Representation with a Unique Subjective Decision Tree *

Norio TAKEOKA

Faculty of Economics, Ritsumeikan University,
1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan

June 16, 2006

Abstract

In a dynamic setting, a decision tree, that is, a pair consisting of a state space and a filtration, has been taken as a primitive for modeling uncertainty. This assumption implicitly requires the analyst to know not only the uncertainties a decision maker perceives, but also how she anticipates those uncertainties to be resolved over time. This is problematic because a decision tree is in the mind of the DM and hence is not directly observable to the analyst. Without assuming any objective states, we derive a unique subjective decision tree from preference over suitable choice objects. This result is a three-stage extension of Dekel, Lipman and Rustichini (2001).

Keywords: preference for flexibility, subjective state space, subjective decision tree.
JEL classification: D81

1 Introduction

1.1 Motivation and Results

Since Savage [10], a state space has been used as the standard tool for modeling uncertainty, and has been taken as a primitive. This assumption implicitly requires the analyst (or outside observer) to know all the uncertainties a decision maker (DM) perceives. This is problematic because states express the DM's perception of the world and hence are not directly observable to the analyst. Kreps [5, 6] address the question of whether we can derive *subjective state spaces*. Dekel, Lipman and Rustichini [1] (hereafter DLR) refine Kreps's idea and derive a unique subjective state space.

*E-mail: takeoka@ec.ritsumei.ac.jp. I would like to thank Larry Epstein for his illuminating guidance and invaluable suggestions. I am thankful also to Alfredo Di Tillio, Takashi Hayashi, Kazuya Hyogo, and Massimo Marinacci for helpful comments. All remaining errors are mine.

DLR consider preference over menus of lotteries over alternatives. If a DM is uncertain about her future preference over lotteries, then her ranking of menus should reflect this uncertainty. Under the hypothesis that the DM's future preference satisfies the expected utility axioms, DLR identify a unique set of future preferences from the ranking of menus, and call it the subjective state space.

In a dynamic setting, the standard tool for modeling uncertainty is a decision tree, that is a pair consisting of a state space and a filtration. It has been taken as a primitive. This assumption is more problematic than that in the static setting because the analyst must know not only all the uncertainties the DM perceives, but also how she perceives those uncertainties to be resolved over time. We are led to ask whether a decision tree can be subjective. The derivation of a *subjective decision tree* is the focus of the paper.

We provide an extension of DLR to a three-stage setting to model a DM who anticipates subjective uncertainty to be resolved gradually over time. Let Z be a finite set of alternatives and $\Delta(Z)$ be the set of all lotteries over Z. Let $\mathcal{K}(\cdot)$ denote the set of non-empty compact subsets of a metric space '$\cdot$'. We consider preference $\succeq$ over $\mathcal{D} \equiv \mathcal{K}(\mathcal{K}(\Delta(Z)))$, that is, the set of menus of menus of lotteries.

We axiomatize preference having the following representation: there exist a product state space $S_1 \times S_2$, a countably additive non-negative measure μ_0 over S_1, a conditional probability system $\mu_1 : S_1 \to \Delta(S_2)$, and a state-dependent mixture linear function $u : S_2 \times \Delta(Z) \to \mathbb{R}$ such that the functional form $U_0 : \mathcal{D} \to \mathbb{R}$,

$$\begin{aligned} U_0(x_0) &= \int_{S_1} \max_{x_1 \in x_0} U_1(x_1, s_1)\, \mathrm{d}\mu_0(s_1), \text{ where} \qquad (1) \\ U_1(x_1, s_1) &= \int_{S_2} \max_{l \in x_1} u(l, s_2)\, \mathrm{d}\mu_1(s_2|s_1), \end{aligned}$$

represents preference.

An interpretation of the representation is as follows: the DM is not sure of her own future preference over lotteries except that it conforms to the expected utility axioms. This subjective uncertainty is captured by μ_0 and μ_1 on $S_1 \times S_2$. The DM behaves as if she anticipates subjective signals to arrive gradually over time, and chooses menus so as to maximize the additive utility across states.

Note that the only relevant part of a signal is the conditional preference that it generates. The representation U_0 with components $(S_1 \times S_2, \mu_0, \mu_1, u)$ determines a set of sequences of "ex-post" preferences. Each $s_1 \in S_1$ determines the conditional preference $\succeq_{s_1}$ on $\mathcal{K}(\Delta(Z))$ induced by $U_1(\cdot, s_1)$. Similarly, each $s_2 \in S_2$ determines the preference $\succeq_{s_2}$ on $\Delta(Z)$ induced by $u(\cdot, s_2)$. From the conditional probability system $\mu_1 : S_1 \to \Delta(S_2)$, we can define the set of all admissible sequences of ex-post preferences $(\succeq_{s_1}, \succeq_{s_2})$. This set, denoted by $\mathbb{S}$, is called the *subjective state space*.

The *subjective filtration* $\{\mathbb{F}_t\}_{t=0}^2$ over $\mathbb{S}$ is naturally determined from the time line. Under a suitable condition, the *subjective decision tree* $(\mathbb{S}, \{\mathbb{F}_t\}_{t=0}^2)$ is uniquely derived from preference.

1.2 An Example: Dinner Choice in the Morning

Now we illustrate how the subjective decision tree can be deduced from preference over menus of menus. The following is a dynamic version of the example in Kreps [5]. Imagine the situation where a DM chooses either chicken or fish for dinner this evening. There are three periods, morning, noon, and evening.

Consider the following four menus of menus:

$$\{\{\text{chicken}\}\},\ \{\{\text{fish}\}\},\ \{\{\text{chicken}\},\{\text{fish}\}\},\ \{\{\text{chicken, fish}\}\}.$$

If the DM chooses $\{\{\text{chicken}\}\}$, there is no room for subsequent choices. In other words, she has to commit herself to chicken in the morning. For example, this option is interpreted as reserving a chicken dinner right now. Similarly, $\{\{\text{fish}\}\}$ can be interpreted as reserving a fish dinner. If the DM chooses $\{\{\text{chicken}\},\{\text{fish}\}\}$, she does not have to commit herself right now to either chicken or fish. She can rather postpone her decision until noon. In terms of flexibility, she may prefer this option to both $\{\{\text{chicken}\}\}$ and $\{\{\text{fish}\}\}$. In terms of flexibility, the option $\{\{\text{chicken, fish}\}\}$ may be more preferable to $\{\{\text{chicken}\},\{\text{fish}\}\}$ because the DM can delay a decision until the evening rather than until noon.

Consider two possible rankings:

$$\{\{\text{chicken, fish}\}\} \sim \{\{\text{chicken}\},\{\text{fish}\}\} \succ \{\{\text{chicken}\}\} \sim \{\{\text{fish}\}\}, \tag{2}$$

$$\{\{\text{chicken, fish}\}\} \succ \{\{\text{chicken}\},\{\text{fish}\}\} \sim \{\{\text{chicken}\}\} \sim \{\{\text{fish}\}\}. \tag{3}$$

Ranking (2) says that the DM desires flexibility at noon, but she desires no more flexibility afterwards. The strict ranking, $\{\{\text{chicken}\},\{\text{fish}\}\} \succ \{\{\text{chicken}\}\}$, can be justified by the following story: the DM anticipates that one of two subjective signals arrives at noon. She is aware that she may change her mind according to those signals. After receiving one signal, she feels chicken is more preferable to fish, while this ranking is reversed after receiving the other signal. That is, we hypothesize that preference for flexibility at noon comes from her awareness of this uncertainty. Presumably, ranking (2) reveals a subjective decision tree as Figure 1.

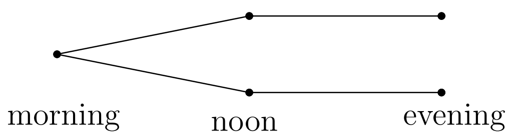

Figure 1: Subjective decision tree deduced from ranking (2)

Ranking (3) says that the DM has no preference for flexibility at noon, but she desires flexibility in the evening. The strict preference, $\{\{\text{chicken, fish}\}\} \succ \{\{\text{chicken}\},\{\text{fish}\}\}$, can be interpreted as above. That is, the DM anticipates at least two subjective signals in the evening and is aware of preference change according to those signals. Ranking (3) will imply a subjective decision tree as Figure 2.

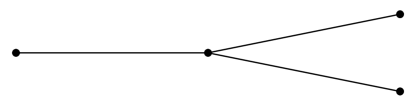

Figure 2: Subjective decision tree deduced from ranking (3)

1.3 Related Literature

Kreps [5, 6] provide an axiomatic foundation for subjective state spaces. By enriching choice objects to lotteries over alternatives, DLR show uniqueness of the subjective state space. More precisely, DLR consider preference over $\mathcal{P}(\Delta(Z))$, where $\mathcal{P}(\cdot)$ denotes the set of all non-empty subsets of a set "$\cdot$". They provide a set of axioms that guarantees the following additive representation:

$$U(x_1) = \int_S \sup_{l \in x_1} u(l, s)\, \mathrm{d}\mu(s), \tag{4}$$

where S is a state space, μ is a countably additive non-negative measure on S, and $u(\cdot, s) : \Delta(Z) \to \mathbb{R}$ is a state-dependent mixture linear function. DLR show that the set of ex post preferences induced from $\{u(\cdot, s)\}_{s \in S}$ is uniquely determined from preference, and call it the *subjective state space*.

Takeoka [12] also provides a three-stage extension of DLR and derives subjective decision trees. To identify subjective beliefs as well, he assumes some objective states and considers preference over menus of menus of Anscombe-Aumann acts. Under the hypothesis that ex post preferences satisfy the subjective expected utility axioms, he models a DM who is certain about future risk preference, but not sure of future beliefs about objective states.

The difference from our paper is that Takeoka [12] is not a generalization of DLR to a three-stage setting. Notice that $\mathcal{K}(\mathcal{K}(\Delta(Z)))$ can be regarded as a subdomain of the set of menus of menus of Anscombe-Aumann acts. On this subdomain, his model collapses to the functional form

$$U_0(x_0) = \max_{x_1 \in x_0} \max_{l \in x_1} u(l),$$

where subjective uncertainty does not play any role. This is because subjective uncertainty concerns future beliefs about objective states in his model. A multi-stage generalization of DLR, provided in this paper, is of independent interest as a foundation for subjective decision trees.

Rustichini [9] addresses a multi-period extension of DLR. Let C be the set of consumptions and C^∞ be the set of infinite consumption streams. He considers $\mathcal{P}(C^\infty)$ as the set of choice objects. His model does not deliver subjective decision trees because it is essentially static in the sense that all subjective uncertainties are resolved in the next period. Modica [7] considers preference over $\mathcal{P}(\mathcal{P}(Z))$. This model, however, cannot pin down representations as in Kreps [5]. Kraus and Sagi [4] consider preferences without the completeness axiom, and accommodate preference for flexibility in a dynamic setting.

2 Model

2.1 Domain

Let Z be a finite set of alternatives with $\#Z = m$ and $\Delta(Z)$ be the set of all Borel probability measures over Z.[1] The set $\Delta(Z)$ is regarded as the $(m-1)$-dimensional unit simplex.

