D^\epsilon) = 0. \end{aligned}$$

Since D^ϵ lies in Σ_2, m_1 is not absolutely continuous with respect to m_2 on Σ_2.

It follows that if $m_1 << m_2$ on Σ_2, we have (4.6) and hence absolute continuity on the entire Borel σ-algebra.[20] ■

It is possible to define comparative notions corresponding to higher order aspects of flexibility and to obtain similar characterizations in terms of the canonical representations m_1 and m_2. (At level n, the characterization would involve

[20] In fact, since β'' does not play much of a role in the preceding, the argument is readily adapted to provide an alternative proof of the implication of absolute continuity on Σ_1.

absolute continuity of the m_i's on Σ_n, the collection of all sets of the form $\{y : y \cap x = \varnothing,\ y \cap x_1 \neq \varnothing,\ ..., y \cap x_n \neq \varnothing\}$ as $x, x_1, ..., x_n$ vary over all menus satisfying some constraints.) Since intuition about these higher order "moments" and also their economic significance seem weaker, we do not pursue them here. Rather we conclude by describing an implication of absolute continuity on the entire Borel σ-algebra.

Say that 2 is *more Y-alternating than* 1 if

$$P' \geq_Y P \implies [P' \succ_1 P \Longrightarrow P' \succ_2 P].$$

If 2 is more Y-alternating than 1, then 2 desires more flexibility than 1 and 2 is also more averse to flexibility risk. (This follows from the facts that $\delta_{x \cup x_1} \geq_Y \delta_x$ and $P' \geq_Y P$ for the random menus defined in (2.2).) Absolute continuity on the Borel σ-algebra is sufficient for comparative Y-alternating.[21]

Theorem 4.2. *If $m_1 << m_2$ on the Borel σ-algebra, then 2 is more Y-alternating than 1.*

Proof : By (3.8), we can write $W_i(P) =$

$$\begin{aligned} \int\int \max_{\beta \in x} \mathbf{1}_y(\beta)\, dm_i(y)\, dP(x) &= \int\int \max_{\beta \in x} \mathbf{1}_y(\beta)\, dP(x)\, dm_i(y) \\ &= \int P(\{x : x \cap y \neq \varnothing\})\, dm_i(y). \end{aligned}$$

Therefore, $W_1(P') - W_1(P) =$
$\int (P'(\{x : x \cap y \neq \varnothing\}) - P(\{x : x \cap y \neq \varnothing\}))\, dm_1(y) > 0 \implies$
$P'(\{x : x \cap y \neq \varnothing\}) - P(\{x : x \cap y \neq \varnothing\}) > 0$ on a set $A \subset Y$, $m_1(A) > 0 \implies$
$P'(\{x : x \cap y \neq \varnothing\}) - P(\{x : x \cap y \neq \varnothing\}) > 0$ on a set $A \subset Y$, $m_2(A) > 0 \Longrightarrow$
$W_2(P') - W_2(P) > 0$. ■

A. APPENDIX: Proof of Theorem 2.1

Some standard notation is adopted: For any metric space X, $ba(X)$ denotes the set of finitely additive signed measures (or charges) of finite variation on the Borel σ-algebra; $ba_+(X)$ is the subset of positive charges and $ba_+^1(X) = \{\mu \in ba_+(X) : \mu(X) = 1\}$. Similarly for the sets of countably additive measures $ca(X)$, $ca_+(X)$ and $ca_+^1(X)$. Note that $ca_+^1(X) = \Delta(X)$.

[21] Necessity is an open question.

Lemma A.1. *If Y satisfies conditions Y1 and Y2, then V^Y is a measurable subset of V.*

Proof : Let $T_s : V \to \mathcal{K}(\mathcal{B})$ be defined as in (3.1), $T_s(v) = \{\beta : v(\beta) \geq s\}$, $0 < s \leq 1$. Since T_s is measurable (see the arguments following (3.1)), $T_s^{-1}(Y)$ is measurable for each s. Note that

$$V^Y = \cap_{s\in(0,1]} T_s^{-1}(Y).$$

Therefore, it suffices to show that

$$\cap_{s\in(0,1]} T_s^{-1}(Y) = \cap_{\kappa\in\mathbb{Q}\cap(0,1]} T_\kappa^{-1}(Y), \tag{A.1}$$

