Lexicographic Expected Utility with a Subjective State Space*

Youichiro Higashi Kazuya Hyogo[†]

August 25, 2008

Abstract

This paper provides a model that allows for a criterion of admissibility based on a subjective state space. For this purpose, we build a non-Archimedean model of preference with subjective states, generalizing Blume, Brandenburger, and Dekel [2], who present a non-Archimedean model with exogenous states; and Dekel, Lipman, and Rustichini [4], who present an Archimedean model with an endogenous state space. We interpret the representation as modeling an agent who has several "hypotheses" about her state space, and who views some as "infinitely less relevant" than others.

JEL classification: D81

Keywords: admissibility, subjective state space, non-Archimedean preferences, lexicographic expected utility.

1 Introduction

Admissibility has been widely used as a criterion of rationality in decision and game theory.[1] It is the requirement that "weakly dominated" actions should not be taken. That is, one action should be preferred to another if the outcome of the first action

*We are grateful to Larry Epstein for his patient encouragement and invaluable suggestions. We have benefited from comments by Sophie Bade, Barton Lipman, Yutaka Nakamura, Jawwad Noor, John Quiggin, and Norio Takeoka. We acknowledge helpful feedback from audiences of the 2006 JEA Spring Meeting, RUD 2006, SWET 2007, Hitotsubashi, Kyoto, and Toyama Universities. Hyogo gratefully acknowledges the financial support by KAKENHI (19830099).

[†]Higashi is at the Faculty of Economics, Hosei University, 4342 Aihara-cho, Machida, Tokyo 194-0298, Japan, yhigashi@hosei.ac.jp; Hyogo is at the Faculty of Economics, Ryukoku University, 67 Fukakusa Tsukamoto-cho, Fushimi-ku, Kyoto 612-8577, Japan, hyogo@econ.ryukoku.ac.jp.

[1]See for example, Arrow [1], Luce and Raiffa [11], Kohlberg and Mertens [8], and Brandenberger, Friedenberg, and Keisler [3].

is at least as good as that of the second action for each state, and strictly better for at least one state.

For the criterion of admissibility, it is crucial for the modeler to identify what uncertainties the agent perceives in her mind. However, the state space in the agent's mind is not directly observable to the modeler. The aim of this paper is to build a model that allows for the criterion of admissibility based on the subjective state space.

In the theory of subjective probability, Savage derives unique probability over objective states from preference and provides an axiomatic foundation for subjective expected utility theory. Subjective expected utility models satisfy admissibility only if there is no null state. This assumption is restrictive because such preferences would rule out pure strategy equilibria in games.[2] In an Anscombe–Aumann framework, Blume, Brandenburger, and Dekel [2] (henceforth, BBD) develop a non-Archimedean subjective probability model that allows for both the criterion of admissibility and "null" events, although not in the sense of Savage. In their model, the agent has a lexicographic hierarchy of subjective probabilities over objective states and may think that some states are "infinitely less relevant" than others. Unless two actions are indifferent in terms of all states in the first hierarchy, the agent does not care about outcomes in the other states. The agent thinks "null" states as infinitely less relevant, but does not entirely exclude them from consideration.

A restrictive feature of BBD is the exogenous state space. Kreps [9, 10] shows how the ranking of menus of alternatives reveals subjective uncertainty. Building on that, Dekel, Lipman, and Rustichini [4] (henceforth, DLR) endogenize the state space in an Archimedean framework. DLR take preference over menus of lotteries as a primitive and derive a unique subjective state space, corresponding to possible future preferences over lotteries. In this paper, we provide a non-Archimedean model with subjective states, which in principle enables us to use admissibility criterion based on the subjective state space. Our model is related to BBD in the same way that DLR is related to Savage.

As in DLR, this paper considers preference over menus of lotteries. By weakening their axiom Continuity, we provide a lexicographic representation $(S, U, \{\mu_k\}_{k=1}^K)$: is a tuple consisting of a nonempty finite state space S, a state dependent utility function $U : \Delta(B) \times S \to R$, and a hierarchy $\{\mu_k\}_{k=1}^K$ of (signed) measures such that for every menu x and y

$$x \succsim y \Leftrightarrow \left(\sum_{s \in S_k} \mu_k(s) \max_{\beta \in x} U(\beta, s) \right)_{k=1}^K \geq_L \left(\sum_{s \in S_k} \mu_k(s) \max_{\beta \in y} U(\beta, s) \right)_{k=1}^K ;$$

here $\Delta(B)$ is the set of lotteries over a finite set of prizes B, and $\geq_L$ compares each level of the hierarchy lexicographically. In the special case of our model where flexibility is valued, all μ_k's are positive measures.

[2]In complete information games, one can think of states as other agents' pure strategy profiles.

The interpretation of the representation above is as follows. The agent anticipates that after a state in S will be realized, she chooses the best lottery out of the menu. The difference from DLR is how she perceives subjective contingencies ex ante. That is, in her mind, the agent has multiple hypotheses about subjective states. The measure μ_1 indicates her primary hypothesis about subjective states. She has a secondary hypothesis μ_2. If two menus are indifferent according to her primary hypothesis, she uses the secondary hypothesis in order to compare the menus. She has a tertiary hypothesis, which is represented by μ_3, and so on. Since μ_k matters for the ranking of any two menus only if those menus are indifferent according to $\mu_1, \cdots, \mu_{k-1}$, we interpret that the hypothesis μ_k is thought of as "infinitely less relevant" than $\mu_1, \cdots, \mu_{k-1}$.

For uniqueness of the representation, the relevant part is the ex post preference $\succsim_s^*$ over $\Delta(B)$ determined by each $s \in S$. Under suitable conditions, we show uniqueness of the hierarchy of (incomplete) subjective state spaces $\{\succsim_s^* | \, s \in \cup_{j=1}^k \mathrm{supp}(\mu_j)\}_{k=1}^K$.

The organization of the paper is as follows. Section 2 introduces the DLR model. In section 3, we provide an example showing that Continuity is not always compelling. Section 4 states the main results. All proofs are collected in the appendix.