Let $\mathcal{K}(\cdot)$ denote the set of all non-empty compact subsets of a metric space '$\cdot$'. We consider preference over $\mathcal{D} \equiv \mathcal{K}(\mathcal{K}(\Delta(Z)))$. Elements in $\mathcal{D}$, denoted by $x_0, y_0, \cdots$, are interpreted as menus of menus of lotteries. Endow $\mathcal{D}$ with the Hausdorff metric. Details are found in Section A.

We hypothesize that the DM behaves as if she has in mind the following timing of decisions:

Period 0**:** choose a menu of menus x_0

Period 1^-**:** receive a subjective signal s_1

Period 1**:** choose a menu $x_1 \in x_0$

Period 2^-**:** receive another subjective signal s_2

Period 2**:** choose a lottery $l \in x_1$

Notice that this time line, beyond period 0, is not part of the formal model. Rather, as shown below, the time line including subjective signals is derived as a representation theorem.

Since DLR use $\mathcal{K}(\Delta(Z))$ as the domain, one might wonder whether $\mathcal{K}(\Delta(\mathcal{K}(\Delta(Z))))$ is more natural as the dynamic counterpart of DLR. There are two reasons why we adopt $\mathcal{K}(\mathcal{K}(\Delta(Z)))$. First of all, DLR consider lotteries because richness of $\Delta(Z)$ makes possible to show uniqueness of the representation. In our case, $\mathcal{K}(\Delta(Z))$ already has a rich structure, and hence we can show uniqueness without additional lotteries. Second, DLR adopt $\mathcal{K}(\Delta(Z))$ so as to justify one of their axioms, called Independence. The counterpart of the axiom can be justified also on $\mathcal{K}(\mathcal{K}(\Delta(Z)))$.

2.2 Axioms

Preference in period 0, that is, $\succeq$ on $\mathcal{D}$, should reflect how the DM anticipates subjective signals to arrive over time. To capture subjective decision trees, we impose the following six axioms on $\succeq$. The first five axioms are formally identical to those of DLR, but are imposed here on $\mathcal{K}(\mathcal{K}(\Delta(Z)))$ rather than on $\mathcal{K}(\Delta(Z))$.

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

[1] For any given metric space X, a topology on the set of all Borel probability measures on X, denoted by $\Delta(X)$, is always understood to be the weak convergence topology.

Axiom 2 (Continuity). For all $x_0 \in \mathcal{D}$, $\{z_0 \in \mathcal{D}|x_0 \succeq z_0\}$ and $\{z_0 \in \mathcal{D}|z_0 \succeq x_0\}$ are closed.

Axiom 3 (Nondegeneracy). There exist $x_0, x_0' \in \mathcal{D}$ such that $x_0 \succ x_0'$.

Define the mixture

$$\lambda x_1 + (1-\lambda)x_1' \equiv \{\lambda l + (1-\lambda)l' | l \in x_1, l' \in x_1'\},$$

for any $x_1, x_1' \in \mathcal{K}(\Delta(Z))$ and $\lambda \in [0,1]$, and

$$\lambda x_0 + (1-\lambda)x_0' \equiv \{\lambda x_1 + (1-\lambda)x_1' | x_1 \in x_0, x_1' \in x_0'\},$$

for any $x_0, x_0' \in \mathcal{D}$ and $\lambda \in [0,1]$.

Axiom 4 (Independence). For all $x_0, y_0, z_0 \in \mathcal{D}$ and for all $\lambda \in (0,1]$,

$$x_0 \succ y_0 \;\Rightarrow\; \lambda x_0 + (1-\lambda)z_0 \succ \lambda y_0 + (1-\lambda)z_0.$$

Independence can be justified by adapting DLR's argument twice. First, for any $x_0, z_0 \in \mathcal{D}$ and $\lambda \in [0,1]$, consider the lottery $\lambda \circ x_0 + (1-\lambda) \circ z_0$, which assigns x_0 with probability λ and z_0 with probability $(1-\lambda)$. vNM independence axiom implies that, for any $\lambda \in (0,1]$, if x_0 is strictly preferred to y_0, then $\lambda \circ x_0 + (1-\lambda) \circ z_0$ is strictly preferred to $\lambda \circ y_0 + (1-\lambda) \circ z_0$.

Second, we argue that the DM is indifferent between $\lambda \circ x_0 + (1-\lambda) \circ z_0$ and $\lambda x_0 + (1-\lambda) z_0$. This indifference says that the DM does not care when the randomization $(\lambda, 1-\lambda)$ is realized. This is appealing if the DM surely believes that her future preference in period 1, that is, preference over $\mathcal{K}(\Delta(Z))$, satisfies Independence in the sense of DLR. As DLR argue, this assumption is in turn appealing whenever the DM is sure that her future preference in period 2, that is, preference over $\Delta(Z)$, satisfies the expected utility axioms. Thus Axiom 4 follows from the above two steps together with the hypothesis that the DM's future preferences in period 2 satisfy the expected utility axioms.

The next axiom says that a bigger menu of menus is always weakly preferred.

Axiom 5 (Monotonicity). For all $x_0, y_0 \in \mathcal{D}$, $x_0 \subset y_0 \Rightarrow y_0 \succeq x_0$.

A bigger menu of menus allows the DM to leave more options open until period 1. Hence Monotonicity is consistent with preference for flexibility.

The last axiom has no counterpart in DLR. It is relevant only in a dynamic setting and states that the DM always prefers to delay a decision.

Axiom 6 (Aversion to Commitment). For all $y_0 \in \mathcal{D}$ and all finite $x_0 \in \mathcal{D}$, $y_0 \cup \{\cup_{x_1 \in x_0} x_1\} \succeq y_0 \cup x_0$.

If $y_0 \cup \{\cup_{x_1 \in x_0} x_1\}$ is chosen over $y_0 \cup x_0$, the DM can choose a weakly bigger menu in period 1. In other words, $y_0 \cup \{\cup_{x_1 \in x_0} x_1\}$ allows her to leave more options until period 2. Hence Aversion to Commitment is intuitive in terms of preference for flexibility.

3 Representations

3.1 Additive EU Representation

We briefly describe the additive representation on $\mathcal{K}(\Delta(Z))$, introduced by DLR. Consider the functional form $U_1 : \mathcal{K}(\Delta(Z)) \to \mathbb{R}$ defined by

$$U_1(x_1) \equiv \int_{S_2} \max_{l \in x_1} u(l, s_2)\, \mathrm{d}\mu_1(s_2), \tag{5}$$

where S_2 is a state space, μ_1 is a countably additive non-negative measure over S_2, and $u : \Delta(Z) \times S_2 \to \mathbb{R}$ is a state-dependent mixture linear function.

Notice that the payoff-relevant information conveyed by a signal is the ex-post risk preference on $\Delta(Z)$ – a signal itself does not matter. Thus we can effectively identify the subjective uncertainties with the set of "ex-post" preferences as we now describe. Take any functional form U_1 with components (S_2, μ_1, u) as above. The preference $\succeq_{s_2}$ induced on $\Delta(Z)$ conditional on $s_2 \in S_2$ is

$$l \succeq_{s_2} l' \;\Leftrightarrow\; u(l, s_2) \geq u(l', s_2).$$

Let $\mathbb{S}_2$ be the set of all conditional preferences, that is,

$$\mathbb{S}_2 \equiv \{\succeq_{s_2} | s_2 \in S_2\},$$

which is called the *subjective state space*.

Potentially, there are many functional forms (5) representing the same preference on $\mathcal{K}(\Delta(Z))$. For example, we can consider a copy S_1' of state space S_1 and split the weight μ_1 between the two state spaces. Then the functional form with $S_1 \cup S_1'$ represents the same preference. To obtain uniqueness, we pay attention to a functional form such that $\succsim_{s_2} \neq \succsim_{s_2'}$ if $s_2 \neq s_2'$.

There is another source for non-uniqueness of the representation. Take a functional form U_1 with components (S_2, μ_1, u). We can add some irrelevant states to S_2 with assuming that μ_1 assigns probability zero to those states. Then the functional form with the new state space also represents the same preference.

To exclude this trivial non-uniqueness, DLR pay attention to "relevant" states. Given (S_2, μ_1, u) with finite S_2, say that $\succeq_{s_2} \in \mathbb{S}_2$ is *relevant* if there exist $x_1, y_1 \in \mathcal{K}(\Delta(Z))$ with $U_1(x_1) \neq U_1(y_1)$ such that

$$\max_{l \in x_1} u(l, s_2') = \max_{l \in y_1} u(l, s_2')$$

for all $s_2' \in S_2$ with $\succeq_{s_2'} \neq \succeq_{s_2}$.[2]

[2]When S_2 is infinite, $\succeq_{s_2} \in \mathbb{S}_2$ is said to be *relevant* if, for all neighborhood N of s_2, there exist $x_1, y_1 \in \mathcal{K}(\Delta(Z))$ with $U_1(x_1) \neq U_1(y_1)$ such that $\max_{l \in x_1} u(l, s_2') = \max_{l \in y_1} u(l, s_2')$ for all $s_2' \in S_2 \setminus N$.

DLR use this notion to show uniqueness of the subjective state space without additivity across states.[3] In case of additive representations with finitely supported non-negative measures, relevance is equivalent to say that s_2 belongs to the support of μ_1. See Appendix B for details.

The above argument leads to the next definition.

Definition 3.1. Preference on $\mathcal{K}(\Delta(Z))$ admits an additive EU representation with a non-negative measure if (i) the functional form $U_1 : \mathcal{K}(\Delta(Z)) \to \mathbb{R}$ with components (S_2, μ_1, u) represents preference, (ii) every state $s_2 \in S_2$ is relevant, and (iii) $\succeq_{s_2} \neq \succeq_{s_2'}$ if $s_2 \neq s_2'$.

DLR show that Order, Continuity, Nondegeneracy, Independence and Monotonicity on preference over $\mathcal{K}(\Delta(Z))$ are necessary and sufficient for an additive EU representation with a non-negative measure. Furthermore, they show that the subjective state space $\mathbb{S}_2$ is uniquely derived from preference if S_2 is finite. When S_2 is infinite, uniqueness is up to the closure of $\mathbb{S}_2$.

3.2 Recursive Additive EU Representation

We describe a representation of preference on the domain $\mathcal{D} \equiv \mathcal{K}(\mathcal{K}(\Delta(Z)))$. Consider the functional form $U_0 : \mathcal{D} \to \mathbb{R}$ defined by

$$U_0(x_0) \equiv \int_{S_1} \max_{x_1 \in x_0} U_1(x_1, s_1) \, \mathrm{d}\mu_0(s_1), \tag{6}$$

where

$$U_1(x_1, s_1) \equiv \int_{S_2} \max_{l \in x_1} u(l, s_2) \, \mathrm{d}\mu_1(s_2|s_1),$$

$S_1 \times S_2$ is a state space, μ_0 is a countably additive non-negative measure on S_1, $\mu_1 : S_1 \to \Delta(S_2)$ is a conditional probability system, and $u : S_2 \times \Delta(Z) \to \mathbb{R}$ is a state-dependent mixture linear function.

An interpretation of the functional form is as follows: the DM behaves as if she has in mind the time line described in Section 2.1, and anticipates to receive subjective signals gradually over time. Subjective uncertainty concerns future risk preferences, that is, she is not sure of future preference over lotteries. The DM expects that, at decision nodes (period 1 and period 2), she makes decisions so as to maximize the signal-dependent additive utility function, $U_1(\cdot, s_1)$ or $u(\cdot, s_2)$.