where $\mathbb{Q}$ denotes the set of rational numbers. Evidently, only $\supset$ requires proof. Let $v \in \cap_{\kappa\in\mathbb{Q}\cap(0,1]} T_\kappa^{-1}(Y)$, that is,

$$\{\beta : v(\beta) \geq \kappa\} \in Y \text{ for every } \kappa \in \mathbb{Q} \cap (0,1],$$

and let $0 < s \leq 1$. Take a sequence of rationals $\kappa_n \nearrow s$. Then

$$y_n = \{\beta : v(\beta) \geq \kappa_n\} \searrow y = \{\beta : v(\beta) \geq s\}, \tag{A.2}$$

which belongs to $cl(Y)$. Moreover, $\beta_* \notin y$ because $s > 0$. Therefore, $y \in cl(Y) \cap \mathcal{K}_- = Y$. Hence $v \in T_s^{-1}(Y)$, proving (A.1). ■

Proof of Theorem 2.1 (a):

NECESSITY: Y-Dominance: As argued in the proof of Theorem 4.2, utility can be written in the form

$$W(P) = \int P(\{x : x \cap y \neq \varnothing\})\, dm_\mu(y).$$

Therefore, $W(P') - W(P) =$
$$\int (P'(\{x : x \cap y \neq \varnothing\}) - P(\{x : x \cap y \neq \varnothing\}))\, dm_\mu(y) \geq 0 \text{ if } P' \geq_Y P.$$

SUFFICIENCY: By Ex Ante vNM, $\succeq$ can be represented by

$$\mathcal{W}(P) = \int_{\mathcal{K}(\mathcal{B})} W(x)\, dP(x),$$

for some bounded and measurable $W : \mathcal{K}(\mathcal{B}) \rightarrow \mathbb{R}$. By Y1, $x \geq_Y \{\beta_*\}$ for every menu x. Therefore, $x \succeq \{\beta_*\}$ by Y-Dominance. Since the latter implies Monotonicity, $x \preceq \mathcal{B}$. Conclude, by Nondegeneracy, that $\{\beta_*\} \prec \mathcal{B}$. Therefore, we can normalize W so that

$$W(\{\beta_*\}) = 0 \text{ and } W(\mathcal{B}) = 1.$$

The remainder of the proof consists in showing that W has the form

$$W(x) = \int_V \max_{\beta \in x} v(\beta)\, d\mu(v),$$

for some $\mu \in \Delta(V)$ with $\mu\left(V^Y\right) = 1$.

For each $\lambda \in ca(\mathcal{K}(\mathcal{B}))$, define the functional ϕ_λ on V^Y by

$$\phi_\lambda(v) = \int \max_{\beta \in x} v(\beta)\, d\lambda(x).$$

Then each ϕ_λ is bounded and measurable.

In particular, ϕ_P is defined thereby for each P in $\Delta(\mathcal{K}(\mathcal{B})) = ca_+^1(\mathcal{K}(\mathcal{B}))$. The purpose of this definition can be understood by recalling DLR's arguments for their setting where preference is defined over menus. After showing that their axioms imply that any menu x is indifferent to its convex hull $co(x)$, they employ the one-to-one relation between convex menus and support functions. There are two major differences here. First, because $co(x) \succ x$ is possible, nonconvex menus must also be considered, and these cannot be distinguished by the usual (linear-based) support functions of the theory of convexity. Second, the basic objects of choice here are lotteries over menus rather than menus. The needed modification of the DLR argument, expressed in the next lemma, is to associate each lottery P with its "expected support function" ϕ_P. Part (ii) describes this correspondence; the other parts assert that the mapping $P \longmapsto \phi_P$ preserves the order, mixture space and topological structures, and (iv) translates the normalization (2.4).

Lemma A.2. *(i)* $\phi_{P'}(\cdot) \geq \phi_P(\cdot) \Longleftrightarrow P' \geq_Y P$.
(ii) $\phi_{P'}(\cdot) = \phi_P(\cdot) \Longrightarrow P' \sim P$.
(iii) $\phi_{\alpha P + (1-\alpha)P'}(\cdot) = \alpha\phi_P(\cdot) + (1-\alpha)\phi_{P'}(\cdot)$.
(iv) $\phi_{P_*}(\cdot) = 0$ *and* $\phi_{P^*}(\cdot) = 1$, *where* $P_* = \delta_{\{\beta_*\}}$ *and* $P^* = \delta_{\mathcal{B}}$.