2 The DLR Model

DLR include the following primitives:

- B: finite set of prizes, let $|B| = B$
- $\Delta(B)$: set of probability measures over B, it is compact metric under the weak convergence topology; a generic element is denoted by β and referred to as a *lottery*
- $\mathcal{X}$: set of closed nonempty subsets of $\Delta(B)$, it is endowed with the Hausdorff topology; a generic element is denoted by x and called a *menu*[3]
- preference $\succsim$ is defined on $\mathcal{X}$

The interpretation is as follows: At time 0 (ex ante), the agent chooses a menu according to $\succsim$. At time 1 (ex post), a subjective state is realized and then she chooses a lottery out of the previously chosen menu. Note that the ex post stage is not a primitive of the formal model. However, since the agent is forward looking, her ex ante choice of menus reflects her subjective perception of states. Therefore, preference $\succsim$ over menus reveals a subjective state space.

The following are the main axioms in DLR.

[3]DLR do not restrict menus to be closed. If we allow any subset to be a menu, then we have to modify the definition of *critical set*. Under slight modification, all results remain the same.

Axiom 1 (Order)**.** $\succsim$ is complete and transitive.

We define the mixture of two menus for a number $\lambda \in [0,1]$ by

$$\lambda x + (1-\lambda)x' = \{\lambda\beta + (1-\lambda)\beta' | \beta \in x,\ \beta' \in x'\}.$$

The following is a version of the Independence Axiom adapted to a model with preference over menus.

Axiom 2 (Independence)**.** For all $x, y, z \in \mathcal{X}$ and $\lambda \in (0,1)$,

$$x \succsim y \Leftrightarrow \lambda x + (1-\lambda)z \succsim \lambda y + (1-\lambda)z.$$

Axiom 3 (Nontriviality)**.** There exist x and x' such that $x \succ x'$.

Axiom 4 (Continuity)**.** For every menu x, the sets $\{x' \in \mathcal{X} | x' \succsim x\}$ and $\{x' \in \mathcal{X} | x \succsim x'\}$ are closed.

The next axiom is introduced by Dekel, Lipman, and Rustichini [5] (henceforth, DLR2) to ensure, together with the other axioms, the finiteness of the state space. Let $\text{conv}(x)$ denote the convex hull of x.

Definition 1. A set $x' \subset \text{conv}(x)$ is *critical for* x if for all menus y with $x' \subset \text{conv}(y) \subset \text{conv}(x)$, we have $y \sim x$.

Axiom 5 (Finiteness)**.** Every menu has a finite critical subset.

The intuition is that when the agent faces a menu and contemplates future contingencies, she cares about only finite possibilities. Note that the set of states she cares about could depend on the menu. Therefore, this axiom does not imply finiteness of the subjective state space by itself.

Now, we explain a finite state space version of DLR's model. Let S be a state space. A function $U : \Delta(B) \times S \to R$ is a *state dependent utility function* if $U(\beta, s)$ has an expected utility form, that is, for $\beta \in \Delta(B)$,

$$U(\beta, s) = \sum_{b \in B} \beta(b) U(b, s).$$

Consider the functional form $W : \mathcal{X} \to R$ defined by

$$W(x) = \sum_{s \in S} \mu(s) \max_{\beta \in x} U(\beta, s), \tag{1}$$

where μ is a measure on S.

Note that S is just an index set though we call it the state space. Given the pair (S, U), define the ex post preference $\succsim_s^*$ over $\Delta(B)$ by

$$\beta \succsim_s^* \beta' \Leftrightarrow U(\beta, s) \geq U(\beta', s),$$

and let

$$P(S,U) = \{\succsim_s^* | s \in S\}.$$

Following DLR, we refer to the set of ex post preferences $P(S,U)$ as the *subjective state space.*

In general, there are many functional forms (1) that represent the same preference on $\mathcal{X}$. In order to obtain the uniqueness property, DLR concentrate on "relevant" subjective states: given a representation of the form (1), a state s is *relevant* if there exist menus x and y such that $x \nsim y$ and that for every $s' \neq s$, $\max_{\beta\in x} U(\beta, s') = \max_{\beta\in y} U(\beta, s')$.

Definition 2. A *finite additive representation* (S, U, μ) is a tuple consisting of a nonempty finite state space S, a state dependent utility function $U : \Delta(B) \times S \to R$, and a measure μ such that (i) $\succsim$ is represented by the functional form $W : \mathcal{X} \to R$, (ii) every state $s \in S$ is relevant, and (iii) if $s \neq s'$, then $\succsim_s^* \neq \succsim_{s'}^*$.

DLR and DLR2 prove

Theorem 2.1. *$\succsim$ satisfies Order, Independence, Nontriviality, Continuity, and Finiteness if and only if it has a finite additive representation.*

Corollary 2.2. *Suppose $\succsim$ has a finite additive representation. Then all finite additive representations of $\succsim$ have the same subjective state space.*

Axiom 6 (Monotonicity)**.** If $x \subset x'$, then $x' \succsim x$.

Monotonicity states that the agent values the flexibility of having more options. The consequence of Monotonicity is the following.

Corollary 2.3. *$\succsim$ satisfies Monotonicity and the axioms in Theorem 2.1 if and only if it has a finite additive representation with a positive measure μ.*

3 Continuity and a hierarchy of hypothesis

In this section, we argue that the axiom Continuity is not always compelling.

The intuition against Continuity is as follows: Suppose that a menu x is strictly preferred to a menu x'. Consider an agent who perceives some subjective contingencies and who has, in her mind, several hypotheses about these contingencies. Think of a hypothesis as a (singed) measure over contingencies that is used to weight the valuation of outcomes across states.[4] She may view one hypothesis as "infinitely less relevant" than another. Think of this as being captured by a hierarchy of hypotheses. Then there is a critical level k^* such that x and x' are indifferent according to each

[4]As explained later, a hypothesis in the formal model does not correspond to beliefs about states, and thus, we refer instead to "weights."

hypothesis at level k less than k^*, but x is strictly better than x' according to the hypothesis at level k^*. Now consider a "small" variation of x, denoted by x_ϵ. Then she should rank x_ϵ strictly better than x' using only the contingencies derived by the hypothesis at level k^*. However, the critical level for comparing x' and x_ϵ may be different than k^*; x' could be better than x_ϵ according to the hypothesis at the new critical level. Therefore, the small deviation might change the ranking between the menus.

The following examples are provided to illustrate this intuition.

Example: Consider an agent who used to like peanut butter very much, but who now has an allergy to peanut. Moreover, when she chooses an orange, she will pick the one that is more likely to be sweet.