One might wonder why u is independent of S_1. Alternatively, $u : S_1 \times S_2 \times \Delta(Z) \to \mathbb{R}$ seems more general. The above formulation is without loss of generality. Indeed, if u depends also on s_1, we can always redefine S_2 by the new state space $S_2^* \equiv S_1 \times S_2$. Then, $u : S_2^* \times \Delta(Z) \to \mathbb{R}$ is independent of S_1, and μ_1 is naturally identified with a function from S_1 into $\Delta(S_2^*)$.

[3]See the definition of weak EU representations in DLR (p. 903).

For each $s_1 \in S_1$, the conditional preference $\succeq_{s_1}$ induced on $\mathcal{K}(\Delta(Z))$ is defined by $U_1(\cdot, s_1)$. Let $\mathbb{S}_1$ be the set of all conditional preferences, that is,

$$\mathbb{S}_1 \equiv \{\succeq_{s_1} | s_1 \in S_1\}.$$

Though relevance of $\succeq_{s_1}$ can be defined as in the previous section, we use an alternative notion. Say that μ_0 on S_1 *has full support* if $\mu_0(G) > 0$ for all open subsets G of S_1. If $s_1 \in S_1$ does not belong to the support of μ_0, s_1 is not relevant. Thus the full support condition is weaker than the condition that every $s_1 \in S_1$ is relevant. As shown in the next section, the full support condition is enough to show uniqueness if supports of μ_0 and of all $\mu_1(s_1)$ are finite.

The following is a counterpart of Definition 3.1 in a three-period setting:

Definition 3.2. Preference $\succeq$ on $\mathcal{D}$ admits a recursive additive EU representation if (i) functional form (6) with components $(S_1 \times S_2, \mu_0, \mu_1, u)$ represents preference, (ii) μ_0 has full support, (iii) $\succeq_{s_1} \neq \succeq_{s_1'}$ if $s_1 \neq s_1'$, and (iv) for all $s_1 \in S_1$, $\succeq_{s_1}$ has an additive EU representation $(S_2(s_1), \mu_1(s_1), u(s_1))$.

Now we are ready to state the main theorem.

Theorem 3.1. *The following statements are equivalent:*

(a) *Preference $\succeq$ on $\mathcal{D}$ satisfies Order, Continuity, Nondegeneracy, Independence, Monotonicity, and Aversion to Commitment.*

(b) *Preference $\succeq$ on $\mathcal{D}$ admits a recursive additive EU representation.*

A proof can be found in Section C.1.

4 Uniqueness

The significance of a signal is the conditional preference generated by the signal. The DM should care about conditional preferences rather than signals themselves. Thus the subjective decision tree is effectively identified with the set of sequences of ex post preferences. We show uniqueness of subjective decision trees under a suitable condition.

For any recursive additive EU representation $(S_1 \times S_2, \mu_0, \mu_1, u)$, recall that $\succeq_{s_1}$ and $\succeq_{s_2}$ denote conditional preferences induced by $s_1 \in S_1$ and by $s_2 \in S_2$, respectively. For any probability measure ν, let $\text{supp}(\nu)$ denote the support of ν. Let

$$\begin{aligned}
\mathbb{S}_1 &\equiv \{\succeq_{s_1} | s_1 \in S_1\}, \\
\mathbb{S}_2(s_1) &\equiv \{\succeq_{s_2} | s_2 \in \text{supp}(\mu_1(s_1))\} \text{ for each } s_1 \in S_1, \text{ and} \\
\mathbb{S}_2 &\equiv \bigcup_{s_1 \in S_1} \mathbb{S}_2(s_1).
\end{aligned}$$

The set $\mathbb{S}_1$ consists of all ex post preferences over menus, induced by the first signals. In other words, $\mathbb{S}_1$ captures the subjective contingencies the DM expects to be resolved by period 1. Next, $\mathbb{S}_2(s_1)$ is the set of possible preferences over lotteries when the DM observes s_1. This set captures the remaining subjective contingencies the DM expects to face after seeing the signal s_1. Finally, $\mathbb{S}_2$ is the set of all subjective contingencies the DM anticipates to face at the beginning of period 2.

Now define the *subjective state space*, denoted by $\mathbb{S}$, as the set of all admissible sequences of ex-post preferences, that is,

$$\mathbb{S} \equiv \{(\succeq_{s_1}, \succeq_{s_2}) \in \mathbb{S}_1 \times \mathbb{S}_2 | \succeq_{s_2} \in \mathbb{S}_2(s_1) \text{ for some } s_1 \in S_1\}. \tag{7}$$

The *subjective filtration* $\{\mathbb{F}_t\}_{t=0}^2$ over $\mathbb{S}$ is determined according to the time line. That is,

$$\begin{aligned} \mathbb{F}_0 &\equiv \{\mathbb{S}\}, \\ \mathbb{F}_1 &\equiv \left\{ \{(\succeq_1, \succeq_2) | (\succeq_1, \succeq_2) \in \mathbb{S}\} \middle| \succeq_1 \in \mathbb{S}_1 \right\}, \text{ and} \\ \mathbb{F}_2 &\equiv \left\{ \{(\succeq_1, \succeq_2)\} \middle| (\succeq_1, \succeq_2) \in \mathbb{S} \right\}. \end{aligned} \tag{8}$$

The pair $(\mathbb{S}, \{\mathbb{F}_t\}_{t=0}^2)$ is called *the subjective decision tree.* The next theorem states that the subjective decision tree is uniquely derived from preference under the finite support condition.

Theorem 4.1. *If two recursive additive EU representations, $(S_1^i \times S_2^i, \mu_0^i, \mu_1^i, u^i)$, $i = 1, 2$, represent the same preference $\succeq$ on $\mathcal{D}$, and if S_1^i and S_2^i are finite for $i = 1, 2$, then $(\mathbb{S}^1, \{\mathbb{F}_t^1\}_{t=0}^2) = (\mathbb{S}^2, \{\mathbb{F}_t^2\}_{t=0}^2)$.*

Theorem 4.1 is a direct consequence of the next proposition. A proof of the proposition can be found in Section C.2.

Proposition 4.1. *If two recursive additive EU representations, $(S_1^i \times S_2^i, \mu_0^i, \mu_1^i, u^i)$, $i = 1, 2$, represent the same preference $\succeq$ on $\mathcal{D}$, and if S_1^i and S_2^i are finite for $i = 1, 2$, then:*

(i) $\mathbb{S}_2^1 = \mathbb{S}_2^2$*; and*

(ii) $\mathbb{S}_1^1 = \mathbb{S}_1^2$.

Part (i) says that the sets of possible future risk preferences the DM anticipates in period 2 are identical between the two representations. Part (ii) says that the sets of possible future preferences over menus coincide between the two representations.

Theorem 4.1 follows from Proposition 4.1 together with the uniqueness result of DLR. Under the condition of the theorem, Proposition 4.1 ensures that $\succeq$ on $\mathcal{D}$ uniquely determines the set of ex post preferences $\succeq_{s_1}$, that is, $\mathbb{S}_1^1 = \mathbb{S}_1^2$. Moreover, DLR show that every $\succeq_{s_1}$ having an additive EU representation determines a unique set of the ex post preferences $\succeq_{s_2}$. Thus, by construction of the subjective decision tree, that is, (7) and (8), the two subjective decision trees coincide as desired.

Uniqueness is shown under the finite support condition. This is because part (ii) of Definition 3.2 is weaker than to say that every state $s_1 \in S_1$ is relevant. Under the finite support condition, those two conditions are equivalent, and we can show uniqueness of the subjective decision tree.

5 Concluding Remarks

In this paper, we have extended DLR to a three-stage setting, and derived the pair $(\mathbb{S}, \{\mathbb{F}_t\}_{t=0}^2)$, that is, the subjective state space and the subjective filtration, from preference over menus of menus of lotteries. This result shows that foundations do exist for subjective decision trees in a three period setting. We have shown uniqueness of subjective decision trees under the finite support condition.

One might wonder if a T-stage model, that is, preference over menus of menus of ... menus of lotteries is considered. The three-stage setting is the minimal extension of DLR that allows one to address foundations for subjective decision trees. Therefore, though T-stage generalization might be axiomatized, we would view it more as a technical extension than as a conceptual one. We do not have a general representation result to offer at this time, and leave it for a future research.

A Hausdorff metric

Let

$$d(l, x_1) \equiv \min_{l' \in x_1} d(l, l'), \text{ and } e(x_1', x_1) \equiv \max_{l' \in x_1'} d(l', x_1),$$

where d is a metric on $\Delta(Z)$. For each $x_1, y_1 \in \mathcal{K}(\Delta(Z))$, define

$$d_h(x_1, y_1) \equiv \max[e(x_1, y_1), e(y_1, x_1)].$$

We call d_h the *Hausdorff metric*. It is known also that $\mathcal{K}(\Delta(Z))$ is a compact metric space under the Hausdorff metric d_h.

Similarly, $\mathcal{K}(\mathcal{K}(\Delta(Z)))$ can be endowed with the Hausdorff metric. Let $D(x_1, x_0) \equiv \min_{x_1' \in x_0} d_h(x_1, x_1')$ and $E(x_0', x_0) \equiv \max_{x_1' \in x_0'} D(x_1', x_0)$. For each $x_0, y_0 \in \mathcal{D}$, let

$$d_H(x_0, y_0) \equiv \max[E(x_0, y_0), E(y_0, x_0)].$$

Since $\mathcal{K}(\Delta(Z))$ is metric and compact under the Hausdorff metric d_h, so is $\mathcal{K}(\mathcal{K}(\Delta(Z)))$ under d_H.

B Full Support and Relevance

Consider components (S_2, μ_1, u) of the functional form (5). Let

$$S_2^* \equiv \left\{ u \in \mathbb{R}^m \,\middle|\, \sum_{i=1}^m u_i = 0,\ \sum_{i=1}^m |u_i| = 1 \right\}.$$

This set is regarded as the set of all non-trivial expected utilities on $\Delta(Z)$. Each $s_2 \in S_2$ is identified with a point in S_2^*. Let $\varphi : S_2 \to S_2^*$ be this identification mapping.