Proof : (ii) is clear if (i) holds. (iii) and (iv) follow from the definition of ϕ_P and (2.4). It remains to prove (i).

$\Longrightarrow$: Suppose that

$$\int \max_{\beta \in x} v(\beta)\, dP'(x) \geq \int \max_{\beta \in x} v(\beta)\, dP(x) \text{ for all } v \in V^Y.$$

Since $\mathbf{1}_y \in V^Y$ for each $y \in Y$, it follows that for every $y \in Y$,

$$P'(\{x \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) \geq P(\{x \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}).$$

$\Longleftarrow$: From the necessity part of the theorem, $\mathcal{W}(P) = \int_{\mathcal{K}(\mathcal{B})} \left(\int_{V^Y} \max_{\beta \in x} v(\beta)\, d\mu(v) \right) dP(x)$ satisfies Y-Dominance for any positive charge μ. Letting $\mu = \delta_v$ for $v \in V^Y$ gives the required property. ■

Denote by ϕ the homeomorphism between $\Delta(\mathcal{K}(\mathcal{B}))$ and $\Phi^0 \subset B_b(V^Y)$,

$$\Phi^0 = \{\phi_P : P \in \Delta(\mathcal{K}(\mathcal{B}))\} \subset B_b(V^Y),$$

that is provided by the lemma.[22] Now mimic DLR. Define W^0 on Φ^0 so that the following diagram commutes:

$$\begin{array}{ccccc} \Delta(\mathcal{K}(\mathcal{B})) & \xrightarrow{\phi} & \Phi^0 & \subset & B_b(V^Y) \\ & \searrow_{\mathcal{W}} & \downarrow \mathcal{W}^0 & & \swarrow \widehat{\mathcal{W}} \\ & & \mathbb{R}^1 & & \end{array}$$

In other words, let

$$\mathcal{W}^0(\phi_P) = \mathcal{W}(P).$$

Lemma A.2(ii) implies that $\mathcal{W}^0$ is well-defined. Note further that

$$\mathcal{W}^0(0) = \mathcal{W}\left(\delta_{\{\beta_*\}}\right) = 0.$$

The next step of the proof amounts to showing that $\mathcal{W}^0$ can be extended to a positive linear functional on all of $B_b(V^Y)$. Then application of a Riesz representation theorem delivers a charge $\mu \in ba_+(V^Y)$. One important difference from DLR is that their model leads to linear functionals on $C_b(S)$ for some compact space S, whereas we are dealing with linear functionals on $B_b(V^Y)$, and V^Y is

[22] $B_b(V^Y)$ denotes the set of bounded measurable functions on V^Y.

normal Hausdorff but not compact (or even locally compact). Thus the appropriate Riesz representation theorem delivers only a charge in our case, while DLR obtain a countably additive measure (compare Theorems 11.38 and 11.41 in [1]).

Extend $\mathcal{W}^0$ in steps from the domain Φ^0. Let $r\Phi^0 = \{r\phi_P = \phi_{rP} : P \in \Delta(\mathcal{K}(\mathcal{B}))\}$,

$$\Phi^+ = \cup_{r\geq 0} r\Phi^0, \text{ and}$$

$$\begin{aligned}\Phi = \Phi^+ - \Phi^+ &= \{\phi_{rP} - \phi_{r'P'} : P, P' \in \Delta(\mathcal{K}(\mathcal{B})), \ r, r' \geq 0\} \\ &= \left\{\phi_{rP-r'P'} : P, P' \in \Delta(\mathcal{K}(\mathcal{B})), \ r, r' \geq 0\right\} \\ &= \{\phi_\lambda : \lambda \in ca(\mathcal{K}(\mathcal{B}))\}.\end{aligned}$$

Lemma A.3. *(i)* Φ^0 *is convex. (ii)* $0 \in \Phi^0$.
(iii) $\alpha\phi_P = \phi_{\alpha P + (1-\alpha)P_*} \in \Phi^0$ *for every* $0 \leq \alpha \leq 1$*, where* $P_* = \delta_{\{\beta_*\}}$.
(iv) For any ϕ_λ *in* Φ*, there exists* $a > 0$*, and* P, P' *in* $\Delta(\mathcal{K}(\mathcal{B}))$*, such that* $\phi_\lambda = \phi_{a(P-P')}$.
(v) Φ *is a vector subspace of* $B_b(V^Y)$.