There are three alternatives: the first one is an orange o_ϵ which turns out to be sweet with probability $0.9+\epsilon$ and sour with $0.1-\epsilon$; the second one is an orange o which turns out to be sweet with probability 0.9 and sour with 0.1; the last one is bread with peanut butter, which is denoted by p.

She may then have the following ranking: for every $\epsilon \in (0, 0.1]$

$$\{o_\epsilon\} \succ \{o, p\} \succ \{o\}. \tag{2}$$

The intuition is that she has two hypotheses for her allergy: the first is that the allergy continues, and the second is that the allergy disappears. However, she thinks that it is infinitely less relevant to take into account the possibility that her allergy disappears. That is, she would rank the two hypotheses hierarchically in her mind.

First, consider the first and second menus. Since flexibility provided by bread with peanut butter is irrelevant in the primary hypothesis, the ranking of the first and second menus follows the taste of orange. Hence, the agent prefers the first menu to the second one.

Next, consider the second and third menus. At first, the agent uses the primary hypothesis to rank the menus. Since the two menus are indifferent in the primary hypothesis, the ranking of menus in the secondary hypothesis is relevant for her choice among menus. Thus, she wants to retain the opportunity to have peanut butter. The agent prefers the second menu to the third one.

Ranking (2) violates Continuity.

4 Lexicographic Representation

In the previous section, the difficulties for Continuity arise out of the strict preference relation. Therefore, we impose "continuity" only for indifference sets.

Axiom 7 (Indifference Continuity)**.** For every menu x, the indifference set $\{x' \in \mathcal{X} | x' \sim x\}$ is closed.

There is no corresponding axiom in BBD. The reason is that BBD assume that the state space is exogenous and finite. In our model, the state space is derived endogenously from preference.

Since we weaken Continuity, a numerical representation is not always possible. We consider a lexicographic representation that compares a vector of utilities assigned to a menu by $\geq_L$.[5] More formally, let S and $U : \Delta(B) \times S \to R$ be a state space and a state dependent utility function. Consider the vector-valued function $V : \mathcal{X} \to R^K$ defined by

$$V(x) = \left(\sum_{s \in S} \mu_k(s) \max_{\beta \in x} U(\beta, s) \right)_{k=1}^{K}, \tag{3}$$

where $\{\mu_k\}_{k=1}^K$ is a hierarchy of measures. This vector-valued function is the counterpart of the DLR functional form (1). We also need a counterpart of "relevance": given a representation of the form (3), a state s is *relevant* if there exist menus x and y such that $x \not\sim y$ and that for every $s' \neq s$, $\max_{\beta \in x} U(\beta, s') = \max_{\beta \in y} U(\beta, s')$.

Definition 3. A *lexicographic representation* $\left(S, U, \{\mu_k\}_{k=1}^K\right)$ is a tuple consisting of a nonempty finite state space S, a state dependent utility function $U : \Delta(B) \times S \to R$, and a hierarchy $\{\mu_k\}_{k=1}^K$ of measures such that

$$\text{(i) } x \succsim y \Leftrightarrow \left(\sum_{s \in S} \mu_k(s) \max_{\beta \in x} U(\beta, s) \right)_{k=1}^{K} \geq_L \left(\sum_{s \in S} \mu_k(s) \max_{\beta \in y} U(\beta, s) \right)_{k=1}^{K},$$

(ii) every state $s \in S$ is relevant, and (iii) if $s \neq s'$, then $\succsim_s^* \neq \succsim_{s'}^*$.

The integer K is referred to as the *length* (of the hierarchy).

Now we state our main result:

Theorem 4.1. *$\succsim$ satisfies Order, Independence, Nontriviality, Indifference Continuity, and Finiteness if and only if it has a lexicographic representation.*

For interpretation, note that the ex post behavior is as in DLR: a state s in S will be realized at the beginning of time 1. Then she will choose the best alternative out of the previously chosen menu according to the ex post utility function $U(\cdot, s)$. Moreover, she anticipates this ex post behavior at time 0. The difference from DLR is how she perceives subjective contingencies ex ante. The agent has a hierarchy of measures in her mind. Each level of the hierarchy represents her hypothesis about how she should allow for the future contingencies ex ante. The measure μ_1 indicates her primary hypothesis. She has a secondary hypothesis, which is represented by μ_2. If menus are indifferent according to her primary hypothesis, she compares them according to her secondary hypothesis. She has a tertiary hypothesis, which is represented by μ_3, and

[5] For $a, b \in R^K$, $a \geq_L b$ if and only if whenever $b_k > a_k$, there is a $j < k$ such that $a_j > b_j$.

so on. Since μ_k enters into the ranking of any two menus x and y only if x and y are indifferent as to $\mu_1, \cdots, \mu_{k-1}$, the measure μ_k is relevant but may be thought of as being "infinitely less relevant" than $\mu_1, \cdots, \mu_{k-1}$.

To further illustrate the meaning of "μ_k is infinitely less relevant than μ_{k-1}," consider the special case where there is no overlap among the supports of μ_k's. Suppose that s_{k-1} and s_k belong to supp(μ_{k-1}) and supp(μ_k) respectively. Consider two menus x and y such that the agent expects the same ex post utilities at all states except s_{k-1} and s_k. Then the ex post ranking between x and y at s_{k-1} determines the ex ante ranking regardless of the ex post ranking at s_k. This leads us to say "s_k is infinitely less relevant than s_{k-1}." In the Archimedean case, as in DLR, every state is either relevant or not. Our model admits a richer comparison between subjective states. That is, there may be a state which is relevant but infinitely less relevant than another state.

Uniqueness of the representation does not hold in general. For example, the μ_k's are not uniquely determined by preference, just as in DLR. Secondly, there may be redundancies in the hierarchy [2, p. 66].

To express the uniqueness properties of our representation, define for each $k = 1, \ldots, K$,

$$P_k(S, U, \{\mu_k\}_{k=1}^K) = \{\succsim_s^* |s \in \cup_{j=1}^k \text{supp}(\mu_j)\} \subset P(S, U).$$

Following DLR, we can think of $\{P_k(S, U, \{\mu_k\}_{k=1}^K)\}_{k=1}^K$ as a hierarchy of (incomplete) subjective state spaces. Note that there is a lexicographic representation with minimal length K, denoted by K^*. To avoid the redundancies, we concentrate on lexicographic representations of minimal length K^*.