Proposition B.1. *Suppose either μ_1 has a finite support, or φ is continuous. Then $s_2 \in S_2$ is relevant if and only if s_2 belongs to the support of μ_1.*

Proof. (if part) Take any $\bar{s}_2$ in the support of μ_1. We want to show that $\bar{u} \equiv \varphi(\bar{s}_2)$ is relevant with respect to $U_1(\cdot, \mu_1)$. Let $x_1 \in \mathcal{K}(\Delta(Z))$ be a menu such that, for all $u \in S_2$, u has a unique maximizer on x_1. Let $\bar{l} \in x_1$ be the maximizer for $\bar{u}$. Take any $\varepsilon > 0$ and ε-ball $B_\varepsilon(\bar{u})$ in $\mathbb{R}^m$. Since $S \setminus B_\varepsilon(\bar{u})$ is compact,

$$v \equiv \min_{u \in S \setminus B_\varepsilon(u)} \left| \max_{l \in x_1} u(l) - u(\bar{l}) \right|$$

is a positive number. Let K be a positive number satisfying $\|u\| \le K$ for all $u \in B_\varepsilon(\bar{u})$. There exists a small $\delta > 0$ such that

$$|u(l) - u(\bar{l})| \le K \|l - \bar{l}\| < v,$$

for all $l \in B_\delta(\bar{l})$. Then, for all $u \in S \setminus B_\varepsilon(\bar{u})$ and $l \in B_\delta(\bar{l})$,

$$\max_{l \in x_1} u(l) \ge u(\bar{l}) + v > u(l).$$

Since $\bar{l}$ is a unique maximizer for $\bar{u}$, there exists a sufficiently small $\alpha > 0$ such that any $l \in x_1 \setminus \{\bar{l}\}$ with $\bar{u}(\bar{l}) - \bar{u}(l) < \alpha$ satisfies $l \in B_\delta(\bar{l})$. Now define $y_1 \equiv x_1 \cap \{l \in \Delta(Z) | \bar{u}(l) \le \bar{u}(\bar{l}) - \alpha\}$. Since $\max_{l' \in x_1} u(l') > u(l)$ for all $l \in x_1 \setminus y_1$ and $u \in S \setminus B_\varepsilon(\bar{u})$, we have $\max_{x_1} u = \max_{y_1} u$ for all $u \in S \setminus B_\varepsilon(\bar{u})$. On the other hand, any u in a small neighborhood V of $\bar{u}$ satisfies $\max_{x_1} u > \max_{y_1} u$.

If the support of μ_1 is finite, $\mu_1(\bar{s}_2)$ is positive. Since $\max_{x_1} u(l, \bar{s}_2) > \max_{y_1} u(l, \bar{s}_2)$, the representation implies $U_1(x_1, \mu_1) > U_1(y_1, \mu_1)$. Thus $\bar{u}$ is relevant.

If φ is continuous, the inverse image of V is a neighborhood of $\bar{s}_2$. Since $\bar{s}_2$ belongs to the support of μ_1, this neighborhood has positive measure. Thus the representation implies $U_1(x_1, \mu_1) > U_1(y_1, \mu_1)$, and hence $\bar{u}$ is relevant.

(only-if part) For all x_1, y_1, if $\max_{x_1} u(l, s_2) = \max_{y_1} u(l, s_2)$ for all $s_2 \in \mathrm{supp}(\mu_1)$, then $U_1(x_1, \mu_1) = U_1(y_1, \mu_1)$. Hence any $\varphi(s_2)$ with $s_2 \notin \mathrm{supp}(\mu_1)$ is not relevant. $\square$

C Proofs

C.1 Proof of Theorem 3.1

C.1.1 Outline of the Proof

Necessity is routine. We show sufficiency, that is, $(a) \Rightarrow (b)$. As in DLR, subjective signals in period 2 are identified with vNM indices over $\Delta(Z)$; that is,

$$S_2 \equiv \left\{ u \in \mathbb{R}^m \middle| \sum_{i=1}^{m} u_i = 0,\ \sum_{i=1}^{m} |u_i| = 1 \right\}.$$

This identification is natural because the DM only cares about ex-post preferences conditional on signals rather than the signals themselves. Similarly, since subjective signals in period 1 are identified with beliefs over S_2, we can specify $S_1 \equiv \Delta(S_2)$.

A key step in the DLR's additive EU representation is characterization of a compact and convex menu via its support function. Similarly, we identify a menu of menus x_0 via its support function $\sigma_{x_0} : S_1 \to \mathbb{R}$ defined by

$$\sigma_{x_0}(\mu) \equiv \max_{x_1 \in x_0} \int_{S_2} \max_{l \in x_1} u(l) \mathrm{d}\mu(u).$$

Let $\overline{\mathrm{co}}(\cdot)$ be the closed convex hull operation. Let $\mathrm{CO}(x_0)$ be the set of all menus $\overline{\mathrm{co}}(y_1)$, where y_1 is a compact subset of x_1 and x_1 varies over $\overline{\mathrm{co}}(x_0)$. We show that: (i) $x_0 \sim \mathrm{CO}(x_0)$ and (ii)

$$\sigma_{\mathrm{CO}(x_0)} = \sigma_{\mathrm{CO}(y_0)} \ \Leftrightarrow\ \mathrm{CO}(x_0) = \mathrm{CO}(y_0).$$

The remaining part is to find a unique non-negative measure μ_0 over S_1 such that

$$U_0(\mathrm{CO}(x_0)) \equiv \int_{S_1} \sigma_{\mathrm{CO}(x_0)}(\mu)\, \mathrm{d}\mu_0(\mu)$$

represents preference over $\mathrm{CO}(x_0)$s. We adapt the argument in DLR.

C.1.2 Sufficiency

(i) We show that there exists the required functional form representing preference. As a preliminary result, we first provide a useful implication of Monotonicity and Aversion to Commitment. Say that x_0 *covers* y_0 if, for any $y_1 \in y_0$, there exists $x_1 \in x_0$ such that $y_1 \subset x_1$.

Lemma C.1. *If x_0 covers y_0, then $x_0 \succeq y_0$.*

Proof. Suppose otherwise. Then there exist x_0 and y_0 such that x_0 covers y_0 but $y_0 \succ x_0$. From Continuity and Lemma 0 (p. 1421) of Gul and Pesendorfer [3], there exists a finite subset $y_0^* \subset y_0$ such that $y_0^* \succ x_0$. Denote y_0^* by $\{y_1^i | i = 1, \cdots, I\}$.

Since x_0 covers y_0^*, for any $y_1^i \in y_0^*$, there exists $x_1^i \in x_0$ such that $y_1^i \subset x_1^i$. Let $z_0^i \equiv \{x_1^i \backslash y_1^i, y_1^i\}$. Since $y_0^* \subset \cup_{i=1}^I z_0^i$, Monotonicity implies $\cup_{i=1}^I z_0^i \succeq y_0^*$. By Aversion to Commitment,

$$\{x_1^1\} \cup (\cup_{i=2}^I z_0^i) \succeq \{x_1^1 \setminus y_1^1, y_1^1\} \cup (\cup_{i=2}^I z_0^i) = \cup_{i=1}^I z_0^i.$$

By repeating the same argument finite times, $x_0^* \equiv \{x_1^i | i = 1, \cdots, I\} \succeq \cup_{i=1}^I z_0^i$. Since $x_0^* \subset x_0$, Monotonicity implies $x_0 \succeq x_0^*$. Therefore, $x_0 \succeq y_0^*$. This is a contradiction. □

As Lemma 1 (p. 922) of DLR, Order, Continuity and Independence imply $x_0 \sim \overline{\text{co}}(x_0)$. We can restrict our attention to the sub-domain, $\mathcal{D}_1 \equiv \{x_0 \in \mathcal{D} | x_0 = \overline{\text{co}}(x_0)\}$. Then, $\mathcal{D}_1$ is a compact and convex space.

For any $x_0 \in \mathcal{D}$, let

$$\text{co}_1(x_0) \equiv \{\text{co}(x_1) \in \mathcal{K}(\Delta(Z)) | x_1 \in x_0\}.$$

That is, $\text{co}_1(x_0)$ is the set of all convex hulls $\text{co}(x_1)$ as x_1 varies over x_0. Notice that $\text{co}_1(x_0)$ and $\text{co}(x_0)$ are distinct objects.

Lemma C.2.

(i) *For all* $x_0 \in \mathcal{D}$, $\text{co}_1(x_0) \in \mathcal{D}$.

(ii) *If* $x_0 \in \mathcal{D}$ *is convex,* $\text{co}_1(x_0)$ *is convex.*

(iii) *The mapping,* $\text{co}_1 : \mathcal{D} \to \mathcal{D}$, *is Hausdorff continuous.*

(iv) *For all* $x_0 \in \mathcal{D}$, $x_0 \sim \text{co}_1(x_0)$.

Proof. (i) Consider the convex hull operator $\text{co}(\cdot) : \mathcal{K}(\Delta(Z)) \to \mathcal{K}(\Delta(Z))$. First of all, since $x_1 \in \mathcal{K}(\Delta(Z))$ is a compact subset of $(m-1)$-dimensional Euclidean space, $\text{co}(x_1)$ is also compact. Hence, this operator is well-defined.

In order to show the claim, it suffices to show that $\overline{\text{co}}(\cdot) : \mathcal{K}(\Delta(Z)) \to \mathcal{K}(\Delta(Z))$ is Hausdorff continuous. Since $\Delta(Z)$ is identified with the unit simplex in $\mathbb{R}^m$, the weak convergence topology on $\Delta(Z)$ is equivalent to the Euclidean metric on $\mathbb{R}^m$. Recall the following notation in Section A:

$$d(l, x_1') \equiv \min_{l' \in x_1'} d(l, l'), \text{ and } e(x_1, x_1') \equiv \max_{l \in x_1} d(l, x_1').$$

Step 1: For any convex menu x_1', $d(\cdot, x_1')$ is a convex function.

Take any $l_1, l_2 \in \Delta(Z)$ and $\lambda \in [0, 1]$. Let $\bar{l}_i \equiv \text{argmin}_{l' \in x_1'} d(l_i, l')$, $i = 1, 2$. Then, taking into account that d is the Euclidean norm,

$$\begin{aligned}
\lambda d(l_1, x_1') + (1-\lambda) d(l_2, x_1') &= d(\lambda l_1, \lambda \bar{l}_1) + d((1-\lambda) l_2, (1-\lambda) \bar{l}_2) \\
&\geq d(\lambda l_1 + (1-\lambda) l_2 \lambda \bar{l}_1 + (1-\lambda) \bar{l}_2) \\
&\geq \min_{l' \in x_1'} d(\lambda l_1 + (1-\lambda) l_2, l') \\
&= d(\lambda l_1 + (1-\lambda) l_2, x_1').
\end{aligned}$$

Thus $d(\cdot, x_1')$ is a convex function whenever x_1' is convex.

It follows from the next step that the convex hull operator is Hausdorff continuous.

Step 2: For all $x_1, y_1 \in \mathcal{K}(\Delta(Z))$,

$$d_h(\text{co}(x_1), \text{co}(y_1)) \leq d_h(x_1, y_1). \tag{9}$$

By definition, for any $l \in x_1$,

$$d(l, \text{co}(y_1)) \leq e(x_1, \text{co}(y_1)). \tag{10}$$

Since $d(\cdot, \text{co}(y_1))$ is a convex function by Step 1, (10) holds for any $l \in \text{co}(x_1)$. Thus,

$$e(\text{co}(x_1), \text{co}(y_1)) \leq e(x_1, \text{co}(y_1)).$$

Moreover, since $y_1 \subset \text{co}(y_1)$, $e(x_1, \text{co}(y_1)) \leq e(x_1, y_1)$. Hence, we have

$$e(\text{co}(x_1), \text{co}(y_1)) \leq e(x_1, y_1). \tag{11}$$

By the symmetric argument, (11) holds when x_1 and y_1 are reversed. Hence (9) holds.