Proof : These claims follow from Lemma A.2. For example, for (iv), let $\phi_\lambda = \phi_{rQ-r'Q'}$ and $a > \max\{r, r'\}$. Then $\phi_P = \frac{r}{a}\phi_Q$ and $\frac{r'}{a}\phi_{Q'}$ lie in Φ, by (iii), and $\phi_\lambda = \phi_{a(P-P')}$. ■

Extend $\mathcal{W}^0$ to $\mathcal{W}^1$ on Φ by linearity:

$$\mathcal{W}^1(\phi_\lambda) = r\mathcal{W}(P) - r'\mathcal{W}(P'), \text{ if } \lambda = rP - r'P',$$

or equivalently, $\mathcal{W}^1(\phi_\lambda) = \int W d\lambda$.

Lemma A.4. $\mathcal{W}^1$ *is a positive linear functional on* Φ.

Proof : To show that $\mathcal{W}^1$ is linear, note that

$$\begin{aligned}\mathcal{W}^1(\alpha\phi_\lambda + \alpha'\phi_{\lambda'}) &= \mathcal{W}^1\left(\phi_{\alpha\lambda+\alpha'\lambda'}\right) = \int W d(\alpha\lambda + \alpha'\lambda') \\ &= \alpha\int W d\lambda + \alpha'\int W d\lambda' = \alpha\mathcal{W}^1(\lambda) + \alpha'\mathcal{W}^1(\lambda').\end{aligned}$$

Now show that $\phi_\lambda \geq \mathbf{0} \Longrightarrow \mathcal{W}^1(\phi_\lambda) \geq 0$: By Lemma A.3(iv), $\phi_\lambda = \phi_{a(P-P')}$, and thus $\phi_\lambda \geq \mathbf{0} \Longrightarrow \phi_P \geq \phi_{P'} \Longrightarrow \mathcal{W}(P) \geq \mathcal{W}(P')$, by Lemma A.2(i) and Y-Dominance. Thus $\mathcal{W}^1(\phi_\lambda) = a(\mathcal{W}(P) - \mathcal{W}(P')) \geq 0$. ■

Lemma A.5. *$W(x)=\int_V \max_{\beta\in x} v(\beta)\,d\mu(v)$ for some $\mu\in ba_+^1\left(V^Y\right)$.*

Proof : Note that $B_b\left(V^Y\right)$ is a Riesz space with unit $\mathbf{1}$ and Φ is a vector subspace containing $\mathbf{1}$. Thus the positive linear functional $\mathcal{W}^1$ on Φ admits a positive linear extension $\widehat{\mathcal{W}}$ to $B_b\left(V^Y\right)$ [1, Corollary 6.32]. By the Riesz Representation Theorem [6, IV.5.1], there exists a Borel charge $\mu\in ba\left(V^Y\right)$ such that

$$\widehat{\mathcal{W}}(\phi)=\int_{V^Y}\phi(v)\,d\mu(v).$$

Since $\mu(F)=\widehat{W}(\mathbf{1}_F)$ for any measurable subset F of V^Y and since $\widehat{W}$ is positive, we have $\mu\in ba_+\left(V^Y\right)$.

Consequently,

$$W(x)=\mathcal{W}(\delta_x)=\widehat{\mathcal{W}}\left(\phi_{\delta_x}\right)=\int_{V^Y}\max_{\beta\in x} v(\beta)\,d\mu(v).$$

Recall that $W(\mathcal{B})=1$ and $\max_{\beta\in\mathcal{B}} v(\beta)=1$ for each $v\in V^Y$. Thus

$$\mu\left(V^Y\right)=\int_{V^Y}1d\mu(v)=\int_{V^Y}\max_{\beta\in\mathcal{B}} v(\beta)\,d\mu(v)=W(\mathcal{B})=1. \quad \blacksquare$$

The rest of the proof consists of invoking the Choquet Theorem to get a Borel measure on $\mathcal{K}(\mathcal{B})$, which turns out to be a Y-representation.