Corollary 4.2. *Suppose that $\succsim$ admits a lexicographic representation. Let $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$ and $\left(S', U', \{\mu'_k\}_{k=1}^{K^*}\right)$ be lexicographic representations of $\succsim$ with the minimal length K^*. Then, for $k = 1, \cdots, K^*$,*

$$P_k(S, U, \{\mu_k\}_{k=1}^{K^*}) = P_k(S', U', \{\mu'_k\}_{k=1}^{K^*}).$$

The next axiom is the difference between our model and DLR's finite additive representation.

Axiom 8 (Upper Semicontinuity)**.** For every menu x, the upper contour set $\{x' \in \mathcal{X} | x' \succsim x\}$ is closed.

Theorem 4.3. *$\succsim$ has a lexicographic representation and satisfies Upper Semicontinuity if and only if it has a finite additive representation (as in DLR).*

Since the axiom Indifference Continuity is uncommon, and perhaps original to this paper, we next verify that it is critical in Theorem 4.1, that is, it is not implied by the other axioms. Let $u, v : \Delta(B) \to R$ be continuous linear nonconstant functions.

Consider an order $\succsim^{nsc}$ over menus represented by the following functional form $U^{nsc}: \mathcal{X} \to R$ proposed in Gul and Pesendorfer [6].[6]

Temptation without self-control:

$$U^{nsc}(x) = \max_{\beta \in x} u(l) \text{ subject to } v(\beta) \geq v(\beta') \text{ for every } \beta' \in x.$$

It is easy to check that $\succsim^{nsc}$ satisfies Order, Independence, Nontriviality, and Finiteness. However, in general, it violates Indifference Continuity: let $B = \{b_1, b_2, b_3\}$, $u(\beta) = \beta_2$, and $v(\beta) = \beta_1$, where $\beta_i = \beta(b_i)$. Let $\bar{\beta}, \beta^*, \beta^n$ be lotteries such that

$$\begin{aligned}\bar{\beta} &= (\bar{\beta}_1, \bar{\beta}_2, \bar{\beta}_3)\\ \beta^* &= (\bar{\beta}_1, 1 - \bar{\beta}_1, 0)\\ \beta^n &= (\bar{\beta}_1 - \epsilon^n, 1 - \bar{\beta}_1 + \epsilon^n, 0) \text{ for } n \geq 1,\end{aligned}$$

where $\bar{\beta}_1, \bar{\beta}_2, \bar{\beta}_3 > 0$ and $\bar{\beta}_1 > \epsilon > 0$. The sequence of menus, $\left\{\{\bar{\beta}, \beta^n\}\right\}_{n=1}^{\infty}$, converges to the menu $\{\bar{\beta}, \beta^*\}$ in the Hausdorff topology and $\{\bar{\beta}, \beta^n\} \sim \{\bar{\beta}\}$ for every n. However, we have $\{\bar{\beta}, \beta^*\} \succ \{\bar{\beta}\}$, contradicting Indifference Continuity.[7]

Finally, if we add Monotonicity to the axioms in Theorem 4.1, then all measures are positive:

Corollary 4.4. $\succsim$ *satisfies Monotonicity and the axioms in Theorem 4.1 if and only if it has a lexicographic representation where all measures are positive.*

Appendix: Proofs

A Proof of Theorem 4.1

The necessity of the axioms is easily verified. We show only the sufficiency part.

To begin with, note that Finiteness implies that $x \sim \text{conv}(x)$ for every menu x. In fact, by Finiteness, every menu x has a critical subset x'. Moreover, $x' \subset \text{conv}(\text{conv}(x)) = \text{conv}(x)$. It follows from the definition of critical set that $x \sim \text{conv}(x)$. Thus, we can restrict attention to the set of closed, convex, nonempty subsets of $\Delta(B)$, denoted by $\mathcal{X}^*$.

Let $S^B = \{s \in R^B | \sum_i s_i = 0, \text{ and } \sum_i s_i^2 = 1\}$. For $x \in \mathcal{X}^*$, define a function $\sigma_x : S^B \to R$ by

$$\sigma_x(s) = \max_{\beta \in x} \beta \cdot s$$

The following is Lemma 5 in DLR2 [5]:

[6]The acronym *nsc* means *no self-control*.

[7]While $\succsim^{nsc}$ may violate Indifference Continuity, it satisfies Upper Semicontinuity (see [6, p.1413]).

Lemma 1. *There is a finite subset $S^* \subset S^B$ such that*

$$\forall x, y \in \mathcal{X}^*, \quad [\forall s \in S^*, \sigma_x(s) = \sigma_y(s)] \Rightarrow x \sim y. \tag{4}$$

Proof. In DLR2's argument, Continuity is required to prove their Lemma 3. In fact, Indifference Continuity is enough to show the result. □

If two finite subsets $S^*, S^{**} \subset S^B$ satisfy the condition (4), then so does the intersection $S^* \cap S^{**}$. In the following, we identify S^* as the smallest finite subset of S^B satisfying the condition (4).

Note that S^* is not empty by Nontriviality. Let m (≥ 1) be the cardinality of S^*. Appliying the above lemma, we embed $\mathcal{X}^*$ into m-dimensional vector space. Denote

$$\mathcal{M} = \left\{ (\sigma_x(s))_{s \in S^*} \mid x \in \mathcal{X}^* \right\} \subset R^m.$$

It is a closed convex subset of R^n and contains the 0 vector. Moreover, it is a mixture space in the sense of Hausner [7]. In the following, we identify σ_x with a corresponding element in $\mathcal{M}$.

Preference $\succsim$ on $\mathcal{X}^*$ induce $\succsim^*$ on a mixture space $\mathcal{M}$. That is, $\sigma_x \succsim^* \sigma_y$ if and only if $x \succsim y$. Since $\succsim$ satisfies Order and Independence, $\succsim^*$ also satisfies Order and Independence. The following lemma directly follows from the result of Hausner [7].