(ii) From Dunford and Schwartz [2, Lemma 4 (iii) and (iv), p.415], $\text{co}(\cdot) : \mathcal{K}(\Delta(Z)) \to \mathcal{K}(\Delta(Z))$ is mixture linear, that is, for all $x_1, y_1 \in \mathcal{K}(\Delta(Z))$ and $\lambda \in [0, 1]$,

$$\text{co}(\lambda x_1 + (1 - \lambda)y_1) = \lambda \text{co}(x_1) + (1 - \lambda)\text{co}(y_1).$$

Since a mixture linear operator preserves convexity, $\text{co}_1(x_0)$ is convex as long as x_0 is convex.

(iii) Let $x_0^n \to x_0$ with $x_0^n, x_0 \in \mathcal{K}(\mathcal{K}(\Delta(Z)))$. We want to show $\text{co}_1(x_0^n) \to \text{co}_1(x_0)$. By definition,

$$d_H(\text{co}_1(x_0^n), \text{co}_1(x_0)) = \max \left[\max_{x_1 \in x_0^n} \min_{y_1 \in x_0} d_h(\text{co}(x_1), \text{co}(y_1)),\ \max_{y_1 \in x_0} \min_{x_1 \in x_0^n} d_h(\text{co}(x_1), \text{co}(y_1)) \right].$$

Condition (9) implies

$$d_H(\text{co}_1(x_0^n), \text{co}_1(x_0)) \leq \max \left[\max_{x_1 \in x_0^n} \min_{y_1 \in x_0} d_h(x_1, y_1),\ \max_{y_1 \in x_0} \min_{x_1 \in x_0^n} d_h(x_1, y_1) \right]. \tag{12}$$

By assumption, the right hand side of (12) converges to zero. Hence, $\text{co}_1(x_0^n) \to \text{co}_1(x_0)$.

(iv) Since $x_1 \subset \text{co}(x_1)$, Lemma C.1 implies $\text{co}_1(x_0) \succeq x_0$. We will show $x_0 \succeq \text{co}_1(x_0)$.

Step 1: If $x_0 \in \mathcal{D}$ is finite and if each element $x_1^i \in x_0$ is also finite, then there exits $\lambda \in (0, 1)$ such that $\text{co}(\text{co}_1(x_0)) \subset \lambda \text{co}(x_0) + (1 - \lambda)\text{co}(\text{co}_1(x_0))$.

Take $x_1 \in \text{co}(\text{co}_1(x_0))$. Since $\text{co}_1(x_0)$ is finite, x_1 can be written as a convex combination of elements of $\text{co}_1(x_0)$. That is, $x_1 = \sum_i \alpha_i \text{co}(x_1^i)$, where $x_1^i \in x_0$ and $\alpha_i > 0$ with $\sum_i \alpha_i = 1$. When x_1^i is finite, we can show as in Lemma 1 (p. 922) of DLR that, for all λ_i sufficiently small, $\text{co}(x_1^i) = \lambda_i x_1^i + (1 - \lambda_i)\text{co}(x_1^i)$. Since x_0 is finite, by taking a small $\lambda > 0$, $\text{co}(x_1^i) = \lambda x_1^i + (1 - \lambda)\text{co}(x_1^i)$ for all i. Then,

$$\begin{aligned} x_1 &= \sum_i \alpha_i (\lambda x_1^i + (1 - \lambda)\text{co}(x_1^i)) \\ &= \lambda \sum_i \alpha_i x_1^i + (1 - \lambda) \sum_i \alpha_i \text{co}(x_1^i). \end{aligned}$$

Since $\sum_i \alpha_i x_1^i \in \mathrm{co}(x_0)$ and $\sum_i \alpha_i \mathrm{co}(x_1^i) \in \mathrm{co}(\mathrm{co}_1(x_0))$, $x_1 \in \lambda \mathrm{co}(x_0) + (1-\lambda)\mathrm{co}(\mathrm{co}_1(x_0))$.

Step 2: For any finite $x_0 \in \mathcal{D}$ such that each element $x_1^i \in x_0$ is also finite, $x_0 \succeq \mathrm{co}_1(x_0)$.

Suppose $\mathrm{co}_1(x_0) \succ x_0$. Since $\mathrm{co}(\mathrm{co}_1(x_0)) \sim \mathrm{co}_1(x_0)$ and $\mathrm{co}(x_0) \sim x_0$, $\mathrm{co}(\mathrm{co}_1(x_0)) \succ \mathrm{co}(x_0)$. Independence implies, for any $\lambda \in (0,1]$,

$$\lambda \mathrm{co}(\mathrm{co}_1(x_0)) + (1-\lambda)\mathrm{co}(\mathrm{co}_1(x_0)) \succ \lambda \mathrm{co}(x_0) + (1-\lambda)\mathrm{co}(\mathrm{co}_1(x_0)). \tag{13}$$

On the other hand, Monotonicity and Step 1 imply that, for some $\lambda \in (0,1)$,

$$\lambda \mathrm{co}(x_0) + (1-\lambda)\mathrm{co}(\mathrm{co}_1(x_0)) \succeq \mathrm{co}(\mathrm{co}_1(x_0)).$$

Since $\lambda \mathrm{co}(\mathrm{co}_1(x_0)) + (1-\lambda)\mathrm{co}(\mathrm{co}_1(x_0)) = \mathrm{co}(\mathrm{co}_1(x_0))$, this ranking contradicts (13).

Step 3: For any $x_0 \in \mathcal{D}$, $x_0 \succeq \mathrm{co}_1(x_0)$.

Take any $x_0 \in \mathcal{D}$. By the property of Hausdorff metric, there exists a sequence $\{x_0^n\}_{n=1}^{\infty}$ such that (1) $x_0^n \to x_0$, (2) x_0^n is finite, and (3) each element of x_0^n is also finite. From Step 2, $x_0^n \succeq \mathrm{co}_1(x_0^n)$. Part (iii) and Continuity imply $x_0 \succeq \mathrm{co}_1(x_0)$. □

For all $x_0 \in \mathcal{D}$, let

$$I(x_1) \equiv \{y_1 \in \mathcal{K}(\Delta(Z)) \,|\, y_1 \subset x_1\}.$$

Let $I(x_0) \equiv \cup_{x_1 \in x_0} I(x_1)$. Thus, $I(x_0)$ is the set of all menus y_1, where y_1 is included in some menu in x_0.

Lemma C.3.

(i) *For all $x_0 \in \mathcal{D}$, $I(x_0) \in \mathcal{D}$.*

(ii) *If x_0 is convex, $I(x_0)$ is also convex.*

(iii) *The mapping, $I : \mathcal{D} \to \mathcal{D}$, is Hausdorff continuous.*

(iv) *For all $x_0 \in \mathcal{D}$, $x_0 \sim I(x_0)$.*

Proof. (i) Since $I(x_0) \subset \mathcal{K}(\Delta(Z))$, it suffices to show that $I(x_0)$ is closed. Let $x_1^n \to x_1$ with $x_1^n \in I(x_0)$. Then there is a sequence $\{y_1^n\}$ in x_0 satisfying $x_1^n \subset y_1^n$. Since x_0 is compact, without loss of generality we can assume y_1^n converges to some point $y_1 \in x_0$. Suppose there exists $l \in x_1 \setminus y_1$. Since y_1 is compact, there is an open neighborhood of l, denoted by $U(l)$, such that $U(l) \cap y_1 = \emptyset$. For all sufficiently large n, there is $l^n \in U(l) \cap x_1^n$ because $x_1^n \to x_1$. Since $x_1^n \subset y_1^n$, $l^n \in y_1^n$. This contradicts the fact that $y_1^n \to y_1$. Thus $x_1 \subset y_1$, which implies $x_1 \in I(x_0)$, and hence $I(x_0)$ is closed.

(ii) Take $x_1', x_1 \in I(x_0)$. Then there exist $y_1', y_1 \in x_0$ such that $x_1' \subset y_1'$ and $x_1 \subset y_1$. Since x_0 is convex, $\alpha y_1' + (1-\alpha)y_1 \in x_0$ for any $\alpha \in [0,1]$. Clearly, $\alpha x_1' + (1-\alpha)x_1 \subset \alpha y_1' + (1-\alpha)y_1$. Hence, $\alpha x'_1 + (1-\alpha)x_1 \in I(x_0)$.

(iii) First of all, the mapping $I(\cdot)$ is well-defined by part (i). Let $x_0^n \to x_0$. We have a sequence $\{I(x_0^n)\}_{n=0}^{\infty}$. Since $\mathcal{D}$ is a compact metric space, without loss of generality, assume $I(x_0^n) \to y_0$ for some $y_0 \in \mathcal{D}$. We want to show $I(x_0) = y_0$.

Step 1: $y_0 \subset I(x_0)$.

Take any $y_1 \in y_0$. There is $z_1 \in y_0$ such that $y_1 \subset z_1$. Since $I(x_0^n) \to y_0$, we can find a sequence $z_1^n \to z_1$ with $z_1^n \in I(x_0^n)$. Thus, there is a sequence $x_1^n \in x_0^n$ with $z_1^n \subset x_1^n$. Since the sequence $\{x_1^n\}$ is in $\mathcal{K}(\Delta(Z))$, we can assume $x_1^n \to x_1$ for some $x_1 \in \mathcal{K}(\Delta(Z))$. Since $x_0^n \to x_0$ and $x_1^n \to x_1$ with $x_1^n \in x_0^n$, we have $x_1 \in x_0$. Thus, $y_1 \in I(x_0)$ because $y_1 \subset z_1 \subset x_1$.

Step 2: $I(x_0) \subset y_0$.

Take any $z_1 \in I(x_0)$. There exists $x_1 \in x_0$ such that $z_1 \subset x_1$. Since $x_0^n \to x_0$, there exists a sequence $x_1^n \in x_0^n$ with $x_1^n \to x_1$. From a property of the Hausdorff metric, there exists a sequence $z_1^m \to z_1$ such that z_1^m is a finite subset of z_1. Take the open $1/m$-neighborhood of z_1, denoted by $B(z_1, 1/m)$. We can assume without loss of generality that $z_1^m \in B(z_1, 1/m)$ for all $m \geq 1$. Since z_1^m is a finite subset of x_1, there exists a finite subset $y_1^{n_m} \subset x_1^{n_m}$ such that $y_1^{n_m} \in B(z_1, 1/m)$. Since $\mathcal{K}(\mathcal{H})$ is compact, the subsequence $\{y_1^{n_m}\}_{m=0}^{\infty}$ converges to z_1. Since $I(x_0^{n_m}) \to y_0$ and $y_1^{n_m} \to z_1$ with $y_1^{n_m} \in I(x_0^{n_m})$, we have $z_1 \in y_0$.

(iv) Since $x_0 \subset I(x_0)$, Monotonicity implies $I(x_0) \succeq x_0$. Since x_0 covers $I(x_0)$, Lemma C.1 implies $x_0 \succeq I(x_0)$. Thus, $x_0 \sim I(x_0)$. □

From Lemma C.2 (iv) and C.3 (iv), we can pay attention the sub-domain

$$\mathcal{D}_2 \equiv \{x_0 \in \mathcal{D}_1 | x_0 = \mathrm{CO}(x_0)\}.$$

where $\mathrm{CO}(x_0) = \mathrm{co}_1(I(\mathrm{co}_1(x_0)))$. From Lemma C.2 (ii), (iii), C.3 (ii) and (iii), $\mathrm{CO}(\cdot)$ is a Hausdorff continuous operator from $\mathcal{D}_1$ into itself. Thus, D_2 is compact. By definition, $x_1 \in x_0 \in \mathcal{D}_2$ is convex. Moreover, if y_1 is convex and if $y_1 \subset x_1 \in x_0$, then $y_1 \in x_0$.