Lemma A.6. *For any $\mu\in ba_+^1(V)$, there exists a unique Borel probability measure m on $\mathcal{K}(\mathcal{B})$ such that, for every x,*

$$W(x)\equiv\int_V\max_{\beta\in x} v(\beta)\,d\mu(v)=m\left(\{y\in\mathcal{K}(\mathcal{B}):x\cap y\neq\varnothing\}\right). \tag{A.3}$$

Proof : W is right-continuous, that is,

$$\text{if } x_n\searrow x,\text{ then } W(x_n)\searrow W(x).$$

In addition, W is *completely alternating*, that is,

$$\Delta_{x_n}...\Delta_{x_1}W(x)\leq 0,$$

for every $n \geq 1$ and $x, x_1, ..., x_n \in \mathcal{K}(\mathcal{B})$; recall (4.1). Here is a verification: because $x \mapsto \max_{\beta \in x} v(\beta)$ is completely alternating for usc v [14, p.11], we have,

$$\Delta_{x_n} ... \Delta_{x_1} W(x) = \int \left(\Delta_{x_n} ... \Delta_{x_1} \left(\max_{\beta \in x} v(\beta) \right) \right) d\mu(v) \leq 0.$$

Note that $\mathcal{B}$ is compact Polish. By the Choquet Theorem [14, Theorem 1.13], there exists a unique measure m satisfying (A.3), defined on the Borel σ-algebra generated by the Fell topology on $\mathcal{K}(\mathcal{B})$. Since $\mathcal{B}$ is compact metric, the Fell topology is equivalent to the Hausdorff metric topology [1, Section 3.17]. This completes the proof. ■

Remark 1. *Following common terminology in the theory of capacities,* W *is* infinitely alternating *if*

$$W\left(\bigcap_{i=1}^{n} x_i\right) \leq \sum_{\{I : \varnothing \neq I \subseteq \{1,...,n\}\}} (-1)^{|I|+1} W\left(\bigcup_{i \in I} x_i\right), \tag{A.4}$$

for all $n \geq 2$ *and* $x_1, ..., x_n \in \mathcal{K}(\mathcal{B})$. *It is straightforward to show that* W *is completely alternating if and only if it is monotone and infinitely alternating.*

Lemma A.7. *If* $m \in \Delta(\mathcal{K}(\mathcal{B}))$ *satisfies (A.3) for some* $\mu \in ba^1_+\left(V^Y\right)$, *then* $m(Y) = 1$.[23]

Proof : First, we show that for an open or closed subset z of $\mathcal{B}$,

$$\begin{array}{l} \{y \in \mathcal{K}(\mathcal{B}) : z \cap y \neq \varnothing\} \subset \mathcal{K}(\mathcal{B}) \backslash Y \\ \Rightarrow m(\{y \in \mathcal{K}(\mathcal{B}) : z \cap y \neq \varnothing\}) = 0. \end{array} \tag{A.5}$$

Similarly to the proof of Theorem 3.1(a),

$$m(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) = \int_0^1 \mu \circ T_s^{-1}(\{y \in \mathcal{K}(\mathcal{B}) : x \cap y \neq \varnothing\}) \, ds$$

for each $x \in \mathcal{K}(\mathcal{B})$. Since $\mu \in ba^1_+\left(V^Y\right)$, (A.5) holds for closed z.

Next prove (A.5) assuming z is open. Since $\mathcal{B}$ is metrizable, there is a sequence z_n of closed sets such that $z_n \nearrow z$. Then, by the countable additivity of m,

$$\begin{aligned} & m(\{y \in \mathcal{K}(\mathcal{B}) : z \cap y \neq \varnothing\}) \\ = \; & \lim_n m(\{y \in \mathcal{K}(\mathcal{B}) : z_n \cap y \neq \varnothing\}) = 0. \end{aligned}$$

[23] Note that Y2 is used heavily in the proof.

The last equality comes from $\{y \in \mathcal{K}(\mathcal{B}) : z_n \cap y \neq \varnothing\} \subset \{y \in \mathcal{K}(\mathcal{B}) : z \cap y \neq \varnothing\} \subset \mathcal{K}(\mathcal{B}) \setminus Y$.