Lemma 2. *(i) $\succsim^*$ satisfies Order and Independence if and only if there are $K(\leqq n)$ affine functions $V_k : \mathcal{M} \to R$ such that*

$$\sigma_x \succsim^* \sigma_y \Leftrightarrow (V_k(\sigma_x))_{k=1}^K \geq_L (V_k(\sigma_y))_{k=1}^K.$$

(ii) Moreover, there is minimal K, denoted by K^*, less than or equal to n. $\{V'_k\}_{k=1}^{K^*}$ satisfy the above representation in place of $\{V_k\}_{k=1}^{K^*}$ if and only if there are real numbers $a_k > 0$, b_{kj} and c_k such that, for every $\sigma \in \mathcal{M}$,

$$V'_k(\sigma) = a_k V_k(\sigma) + \sum_{j=1}^{k-1} b_{kj} V_j(\sigma) + c_k.$$

Now, we follow again the argument in DLR2. Then, V_k has a well-defined extension to a continuous linear function on R^m. Moreover, there exists a vector $(\mu_k(s))_{s \in S^*} \in R^m$ such that

$$V_k(M) = \sum_{s \in S^*} \mu_k(s) M_s \text{ for every } M \in \mathcal{M}.$$

Define a state dependent utility function $U : \Delta(B) \times S^* \to R$ by $U(\beta, s) = \beta \cdot s$. By construction, every $s \in S^*$ is relevant and $\succsim^*_s \neq \succsim^*_{s'}$ if $s \neq s'$. Thus we have the desired representation $(S^*, U, \{\mu_k\}_{k=1}^K)$.

B Proof of Corollary 4.2

Definition 4. A lexicographic representation $\left(\bar{S}, \bar{U}, \{\bar{\mu}_k\}_{k=1}^{K^*}\right)$ is *canonical* if it satisfies

$$\bar{S} \subset S^B, \text{ and } \bar{U}(\beta, \bar{s}) = \beta \cdot \bar{s} \text{ for every } \beta \in \Delta(B).$$

Lemma 3. *For every* $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$, *there exists a canonical representation* $\left(\bar{S}, \bar{U}, \{\bar{\mu}_k\}_{k=1}^{K^*}\right)$ *such that (i)* $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$ *and* $\left(\bar{S}, \bar{U}, \{\bar{\mu}_k\}_{k=1}^{K^*}\right)$ *represent the same preference, and for every* $k = 1, \cdots, K^*$,

$$\textit{(ii) } P_k(S, U, \{\mu_k\}_{k=1}^{K^*}) = P_k(\bar{S}, \bar{U}, \{\bar{\mu}_k\}_{k=1}^{K^*}).$$

Proof. Let

$$c_s = \frac{1}{B} \sum_{b \in B} U(b, s) \text{ and } c_k = \sum_{s \in S} \mu_k(s) c_s.$$

Note that $U(\cdot, s)$ cannot be a constant function for every $s \in S$ since every $s \in S$ is relevant. Thus $\sum_{b \in B} (U(b, s) - c_s)^2$ has to be strictly positive. Define a function $\phi : S \to R^B$ by

$$\phi(s) = \left(\frac{U(b, s) - c_s}{\sum_{b' \in B} (U(b', s) - c_s)^2} \right)_{b \in B}.$$

It is straightforward that ϕ is one-to-one and $\phi(s)$ belongs to S^B for every $s \in S$.

Let $\bar{S} = \phi(S) \subset S^B$ and $\bar{U}(\beta, \bar{s}) = \beta \cdot \bar{s}$ for every $\bar{s} \in \bar{S}$. Define a measure $\bar{\mu}_k$ over S^M by

$$\bar{\mu}_k(\phi(s)) = \mu_k(s) \sum_{b \in B} (U(b, s) - c_s)^2.$$

By definition, $\succsim_s^* = \succsim_{\phi(s)}^*$ and $\phi(s) \in \operatorname{supp}(\bar{\mu}) \Leftrightarrow s \in \operatorname{supp}(\mu)$. Hence, for every $k = 1, \cdots, K$, $P_k(S, U, \{\mu_k\}_{k=1}^{K^*}) = P_k(\bar{S}, \bar{U}, \{\bar{\mu}_k\}_{k=1}^{K^*})$. Moreover, it holds that

$$\sum_{\phi(s) \in \bar{S}} \bar{\mu}_k(\phi(s)) \max_{\beta \in x} \bar{U}(\beta, \phi(s)) = \{\sum_{s \in S} \mu_k(s) \max_{\beta \in x} U(\beta, s)\} - c_k.$$

Part (ii) of Lemma 2 implies that $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$ and $\left(\bar{S}, \bar{U}, \{\bar{\mu}_k\}_{k=1}^{K^*}\right)$ represents the same preference. □

Lemma 4. *Let* $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$ *and* $\left(S', U', \{\mu'_k\}_{k=1}^{K^*}\right)$ *be canonical lexicographic representations of* $\succsim$ *with the minimal length* K^*. *Then, for* $k = 1, \cdots, K^*$,

$$P_k(S, U, \{\mu_k\}_{k=1}^{K^*}) = P_k(S', U', \{\mu'_k\}_{k=1}^{K^*}).$$

Proof. To begin with, we see that $S = S'$ (and hence $U = U'$). Suppose to the contrary that $S \neq S'$. Without loss of generality, we assume there exists $\bar{s} \in S \setminus S'$. Fix a sphere $y \subset \text{int}\Delta(B)$. Define

$$x = \cap_{s \in S \cup S' \setminus \{\bar{s}\}} \{\beta \in \Delta(B) |\ \beta \cdot s \leqq \max_{\beta' \in y} \beta' \cdot s\}.$$

Since both representations are canonical, it holds that

$$\begin{aligned}
&\max_{\beta \in x} U(\beta, \bar{s}) > \max_{\beta \in y} U(\beta, \bar{s}), \\
&\max_{\beta \in x} U(\beta, s) = \max_{\beta \in y} U(\beta, s) \text{ for every } s \in S, s \neq \bar{s}, \text{ and} \\
&\max_{\beta \in x} U'(\beta, s) = \max_{\beta \in y} U'(\beta, s) \text{ for every } s \in S'.
\end{aligned}$$

Hence, $x \nsim y$ according to $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$, but $x \sim y$ according to $\left(S', U', \{\mu'_k\}_{k=1}^{K^*}\right)$. This is a contradiction.