Order, Continuity, Nondegeneracy and Independence ensure a non-constant mixture linear representation $U : \mathcal{D}_1 \to \mathbb{R}$ because $\mathcal{D}_1$ is a mixture space.

Let

$$S_2 \equiv \left\{ u \in \mathbb{R}^m \middle| \sum_{i=1}^{m} u_i = 0,\ \sum_{i=1}^{m} |u_i| = 1 \right\}.$$

Let $\mathcal{C}(S_2)$ be the set of all real-valued continuous functions on S_2 with the sup-norm. Let $\mathcal{K}_c(\Delta(Z))$ be the set of all compact and convex subsets of $\Delta(Z)$. Notice that $\mathcal{K}_c(\Delta(Z))$ is compact and convex. For all $x_1 \in \mathcal{K}_c(\Delta(Z))$ and $u \in S_2$, define

$$\zeta_{x_1}(u) \equiv \max_{l \in x_1} u(l).$$

Lemma 3, 4, and 8 of DLR show that the function $\zeta : \mathcal{K}_c(\Delta(Z)) \to \mathcal{C}(S_2)$ is injective, mixture linear, and continuous.

Let $S_1 \equiv \Delta(S_2)$ with the weak convergence topology. Let $\mathcal{C}(S_1)$ be the set of all real-valued continuous functions on S_1. For all $x_1 \in \mathcal{K}_c(\Delta(Z))$ and $\mu \in S_1$, let

$$U_1(x_1, \mu) \equiv \int_{S_1} \zeta_{x_1}(u)\,\mathrm{d}\mu(u) = \int_{S_1} \max_{l \in x_1} u(l)\,\mathrm{d}\mu(u).$$

For all $x_0 \in \mathcal{D}_2$ and $\mu \in S_1$, let

$$\sigma_{x_0}(\mu) \equiv \max_{x_1 \in x_0} U_1(x_1, \mu).$$

Now we have the function $\sigma : \mathcal{D}_2 \to \mathcal{C}(S_1)$.

Lemma C.4.

(i) *σ is injective, that is, $\sigma_{x_0} = \sigma_{y_0} \Rightarrow x_0 = y_0$.*

(ii) *For all $x_0, y_0 \in \mathcal{D}_2$, $\lambda\sigma_{x_0} + (1-\lambda)\sigma_{y_0} = \sigma_{\mathrm{CO}(\lambda x_0 + (1-\lambda) y_0)}$.*

(iii) *σ is continuous.*

Proof. (i) Let $x_0 \neq y_0$. Since the symmetric argument works, assume that $x_0 \not\subset y_0$. Then there exists $\bar{x}_1 \in x_0 \setminus y_0$. Since ζ is injective, $\zeta_{\bar{x}_1} \in \zeta(x_0) \setminus \zeta(y_0)$, where $\zeta(x_0)$ and $\zeta(y_0)$ are the images of x_0 and of y_0 under ζ, respectively.

Step 1: $\zeta(y_0) \cap (\mathcal{C}_+(S_2) + \{\zeta_{\bar{x}_1}\}) = \emptyset$, where $\mathcal{C}_+(S_2) \subset \mathcal{C}(S_2)$ is the set of all non-negative continuous functions on S_2.

Suppose otherwise. Then there exists $\zeta_{y_1} \in \zeta(y_0)$ such that $\zeta_{y_1}(u) \geq \zeta_{\bar{x}_1}(u)$ for all $u \in S_2$ and $\zeta_{y_1}(\underline{u}) > \zeta_{\bar{x}_1}(\underline{u})$ for some $\underline{u} \in S_2$. Hence, $\bar{x}_1 \subset y_1$. By definition of $\mathcal{D}_2$, $\bar{x}_1 \in y_0$. This is a contradiction.

Step 2: There exist a linear functional Λ on $\mathcal{C}(S_2)$ and a constant $c \in \mathbb{R}$ such that $\Lambda(\zeta_{\bar{x}_1}) > c > \Lambda(\zeta_{y_1})$ for all $\zeta_{y_1} \in \zeta(y_0)$.

Since ζ is continuous and mixture linear, $\zeta(y_0)$ is compact and convex. Notice that $\mathcal{C}_+(S_2) + \{\zeta_{\bar{x}_1}\}$ is a closed subset of $\mathcal{C}(S_2)$. Moreover, from Step 1, these subsets are disjoint. By the separation hyperplane theorem (See Schaefer [11, p.65, Theorem 9.2 (Second Separation Theorem)]), there exist a linear functional Λ on $\mathcal{C}(S_2)$ and a constant $c \in \mathbb{R}$ strictly separating these sets. Since the constant function equal to zero belongs to $\mathcal{C}_+(S_2)$, $\Lambda(\zeta_{\bar{x}_1}) > c > \Lambda(\zeta_{y_1})$ for all $\zeta_{y_1} \in \zeta(y_0)$.

Step 3: Λ is positive, that is, $\Lambda(f_+) \geq 0$ if $f_+ \in \mathcal{C}_+(S_2)$.

From Step 2, $\Lambda(f_+) > \Lambda(\zeta_{y_1} - \zeta_{\bar{x}_1})$ for all $f_+ \in \mathcal{C}_+(S_2)$ and $\zeta_{y_1} \in \zeta(y_0)$. This means that Λ is bounded from below on $\mathcal{C}_+(S_2)$. Take a lower bound $\alpha \in \mathbb{R}$. Then $\Lambda(f_+) \geq \alpha$ for all $f_+ \in \mathcal{C}_+(S_2)$. Suppose that Λ is not positive. There exists $\bar{f}_+ \in \mathcal{C}_+(S_2)$ with $\Lambda(\bar{f}_+) < 0$. Since $\theta\bar{f}_+ \in \mathcal{C}_+(S_2)$ for all $\theta > 0$, $\Lambda(\theta\bar{f}_+) = \theta\Lambda(\bar{f}_+)$ diverges to $-\infty$ as θ tends to ∞. This contradicts the fact that $\Lambda(f_+) \geq \alpha$ for all $f_+ \in \mathcal{C}_+(S_2)$.

The Riesz Representation theorem (See Rudin [8, p40, Theorem 2.14]) ensures the existence of a positive measure ν on S_2 satisfying

$$\Lambda(f) = \int_{S_2} f(u)\,\mathrm{d}\nu(u) \quad \text{for all } f \in \mathcal{C}(S_2).$$

Let $\bar{\mu} \in \Delta(S_2)$ be the normalization of ν. Since $\Lambda(\zeta_{\bar{x}_1}) > c > \Lambda(\zeta_{y_1})$ for all $\zeta_{y_1} \in \zeta(y_0)$, we have

$$\int_{S_2} \zeta_{\bar{x}_1}(u)\,\mathrm{d}\bar{\mu}(u) > c \geq \max_{y_1 \in y_0} \int_{S_2} \zeta_{y_1}(u)\,\mathrm{d}\bar{\mu}(u).$$

Thus,

$$\sigma_{x_0}(\bar{\mu}) = \max_{x_1 \in x_0} U_1(x_1, \bar{\mu}) > \max_{y_1 \in y_0} U_1(y_1, \bar{\mu}) = \sigma_{y_0}(\bar{\mu}).$$

Since $\sigma_{x_0} \neq \sigma_{y_0}$, σ is injective.

(ii) For each $\mu \in S_1$, let x_1^* and y_1^* satisfy $U_1(x_1^*, \mu) = \max_{x_1 \in x_0} U_1(x_1, \mu)$ and $U_1(y_1^*, \mu) = \max_{y_1 \in y_0} U_1(y_1, \mu)$. Since $\lambda x_1^* + (1-\lambda) y_1^* \in \lambda x_0 + (1-\lambda) y_0$, mixture linearity of $U_1(\cdot, \mu)$ implies,

$$\begin{aligned}
\lambda \sigma_{x_0}(\mu) + (1-\lambda)\sigma_{y_0}(\mu) &= \lambda U_1(x_1^*, \mu) + (1-\lambda) U_1(y_1^*, \mu) \\
&= U_1(\lambda x_1^* + (1-\lambda) y_1^*, \mu) \\
&= \max_{z_1 \in \lambda x_0 + (1-\lambda) y_0} U_1(z_1, \mu) \\
&= \max_{z_1 \in \mathrm{CO}(\lambda x_0 + (1-\lambda) y_0)} U_1(z_1, \mu) \\
&= \sigma_{\mathrm{CO}(\lambda x_0 + (1-\lambda) y_0)}(\mu).
\end{aligned}$$

(iii) It suffices to show that, for all $x_0, y_0 \in \mathcal{D}_2$,

$$d_{\mathrm{supnorm}}(\sigma_{x_0}, \sigma_{y_0}) \leq d_{\mathrm{Hausdorff}}(x_0, y_0).$$

For any fixed $\mu \in S_1$,

$$\begin{aligned}
|\sigma_{x_0}(\mu) - \sigma_{y_0}(\mu)| &= \left| \max_{x_1 \in x_0} \int_{s_2} \zeta_{x_1}(u)\,\mathrm{d}\mu(u) - \max_{y_1 \in y_0} \int_{s_2} \zeta_{y_1}(u)\,\mathrm{d}\mu(u) \right| \\
&= \left| \int_{s_2} \zeta_{x_1^*}(u)\,\mathrm{d}\mu(u) - \int_{s_2} \zeta_{y_1^*}(u)\,\mathrm{d}\mu(u) \right|,
\end{aligned}$$

where $x_1^* \in x_0$ and $y_1^* \in y_0$ are maximizers. Without loss of generality, assume

$$\int_{s_2} \zeta_{x_1^*}(u)\,\mathrm{d}\mu(u) \geq \int_{s_2} \zeta_{y_1^*}(u)\,\mathrm{d}\mu(u). \tag{14}$$

Let $\bar{y}_1$ be a minimizer of the following problem:

$$\begin{aligned}
&\min_{y_1 \in z_0} \; d_{\mathrm{supnorm}}(\zeta_{y_1}, \zeta_{x_1^*}), \\
&z_0 \equiv \left\{ y_1 \in \mathcal{K}_c(\Delta(Z)) \,\middle|\, \int_{S_2} \zeta_{y_1}(u)\,\mathrm{d}\mu(u) \leq \int_{S_2} \zeta_{y_1^*}(u)\,\mathrm{d}\mu(u) \right\}.
\end{aligned}$$

Since z_0 is compact, $\bar{y}_1$ indeed exists. Notice that $y_0 \subset z_0$ by definition of y_1^*. Notice also that (14) implies

$$\int_{S_2} \zeta_{\bar{y}_1}(u)\,\mathrm{d}\mu(u) = \int_{S_2} \zeta_{y_1^*}(u)\,\mathrm{d}\mu(u).$$

Taking into account Lemma 8 (p.927) of DLR,

$$\begin{aligned}
|\sigma_{x_0}(\mu) - \sigma_{y_0}(\mu)| &= \left| \int_{s_2} \zeta_{x_1^*}(u)\,\mathrm{d}\mu(u) - \int_{s_2} \zeta_{\bar{y}_1}(u)\,\mathrm{d}\mu(u) \right| \\
&\leq \int \left|\zeta_{x_1^*}(u) - \zeta_{\bar{y}_1}(u)\right| \mathrm{d}\mu(u) \\
&\leq d_{\text{supnorm}}(\zeta_{x_1^*}, \zeta_{\bar{y}_1}) \\
&\leq \min_{y_1 \in y_0} d_{\text{supnorm}}(\zeta_{x_1^*}, \zeta_{y_1}) \\
&= \min_{y_1 \in y_0} d_{\text{Hausdorff}}(x_1^*, y_1) \\
&\leq d_{\text{Hausdorff}}(x_0, y_0).
\end{aligned}$$

Since this inequality holds for all $\mu \in S_1$, we have

$$d_{\text{supnorm}}(\sigma_{x_0}, \sigma_{y_0}) \equiv \sup_{\mu \in S_1} |\sigma_{x_0}(\mu) - \sigma_{y_0}(\mu)| \leq d_{\text{Hausdorff}}(x_0, y_0).$$

□

Let $C \subset \mathcal{C}(S_1)$ be the range of σ.