The sets (3.3) constitute a base for the Hausdorff metric topology - see [1, Lemma 3.66]. Define $\mathcal{K}_* = \{y \in \mathcal{K}(\mathcal{B}) : \beta_* \in y\}$. Then $\mathcal{K}_*$ and $Y \cup \mathcal{K}_*$ are both closed. Since $\mathcal{K}(\mathcal{B})$ is separable, the open set $\mathcal{K}(\mathcal{B}) \setminus (Y \cup \mathcal{K}_*)$ is the countable union of basic sets. Thus there exist open or closed subsets z_n of $\mathcal{B}$, such that

$$\begin{aligned} m(\mathcal{K}(\mathcal{B}) \setminus (Y \cup \mathcal{K}_*)) &= m\left(\bigcup_{n=1}^{\infty} \{y \in \mathcal{K}(\mathcal{B}) : z_n \cap y \neq \varnothing\}\right) \\ &\leq \sum_{n=1}^{\infty} m(\{y \in \mathcal{K}(\mathcal{B}) : z_n \cap y \neq \varnothing\}) = 0; \end{aligned}$$

equality with zero follows from $\mu\left(V^Y\right) = 1$ and the inclusions $\{y \in \mathcal{K}(\mathcal{B}) : z_n \cap y \neq \varnothing\} \subset \mathcal{K}(\mathcal{B}) \setminus (Y \cup \mathcal{K}_*) \subset \mathcal{K}(\mathcal{B}) \setminus Y$.

Finally, recall that $\mathcal{K}_* = \{y \in \mathcal{K}(\mathcal{B}) : \beta_* \in y\}$. Then,

$$\begin{aligned} 1 &\geq m(Y) = m(Y \cup \mathcal{K}_*) - m(\mathcal{K}_*) \\ &= 1 - m(\{y \in \mathcal{K}(\mathcal{B}) : \{\beta_*\} \cap y \neq \varnothing\}) \\ &= 1 - \mathcal{W}\left(\delta_{\{\beta_*\}}\right) = 1. \quad \blacksquare \end{aligned}$$

Since Y is embedded in V^Y by the identification $y \mapsto \mathbf{1}_y$, m in the previous Lemma can be viewed as an element of $\Delta\left(V^Y\right)$ and hence we have a Y-representation.

Proof of Theorem 2.1(b): By Theorem 3.1(c), $1 = m_{\mu'}(Y) = \int_0^1 \mu' \circ T_s^{-1}(Y)\, dL(s)$. Therefore, $\mu' \circ T_s^{-1}(Y) = 1$ for all $s \in E \subset (0,1]$, where $L(E) = 1$. There exists a countable subset E^* of E that is dense in $(0,1]$. (The open intervals can be enumerated $\{I_n\}$. For every open interval I_n, we can pick $e_n \in I_n \cap E$ - the intersection must be nonempty. Let $E^* = \{e_n\}$.) Since μ' is countably additive,

$$\mu' \circ \cap_{s \in E^*} T_s^{-1}(Y) = 1.$$

But $\cap_{s \in (0,1]} T_s^{-1}(Y) = \cap_{s \in E^*} T_s^{-1}(Y)$. (See the proof of (A.1); the latter refers to the special case where E^* is the set of rationals, but only the denseness of E^* is important.) Therefore,

$$\mu' \circ \cap_{s \in (0,1]} T_s^{-1}(Y) = 1.$$

Finally, note that $V^Y = \cap_{s \in (0,1]} T_s^{-1}(Y)$. $\blacksquare$

References

[1] C.D. Aliprantis and K.C. Border, *Infinite Dimensional Analysis*, Springer, 1994.

[2] S.H. Chew, A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox, *Econometrica* 51 (1983), 1065-1092.

[3] G. Choquet, Theory of capacities, *Annales de l'Institut Fourier* 5 (1953), 131-295.

[4] E. Dekel, An axiomatic characterization of preferences under uncertainty, *J. Econ. Theory* 40 (1986), 304-318.

[5] E. Dekel, B. Lipman and A. Rustichini, Representing preferences with a unique subjective state space, *Econometrica* 69 (2001), 891-934.

[6] N. Dunford and J.T. *Schwartz, Linear Operators, Part I: General Theory*, Wiley, 1988.

[7] L.G. Epstein, M. Marinacci and K. Seo, Coarse contingencies, 2006.

[8] L.G. Epstein and M. Peters, A revelation principle for competing mechanisms, *J. Econ. Theory* 88 (1999), 119-160.

[9] P. Fishburn, *Utility Theory for Decision Making*, John Wiley, 1970.

[10] J. M. Grandmont, Continuity properties of a von Neumann-Morgenstern utility, *J. Econ. Theory* 4 (1972), 45-57.

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

[12] D. Kreps, *Notes on the Theory of Choice*, Westview, 1988.

[13] 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.

[14] I. Molchanov, *Theory of Random Sets*, Springer, 2005.