Finally, we show $P_k(S', U', \{\mu'_k\}_{k=1}^{K^*}) \subset P_k(S, U, \{\mu_k\}_{k=1}^{K^*})$. Then the other direction $P_k(S, U, \{\mu_k\}_{k=1}^{K^*}) \subset P_k(S', U', \{\mu'_k\}_{k=1}^{K^*})$ also holds since $(S, U, \{\mu_k\}_{k=1}^{K^*})$ and $(S', U', \{\mu'_k\}_{k=1}^{K^*})$ are symmetric. Define, as in the proof of Theorem 4.1,

$$\begin{aligned}
\mathcal{M} &= \left\{ (\sigma_x(s))_{s \in S} \mid x \in \mathcal{X}^* \right\}, \\
V_k(\sigma_x) &= \sum_{s \in S} \mu_k(s) \max_{\beta \in x} \beta \cdot s, \text{ and} \\
V'_k(\sigma_x) &= \sum_{s \in S} \mu'_k(s) \max_{\beta \in x} \beta \cdot s.
\end{aligned}$$

Since V_k and V'_k represents the same order, it follows from Part (ii) of Lemma 2 that

$$V'_k(\sigma_x) = a_k V_k(\sigma_x) + \sum_{j=1}^{k-1} b_{kj} V_j(\sigma_x) + c_k \text{ for } k = 1, \cdots, K^*.$$

Thus, $\text{supp}(\mu'_k) \subset \cup_{j=1}^{k} \text{supp}(\mu_j)$, and hence, for every $l \leqq k$,

$$\text{supp}(\mu'_l) \subset \cup_{j=1}^{l} \text{supp}(\mu_j) \subset \cup_{j=1}^{k} \text{supp}(\mu_j).$$

Therefore, $\cup_{j=1}^{k} \text{supp}(\mu'_j) \subset \cup_{j=1}^{k} \text{supp}(\mu_j)$. That is,

$$\begin{aligned}
P_k(S', U', \{\mu'_k\}_{k=1}^{K^*}) =& \{ \succsim_s^* | s \in \cup_{j=1}^{k} \text{supp}(\mu'_j) \} \\
\subset& \{ \succsim_s^* | s \in \cup_{j=1}^{k} \text{supp}(\mu_j) \} = P_k(S, U, \{\mu_k\}_{k=1}^{K^*})
\end{aligned}$$

□

C Proof of Theorem 4.3

Necessity is straightforward. We show only sufficiency.

Let $\left(S, U, \{\mu_k\}_{k=1}^{K^*}\right)$ be a canonical lexicographic representation of minimal length K^*. We see that if $K^* \geq 2$, then Upper Semicontinuity is violated. More specifically, we construct menus x and x' and a sequence of menus $\{x_n\}$ such that $x_n \to x$, $x_n \succ x'$ for every n, and $x' \succ x$.

Fix a sphere $y \subset \text{int}\Delta(B)$. Define

$$x' = \cap_{s \in \cup_{j=1}^2 \text{supp}(\mu_j)} \{\beta \in \Delta(B) | \, \beta \cdot s \leq \max_{\beta' \in y} \beta' \cdot s\}.$$

Case 1: $\text{supp}(\mu_1) \subsetneq \cup_{j=1}^2 \text{supp}(\mu_j)$. Pick a state $s' \in \text{supp}(\mu_2) \setminus \text{supp}(\mu_1)$.

Subcase 1-1: $\mu_2(s') > 0$. Let $\epsilon > 0$. Define

$$x = \cap_{s \in \cup_{j=1}^2 \text{supp}(\mu_j)} \{\beta \in \Delta(B) | \, \beta \cdot s \leq f_1(s)\},$$

where $f_1(s') = \max_{\beta' \in y} \beta' \cdot s' - \epsilon$ and $f_1(s) = \max_{\beta' \in y} \beta' \cdot s$ for $s \neq s'$. We take ϵ small enough so that x is a menu with $\max_{\beta \in x} \beta \cdot s = f_1(s)$ for every $s \in S$. This is possible since S is finite.

Take a state $s^* \in \text{supp}(\mu_1)$. First, we consider the case $\mu_1(s^*) < 0$. Let $\xi \in (0, 1)$. Define

$$x_n = \cap_{s \in \cup_{j=1}^2 \text{supp}(\mu_j)} \{\beta \in \Delta(B) | \, \beta \cdot s \leq g_1(s, n)\},$$

where $g_1(s', n) = \max_{\beta' \in y} \beta' \cdot s' - \epsilon$, $g_1(s^*, n) = \max_{\beta' \in y} \beta' \cdot s^* - \xi^n$, and $g_1(s, n) = \max_{\beta' \in y} \beta' \cdot s$ for $s \neq s', s^*$. We take ξ small enough so that $\max_{\beta \in x_n} \beta \cdot s = g_1(s, n)$ for every s and n. Again, this is possible since S is finite. By construction, $x_n \to x$.

Compare x_n and x'. In the first hierarchy, the difference between the valuations from x_n and x' is

$$\sum_{s \in S} \mu_1(s) \max_{\beta \in x_n} \beta \cdot s - \sum_{s \in S} \mu_1(s) \max_{\beta \in x'} \beta \cdot s = -\mu_1(s^*) \xi^n > 0.$$

Hence, $x_n \succ x'$ for every n.

Next, compare x' and x. Since $s' \notin \text{supp}(\mu_1)$, x' is indifferent to x in the first hierarchy. In the second hierarchy, the difference between the valuations from x and x' is

$$\sum_{s \in S} \mu_2(s) \max_{\beta \in x'} \beta \cdot s - \sum_{s \in S} \mu_2(s) \max_{\beta \in x} \beta \cdot s = \mu_2(s') \epsilon > 0.$$

Therefore, $x' \succ x$.

For $\mu_1(s^*) > 0$, we modify $\{x_n\}$. Define

$$x_n = \cap_{s \in \cup_{j=1}^2 \text{supp}(\mu_j)} \{\beta \in \Delta(B) | \, \beta \cdot s \leq g_2(s, n)\},$$

where $g_2(s', n) = \max_{\beta' \in y} \beta' \cdot s' - \epsilon$, $g_2(s^*, n) = \max_{\beta' \in y} \beta' \cdot s^* + \xi^n$, and $g_2(s, n) = \max_{\beta' \in y} \beta' \cdot s$ for $s \neq s', s^*$. Then the same argument holds.