Lemma C.5.

(i) *C is convex.*

(ii) *The constant function equal to zero, that is, $f(\mu) = 0$ for any $\mu \in S_1$, is in C.*

(iii) *There exists $c > 0$ such that the constant function equal to c, that is, $f(\mu) = c$ for any $\mu \in S_1$, is in C.*

(iv) *The supremum of any two points $f, f' \in C$ is also in C. That is,* $\max[f(\mu), f'(\mu)]$ *is also in C.*

Proof. (i) Take any $f, f' \in C$ and $\lambda \in [0,1]$. There are $x_0, x_0' \in \mathcal{D}_2$ satisfying $f = \sigma_{x_0}$ and $f' = \sigma_{x_0'}$. From Lemma C.4 (ii),

$$\lambda f + (1-\lambda) f' = \lambda \sigma_{x_0} + (1-\lambda)\sigma_{x_0'} = \sigma_{\mathrm{CO}(\lambda x_0 + (1-\lambda) x_0')} \in C.$$

Hence, C is convex.

(ii) Let $x_0 \equiv \{\{(1/m, \cdots, 1/m)\}\} \in \mathcal{D}_2$. Since $u((1/m, \cdots, 1/m)) = 0$ for all $u \in S_2$, $\sigma_{x_0}(\mu) = 0$ for all $\mu \in S_1$.

(iii) Let c be a sufficiently small positive number. Let x_1 be the intersection of $\Delta(Z)$ and the closed c-ball at $(1/m, \cdots, 1/m)$. Let $x_0 \equiv \mathrm{CO}(\{x_1\})$. Since $\max_{l \in x_1} u(l) = c$ for all $u \in S_2$, $\sigma_{x_0}(\mu) = c$ for all $\mu \in S_1$.

(iv) Take any $f, f' \in C$. There exist $x_0, x_0' \in \mathcal{D}_2$ such that $f = \sigma_{x_0}$ and $f' = \sigma_{x_0'}$. Let $x_0'' \equiv \mathrm{CO}(\mathrm{co}(x_0 \cup x_0')) \in \mathcal{D}_2$ and $f'' \equiv \sigma_{x_0''} \in C$. Then, $f''(\mu) = \max[\sigma_{x_0}(\mu), \sigma_{x_0'}(\mu)]$. □

From Lemma C.4 (i), $\sigma : \mathcal{D}_2 \to C$ is bijective. Define $W : C \to \mathbb{R}$ by $W(f) \equiv U(\sigma^{-1}(f))$. From Lemma C.5 (ii) and (iii), $W(0) = 0$ and $W(c) = c$, where 0 and c are identified with the constant function equal to zero and with the constant function equal to c, respectively. Since U and σ are continuous, so is W under the sup-norm. By adapting DLR's argument, we can verify that W is linear in the following sense: for any $\alpha, \beta \in \mathbb{R}_+$, if $f, f', \alpha f + \beta f' \in C$, then

$$W(\alpha f + \beta f') = \alpha W(f) + \beta W(f').$$

We will extend the function $W : C \to \mathbb{R}$ to $\mathcal{C}(S_1)$ step by step. First, restrict W on $C_+ \equiv \{f \in C | f \geq 0\}$. For any $r \geq 0$, let $rC_+ \equiv \{rf | f \in C_+\}$. Let $H \equiv \cup_{r \geq 0} rC_+$ and

$$H^* \equiv H - H = \{f_1 - f_2 \in \mathcal{C}(S_1) | f_1, f_2 \in H\}.$$

For any $f \in H \setminus 0$, there is $r > 0$ satisfying $(1/r)f \in C_+$. Define $W(f) \equiv rW((1/r)f)$. Then, $W : H \to \mathbb{R}$ is well-defined, monotonic, and linear. For any $f \in H^*$, there are $f_1, f_2 \in H$ satisfying $f = f_1 - f_2$. Define $W(f) \equiv W(f_1) - W(f_2)$. Then, $W : H^* \to \mathbb{R}$ is well-defined and linear.

Lemma C.6. *H^* is dense in $\mathcal{C}(S_1)$.*

Proof. From the Stone-Weierstrass theorem, it is enough to show that (i) H^* is a vector sublattice, (ii) for any distinct points $\mu, \mu' \in S_1$, there exists $f \in H^*$ such that $f(\mu) > f(\mu')$, and (iii) H^* contains the constant functions equal to one. By the exactly same argument as Lemma 11 (p.928) in DLR, condition (i) holds. Condition (iii) directly follows from Lemma C.5 (iii) and the definition of H.

To show condition (ii), take distinct points $\mu, \mu' \in S_1$. By the separating hyperplane theorem, there exists a linear functional Γ on S_1 and a constant $c \in \mathbb{R}$ such that $\Gamma(\mu) > c > \Gamma(\mu')$. Without loss of generality, we can assume $c = 0$. Since $\mathcal{C}(S_2)$ is a weak* dense subset of the dual space of S_1 (Dunford and Schwartz [2, Corollary 6, p. 425]), there exists $f \in \mathcal{C}(S_2)$ such that

$$\int_{S_2} f \,\mathrm{d}\mu > 0 > \int_{S_2} f \,\mathrm{d}\mu'.$$

We can assume $\|f\|$ is sufficiently small. From Lemma 11 (p. 928) of DLR, there exist $x_1, y_1 \in \mathcal{K}_c(\Delta(Z))$ such that

$$\int_{S_2} (\zeta_{x_1} - \zeta_{y_1}) \,\mathrm{d}\mu > 0 > \int_{S_2} (\zeta_{x_1} - \zeta_{y_1}) \,\mathrm{d}\mu'.$$

Hence,

$$\int_{S_2} \zeta_{x_1} \,\mathrm{d}\mu > \int_{S_2} \zeta_{y_1} \,\mathrm{d}\mu, \text{ and } \int_{S_2} \zeta_{y_1} \,\mathrm{d}\mu' > \int_{S_2} \zeta_{x_1} \,\mathrm{d}\mu'. \tag{15}$$

If

$$\int_{S_2} \zeta_{x_1} \,\mathrm{d}\mu = \int_{S_2} \zeta_{y_1} \,\mathrm{d}\mu',$$

redefine x_1 as the menu $\{l \in \Delta(Z) \,|\, d(l, x_1) \leq \varepsilon\}$ for some small $\varepsilon > 0$. Then,

$$\int_{S_2} \zeta_{x_1} \, \mathrm{d}\mu > \int_{S_2} \zeta_{y_1} \, \mathrm{d}\mu'. \tag{16}$$

Moreover, as long as $\varepsilon > 0$ is small enough, (15) still holds after this modification. Let $x_0 \equiv \mathrm{CO}(\mathrm{co}(\{x_1, y_1\}))$. Taking (15) and (16) together,

$$\sigma_{x_0}(\mu) = \int_{S_2} \zeta_{x_1} \, \mathrm{d}\mu > \int_{S_2} \zeta_{y_1} \, \mathrm{d}\mu' = \sigma_{x_0}(\mu').$$

Since $\sigma_{x_0} \in H^*$, condition (ii) holds.

□

By the same argument as in Lemma 12 (p. 929) of DLR, it can be shown that there is a constant $K > 0$ such that $W(f) \leq K||f||$ for any $f \in H^*$, because $\mathcal{D}_2$ is compact. By the Hahn-Banach theorem, we can extend W to $\overline{W} : \mathcal{C}(S_1) \to \mathbb{R}$ in a linear, continuous and increasing way. Uniqueness of this extension follows from Lemma C.6.

Since $\overline{W}$ is a positive linear functional on $\mathcal{C}(S_1)$, the Riesz representation theorem (Dunford and Schwartz [2, p.265, Theorem 3]) ensures that there exits a unique countably additive non-negative measure μ_0 on S_1 satisfying

$$\overline{W}(f) = \int_{S_1} f(\mu) \, \mathrm{d}\mu_0(\mu),$$

for all $f \in \mathcal{C}(S_1)$. Especially, μ_0 can be taken to be a probability measure. Thus, for any $x_0 \in \mathcal{D}_2$,

$$U(x_0) = \overline{W}(\sigma_{x_0}) = \int_{S_1} \sigma_{x_0}(\mu) \, \mathrm{d}\mu_0(\mu) = \int_{S_1} \max_{x_1 \in x_0} \int_{S_2} \max_{l \in x_1} u(l) \, \mathrm{d}\mu(u) \, \mathrm{d}\mu_0(\mu). \tag{17}$$

For any probability measure ν, let $\mathrm{supp}(\nu)$ denote the support of ν. Redefine $S_1 \equiv \mathrm{supp}(\mu_0)$ and Define $\mu_1 : S_1 \to \Delta(S_2)$ as the identity mapping, that is, $\mu_1(\mu) = \mu$. Define $u^* : \Delta(Z) \times S_2 \to \mathbb{R}$ by $u^*(l, u) = u(l)$. Then, $(\{S_t\}_{t=1}^2, \{\mu_t\}_{t=0}^1, u^*)$ is the required functional form representing $\succeq$.

(ii) By definition, μ_0 has full support.

(iii) We will show that, if $U_1(\cdot, \mu)$ and $U_1(\cdot, \mu')$ represent the identical preference, then $\mu = \mu'$. For all $u \in S_2$ and $x_1 \in \mathcal{K}(\Delta(Z))$, define

$$\zeta_{x_1}(u) \equiv \max_{l \in x_1} u(l).$$

Then,

$$U_1(x_1, \mu) = \int \zeta_{x_1}(u) \, \mathrm{d}\mu(u), \text{ and } U_1(x_1, \mu') = \int \zeta_{x_1}(u) \, \mathrm{d}\mu'(u). \tag{18}$$

For all $l \in \Delta(Z)$,

$$U_1(\{l\}, \mu) = \int u(l) \, \mathrm{d}\mu(u) = \bar{u}(l),$$

where $\bar{u}$ is the mean vNM-index with respect to μ. Precisely, for all $i = 1, \cdots, m$,

$$\bar{u}_i \equiv \int u_i \, \mathrm{d}\mu(u).$$

By definition, $\bar{u} \in S_2$. Similarly, $\bar{u}' \in S_2$ is defined as the mean vNM index with respect to μ'. Since $\bar{u}$ and $\bar{u}'$ represent the identical preference over $\Delta(Z)$, $\bar{u} = \bar{u}'$.