[15] K. Nehring, Preference for flexibility and freedom of choice in a Savage framework, UC Davis Working Paper, 1996.

[16] K. Nehring, Preference for flexibility in a Savage framework, *Econometrica* 67 (1999), 101-119.

[17] K. Nehring, Diversity and the geometry of similarity, 1999.

[18] T. Norberg, Random capacities and their distributions, *Probab. Th. Rel. Fields* 73 (1986), 281-297.

[19] T. Rader, The existence of a utility function to represent preferences, *Rev. Ec. Stud.* 30 (1963), 229-232.

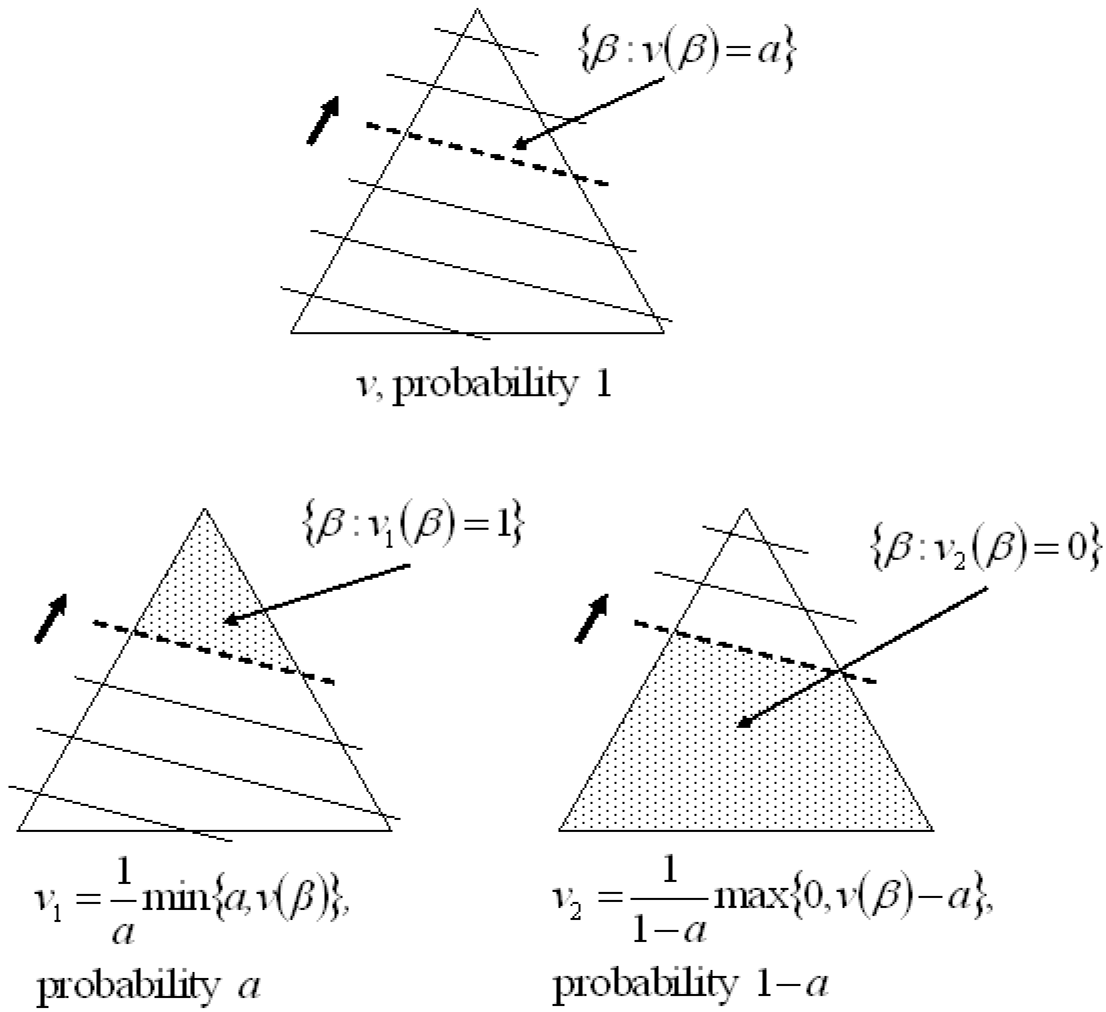

Figure A.1: Two subjective state spaces