Subcase 1-2: $\mu_2(s') < 0$**.** With the following modification, we can make the same argument as in Subcase 1-1. Let

$$x = \cap_{s\in\cup_{j=1}^2 \text{supp}(\mu_j)}\{\beta \in \Delta(B) | \, \beta \cdot s \leq f_2(s)\},$$

where $f_2(s') = \max_{\beta'\in y} \beta' \cdot s' + \epsilon$, and $f_2(s) = \max_{\beta'\in y} \beta' \cdot s$ for $s \neq s'$. If $\mu_1(s^*) > 0$, define

$$x_n = \cap_{s\in\cup_{j=1}^2 \text{supp}(\mu_j)}\{\beta \in \Delta(B) | \, \beta \cdot s \leq g_3(s,n)\},$$

where $g_3(s',n) = \max_{\beta'\in y} \beta' \cdot s' + \epsilon$, $g_3(s^*,n) = \max_{\beta'\in y} \beta' \cdot s^* + \xi^n$, and $g_3(s,n) = \max_{\beta'\in y} \beta' \cdot s$ for $s \neq s', s^*$. If $\mu_1(s^*) < 0$, define

$$x_n = \cap_{s\in\cup_{j=1}^2 \text{supp}(\mu_j)}\{\beta \in \Delta(B) | \, \beta \cdot s \leq g_4(s,n)\},$$

where $g_4(s',n) = \max_{\beta'\in y} \beta' \cdot s' + \epsilon$, $g_4(s^*,n) = \max_{\beta'\in y} \beta' \cdot s^* - \xi^n$, and $g_4(s,n) = \max_{\beta'\in y} \beta' \cdot s$ for $s \neq s', s^*$.

Case 2: $\text{supp}(\mu_1) = \cup_{j=1}^2 \text{supp}(\mu_j)$**.** First, note that $\text{supp}(\mu_1)$ has to contain more than two states since we consider a lexicographic representation of minimal length. Moreover, there exit two states s' and s'' such that $\frac{\mu_2(s')}{\mu_1(s')} \neq \frac{\mu_2(s'')}{\mu_1(s'')}$.

Subcase 2-1: $\mu_1(s) < 0$ **for** $s = s', s''$**.** Label s', s'' as $\frac{\mu_2(s')}{\mu_1(s')} < \frac{\mu_2(s'')}{\mu_1(s'')}$. Let $\epsilon > 0$. Define

$$x = \cap_{s\in \text{supp}(\mu_1)}\{\beta \in \Delta(B) | \, \beta \cdot s \leq f_3(s)\},$$

where $f_3(s') = \max_{\beta'\in y} \beta' \cdot s' - \epsilon$, $f_3(s'') = \max_{\beta'\in y} \beta' \cdot s'' + \frac{\mu_1(s')}{\mu_1(s'')}\epsilon$, and $f_3(s) = \max_{\beta'\in y} \beta' \cdot s$ for $s \neq s', s''$. We take ϵ small enough so that x is a menu with $\max_{\beta\in x} \beta \cdot s = f_3(s)$ for every $s \in S$. This is possible since S is finite.

Let $\xi \in (0,1)$. Define

$$x_n = \cap_{s\in \text{supp}(\mu_1)}\{\beta \in \Delta(B) | \, \beta \cdot s \leq g_5(s,n)\},$$

where $g_5(s',n) = \max_{\beta'\in y} \beta' \cdot s' - \epsilon - \xi^n$, $g_5(s'',n) = \max_{\beta'\in y} \beta' \cdot s'' + \frac{\mu_1(s')}{\mu_1(s'')}\epsilon$, and $g_5(s,n) = \max_{\beta'\in y} \beta' \cdot s$ for $s \neq s', s''$. We take ξ small enough so that $\max_{\beta\in x_n} \beta \cdot s = g_5(s,n)$ for every s and n. Again, this is possible since S is finite. By construction, $x_n \to x$.

Compare x_n and x'. In the first hierarchy, the difference between the valuations from x_n and x' is

$$\sum_{s\in S} \mu_1(s) \max_{\beta\in x_n} \beta \cdot s - \sum_{s\in S} \mu_1(s) \max_{\beta\in x'} \beta \cdot s = -\mu_1(s')\xi^n > 0.$$

Hence, $x_n \succ x'$ for all n.

Next, compare x' and x. In the first hierarchy, x is indifferent to x' because

$$\sum_{s\in S}\mu_1(s)\max_{\beta\in x'}\beta\cdot s-\sum_{s\in S}\mu_1(s)\max_{\beta\in x}\beta\cdot s=\mu_1(s')\epsilon-\mu_1(s'')\frac{\mu_1(s')}{\mu_1(s'')}\epsilon=0.$$

The difference between the valuations from x' and x in the second hierarchy is

$$\begin{aligned}&\sum_{s\in S}\mu_2(s)\max_{\beta\in x'}\beta\cdot s-\sum_{s\in S}\mu_2(s)\max_{\beta\in x}\beta\cdot s\\=&\mu_2(s')\max_{\beta\in y}\beta\cdot s'+\mu_2(s'')\max_{\beta\in y}\beta\cdot s''\\&\quad-\mu_2(s')\left(\max_{\beta\in y}\beta\cdot s'-\epsilon\right)-\mu_2(s'')\left(\max_{\beta\in y}\beta\cdot s''+\frac{\mu_1(s')}{\mu_1(s'')}\epsilon\right)\\=&\frac{\mu_2(s')\mu_1(s'')-\mu_2(s'')\mu_1(s')}{\mu_1(s'')}\epsilon>0.\end{aligned}$$

Therefore, $x'\succ x$.

With the following constructions of x and $\{x_n\}$, the same argument holds for the other cases.

Subcase 2-2: $\mu_1(s)>0$ for $s=s',s''$. Label s',s'' as $\frac{\mu_2(s')}{\mu_1(s')}<\frac{\mu_2(s'')}{\mu_1(s'')}$. Define

$$x=\cap_{s\in\text{supp}(\mu_1)}\{\beta\in\Delta(B)|\,\beta\cdot s\le f_4(s)\},$$

where $f_4(s')=\max_{\beta'\in y}\beta'\cdot s+\epsilon$, $f_4(s'')=\max_{\beta'\in y}\beta'\cdot s''-\frac{\mu_1(s')}{\mu_1(s'')}\epsilon$, and $f_4(s)=\max_{\beta'\in y}\beta'\cdot s$ for $s\ne s',s''$. Define

$$x_n=\cap_{s\in\text{supp}(\mu_1)}\{\beta\in\Delta(B)|\,\beta\cdot s\le g_6(s,n)\},$$

where $g_6(s',n)=\max_{\beta'\in y}\beta'\cdot s'+\epsilon+\xi^n$, $g_6(s'',n)=\max_{\beta'\in y}\beta'\cdot s''-\frac{\mu_1(s')}{\mu_1(s'')}\epsilon$, and $g_6(s,n)=\max_{\beta'\in y}\beta'\cdot s$ for $s\ne s',s''$.

Subcase 2-3: $\mu_1(s')<0$ and $\mu_1(s'')>0$. If $\frac{\mu_2(s')}{\mu_1(s')}<\frac{\mu_2(s'')}{\mu_1(s'')}$, the constructions of x and $\{x_n\}$ are the same as Subcase 2-1.

If $\frac{\mu_2(s')}{\mu_1(s')}>\frac{\mu_2(s'')}{\mu_1(s'')}$, we define

$$x=\cap_{s\in\text{supp}(\mu_1)}\{\beta\in\Delta(B)|\,\beta\cdot s\le f_5(s)\},$$

where $f_5(s')=\max_{\beta'\in y}\beta'\cdot s'+\epsilon$, $f_5(s'')=\max_{\beta'\in y}\beta'\cdot s''-\frac{\mu_1(s')}{\mu_1(s'')}\epsilon$, and $f_5(s)=\max_{\beta'\in y}\beta'\cdot s$ for $s\ne s',s''$. Define

$$x_n=\cap_{s\in\text{supp}(\mu_1)}\{\beta\in\Delta(B)|\,\beta\cdot s\le g_7(s,n)\},$$

where $g_7(s',n)=\max_{\beta'\in y}\beta'\cdot s'+\epsilon-\xi^n$, $g_7(s'',n)=\max_{\beta'\in y}\beta'\cdot s''-\frac{\mu_1(s')}{\mu_1(s'')}\epsilon$, and $g_7(s,n)=\max_{\beta'\in y}\beta'\cdot s$ for $s\ne s',s''$.

D Proof of Corollary 4.4

The necessity is straightforward. We show only the sufficiency.

To begin with, note that Monotonicity implies that μ_1 is a positive measure.

Let $(S, U, \{\mu_j\}_{j=1}^K)$ be a canonical lexicographic representation of $\succsim$ such that μ_j is a positive measure for $j = 1, \cdots, k$ $(\leqq K-1)$. It is enough to show that there is a canonical lexicographic representation $(S, U, \{\mu'_j\}_{j=1}^K)$ of $\succsim$ such that $\mu'_j = \mu_j$ for $j = 1, \cdots, k$ and μ'_{k+1} is a positive measure.

First, we see that if $s \in \text{supp}(\mu_{k+1}) \setminus \cup_{j=1}^k \text{supp}(\mu_j)$, then $\mu_{k+1}(s) > 0$. Suppose to the contrary that $\mu_{k+1}(s) < 0$. Fix a sphere $y \subset \text{int}\Delta(B)$. Define

$$x = \cap_{s' \in \cup_{j=1}^{k+1} \text{supp}(\mu_j) \setminus \{s\}} \{\beta \in \Delta(B) | \ \beta \cdot s' \leqq \max_{\beta' \in y} \beta' \cdot s'\}.$$

Then the representation implies that $y \succ x$ while $y \subset x$, contradicting Monotonicity.

Next, we construct the desired $\{\mu'_j\}_{j=1}^K$. Let $\epsilon > 0$ be such that

$$\min\{\mu_j(s) | j = 1, \cdots, k, \text{ and } s \in \cup_{j=1}^k \text{supp}(\mu_j)\} > \epsilon \max_s |\mu_{k+1}(s)| \, .$$

Define $\{\mu'_j\}_{j=1}^K$ by

$$\mu'_j = \mu_j \text{ for } j \neq k+1,$$
$$\mu'_{k+1} = \sum_{j=1}^k \mu_j + \epsilon \mu_{k+1}.$$

Then it follows from Part (ii) of Lemma 2 that $(S, U, \{\mu'_j\})$ represents the same order as $(S, U, \{\mu_j\})$. By construction, μ'_j is a positive measure for $j = 1, \cdots, k+1$. This completes the proof.

References

[1] Arrow, K, J. Alternative Approaches to the Theory of Choice in Risk-Taking Situations, *Econometrica*, Vol. 19, 404–437, 1951.

[2] Blume, L., A. Brandenburger, E. Dekel, Lexicographic probability and choice under uncertainty, *Econometrica*, Vol. 59, 61–79, 1991.

[3] Brandenberger, A., A. Friedenberg, and H. J. Keisler, Admissibility in Games, *Econometrica*, Vol. 76, 307–352, 2008.

[4] Dekel, E., B. Lipman, and A. Rustichini, Representing preferences with a unique subjective state space, *Econometrica*, Vol. 69, 891–934, 2001.

[5] Dekel, E., B. Lipman, and A. Rustichini, Temptation-Driven Preferences, forthcoming in *Review of Economic Studies.*

[6] Gul, F., and W. Pesendorfer, Temptation and Self-Control, *Econometrica*, Vol. 69, 1403–1435, 2001

[7] Hausner, M., Multidimensional Utilities, In Thrall, R.M., C. H. Cooms, and R. L. Davis, editors, *Decision Process*, Wiley, NY, 1954.

[8] Kohlberg, E., and J. F. Mertens, On the Stratigic Stability of Equilibria, *Econometrica*, Vol. 54, 1003–1037, 1986.

[9] Kreps, D. M., A representation theorem for preference for flexibility, *Econometrica*, Vol. 47, 565–578, 1979.

[10] Kreps, D. M., Static choice and unforeseen contingencies, In Dasgupta, P. and D. Gale, and O. Hart, and E. Maskin, editors, *Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*, 259–281. MIT Press, Cambridge, MA, 1992.

[11] Luce, R., and H. Raiffa, *Games and Decisions*, Wiley, New York, 1957.