Since $U_1(\cdot, \mu)$ and $U_1(\cdot, \mu')$ are mixture linear functions over $\mathcal{K}(\Delta(Z))$ representing the identical preference, there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $U_1(\cdot, \mu') = \alpha U_1(\cdot, \mu) + \beta$. For any lottery l,

$$\begin{aligned} U_1(\{l\}, \mu') &= \alpha U_1(\{l\}, \mu) + \beta \\ \bar{u}'(l) &= \alpha \bar{u}(l) + \beta. \end{aligned}$$

Since $\bar{u} = \bar{u}'$, we must have $\alpha = 1$ and $\beta = 0$, and hence $U_1(\cdot, \mu) = U_1(\cdot, \mu')$. From (18), for all x_1,

$$\int \zeta_{x_1}(u) \, \mathrm{d}\mu(u) = \int \zeta_{x_1}(u) \, \mathrm{d}\mu'(u). \tag{19}$$

Take any $x_1, y_1 \in \mathcal{K}(\Delta(Z))$ and $\alpha, \beta \geq 0$. Equation (19) holds even when ζ_{x_1} is replaced with $\alpha\zeta_{x_1} - \beta\zeta_{y_1}$. From Lemma 11 of DLR, the set of all such functions is a dense subset of the set of all real-valued continuous functions on S_2. Hence, equation (19) holds for any real-valued continuous function instead of ζ_{x_1}. The Riesz representation theorem (Dunford and Schwartz [2, Theorem 3, p. 265]) implies $\mu = \mu'$.

(iv) For any $\mu \in S_1$, let $S_2(\mu) \equiv \mathrm{supp}(\mu)$ and $u^* : \Delta(Z) \times S_2(\mu) \to \mathbb{R}$ be the restriction of u^* on $\Delta(Z) \times S_2(\mu)$. Then, $U_1(\cdot, \mu)$ with components $(S_2(\mu), \mu, u^*)$ represents $\succeq_\mu$. By definition, μ has full support. Since $S_2 = \{u \in \mathbb{R}^m | \sum_{i=1}^m u_i = 0, \, \sum_{i=1}^m |u_i| = 1\}$, $u^*(\cdot, u)$ and $u^*(\cdot, u')$ represent distinct vNM preferences whenever $u \neq u'$. That is, $\succeq_u \neq \succeq_{u'}$. Finally, Proposition B.1 ensures that every $u \in S_2(\mu)$ is relevant because the identification mapping φ is the identity mapping in this case and hence is continuous.

C.2 Proof of Proposition 4.1

(i) Suppose $\mathbb{S}_2^1 \neq \mathbb{S}_2^2$. Assume there exists $\succeq_{\bar{s}_2} \in \mathbb{S}_2^1 \setminus \mathbb{S}_2^2$.[4] We can find a set of positive numbers $\bar{v}$ and $\{v_{s_2}\}_{s_2 \in S_2^2}$ so that

$$x_1 \equiv \left\{l \in \Delta(Z) | u^1(l, \bar{s}_2) \leq \bar{v}\right\} \cap \left(\cap_{s_2 \in S_2^2} \left\{l \in \Delta(Z) | u^2(l, s_2) \leq v_{s_2}\right\}\right)$$

is a non-empty, compact and convex menu such that each lower contour set coincides with a non-trivial part of the boundary of x_1.

For $\varepsilon > 0$, let

$$y_1 \equiv \{l \in \Delta(Z) | u^1(l, \bar{s}_2) \leq \bar{v} - \varepsilon\} \cap \left(\cap_{s_2 \in S_2^2} \left\{l \in \Delta(Z) | u^2(l, s_2) \leq v_{s_2}\right\}\right).$$

Since $\succeq_{\bar{s}_2} \notin \mathbb{S}_2^2$, there exists a sufficiently small $\varepsilon > 0$,

$$\int_{S_2^2} \max_{l \in x_1} u^2(l, s_2) \mathrm{d}\mu_1^2(s_2 | s_1) = \int_{S_2^2} \max_{l \in y_1} u^2(l, s_2) \mathrm{d}\mu_1^2(s_2 | s_1),$$

[4] The symmetric argument works when there exists $\succeq_{\bar{s}_2} \in \mathbb{S}_2^2 \setminus \mathbb{S}_2^1$.

for any $s_1 \in S_1^2$. Hence the representation U_0^2 implies $\{x_1\} \sim \{y_1\}$.

On the other hand, since $y_1 \subset x_1$,

$$\int_{S_2^1} \max_{l \in x_1} u^1(l, s_2) \mathrm{d}\mu_1^1(s_2|s_1) \geq \int_{S_2^1} \max_{l \in y_1} u^1(l, s_2) \mathrm{d}\mu_1^1(s_2|s_1), \tag{20}$$

for any $s_1 \in S_1^1$. Since $\succeq_{\bar{s}_2} \in \mathbb{S}_2^1$, there exists at least one $s_1 \in S_1^1$ such that (20) holds with strict inequality. The representation U_0^1 implies $\{x_1\} \succ \{y_1\}$. This is a contradiction.

(ii) Let

$$S_2^* \equiv \left\{ u \in \mathbb{R}^m \,\middle|\, \sum_{i=1}^m u_i = 0,\ \sum_{i=1}^m |u_i| = 1 \right\}.$$

This is a normalization of vNM-indices on $\Delta(Z)$. From Definition 3.2 (iii), for any $s_2 \in S_2^i$, there exists a unique $u \in S_2^*$ such that both $u^i(\cdot, s_2)$ and u generate the same vNM preference over $\Delta(Z)$. There exist $\alpha_{s_2}^i > 0$ and $\beta_{s_2}^i \in \mathbb{R}$ such that $u^i(\cdot, s_2) = \alpha_{s_2}^i u(\cdot) + \beta_{s_2}^i$. We can define the injective mapping $\varphi^i : S_2^i \to S_2^*$ by the above relation for $i = 1, 2$.

For any x_1 and $s_1 \in S_1^i$,

$$\begin{aligned}
U_1^i(x_1, s_1) &= \sum_{S_2^i} \max_{l \in x_1} u^i(l, s_2) \mu_1^i(s_2|s_1) = \sum_{S_2^i} \max_{l \in x_1} (\alpha_{s_2}^i u(l) + \beta_{s_2}^i) \mu_1^i(s_2|s_1) \\
&= \sum_{S_2^i} \max_{l \in x_1} \alpha_{s_2}^i u(l) \mu_1^i(s_2|s_1) + \sum_{S_2^i} \beta_{s_2}^i \mu_1^i(s_2|s_1) \\
&= \sum_{S_2^i} \max_{l \in x_1} u(l) \alpha_{s_2}^i \mu_1^i(s_2|s_1) + \sum_{S_2^i} \beta_{s_2}^i \mu_1^i(s_2|s_1) \\
&= \sum_{S_2^*} \max_{l \in x_1} u(l) \mu^i(u|s_1) + \sum_{S_2^i} \beta_{s_2}^i \mu_1^i(s_2|s_1),
\end{aligned}$$

where $\mu^i(u|s_1)$ is the non-negative measure on S_2^* induced by $\varphi^i : (S_2^i, \alpha_{(\cdot)}^i \mu_1^i(\cdot|s_1)) \to S_2^*$. Let $\mu_{s_1}^{*i}(u) \equiv \mu^i(u|s_1)/\mu^i(S_2^*|s_1)$ and

$$U_1(x_1, \mu_{s_1}^{*i}) \equiv \sum_{S_2^*} \max_{l \in x_1} u(l) \mu_{s_1}^{*i}(u).$$

Since $U_1^i(x_1, s_1) \geq U_1^i(y_1, s_1)$ if and only if $U_1(x_1, \mu_{s_1}^{*i}) \geq U_1(y_1, \mu_{s_1}^{*i})$, both $U_1^i(\cdot, s_1)$ and $U_1(\cdot, \mu_{s_1}^{*i})$ induce the same preference on $\mathcal{K}(\Delta(Z))$.

Define the injective mapping $\psi^i : S_1^i \to \Delta(S_2^*)$ by $\psi^i(s_1) \equiv \mu_{s_1}^{*i}$. Then,

$$\begin{aligned}
U_0^i(x_0) &= \sum_{S_1^i} \max_{x_1 \in x_0} U_1^i(x_1, s_1) \mu_0^i(s_1) \\
&= \sum_{S_1^i} \max_{x_1 \in x_0} U_1(x_1, \mu_{s_1}^{*i}) \mu^i(S_2^*|s_1) \mu_0^i(s_1) + \sum_{S_1^i} \sum_{S_2^i} \beta_{s_2}^i \mu_1^i(s_2|s_1) \mu_0^i(s_1) \\
&= \sum_{\Delta(S_2^*)} \max_{x_1 \in x_0} U_1(x_1, \mu) \bar{\mu}_0^i(\mu) + \sum_{S_1^i} \sum_{S_2^i} \beta_{s_2}^i \mu_1^i(s_2|s_1) \mu_0^i(s_1),
\end{aligned}$$

where $\bar{\mu}_0^i$ is the non-negative measure induced by the mapping $\psi^i : (S_1^i, \mu^i(S_2^*|\cdot)\mu_0^i(\cdot)) \to \Delta(S_2^*)$.

Let $\mu_0^{*i} \equiv \bar{\mu}_0^i / \bar{\mu}_0^i(S_1^*) \in \Delta(\Delta(S_2^*))$ and

$$U_0^{*i}(x_0) \equiv \sum_{\Delta(S_2^*)} \max_{x_1 \in x_0} U_1(x_1, \mu)\mu_0^{*i}(\mu).$$

Then both $U_0^i(x_0)$ and $U_0^{*i}(x_0)$ represent $\succeq$ and have the same set of ex-post preferences, that is, $\mathbb{S}_1^i = \mathbb{S}_1(U_0^{*i})$ for $i = 1, 2$. Therefore, the proof completes if $\mathbb{S}_1(U_0^{*1}) = \mathbb{S}_1(U_0^{*2})$.

To show the above claim, we prepare the next lemma:

Lemma C.7. *Take any $\mu_0^i \in \Delta(\Delta(S_2^*))$, $i = 1, 2$. If*

$$U_0^i(x_0) \equiv \int_{\Delta(S_2^*)} \max_{x_1 \in x_0} U_1(x_1, \mu)\,\mathrm{d}\mu_0^i(\mu),\ \ i = 1, 2, \tag{21}$$

where

$$U_1(x_1, \mu) \equiv \int_{S_2^*} \max_{l \in x} u(l)\,\mathrm{d}\mu(u),$$

represent the identical preference on $\mathcal{D}$, then $\mu_0^1 = \mu_0^2$.

Proof. For all $\mu \in \Delta(S_2^*)$ and $x_0 \in \mathcal{D}$, let

$$\sigma_{x_0}(\mu) \equiv \max_{x_1 \in x_0} U_1(x_1, \mu).$$

Then

$$U_0^i(x_0) = \int \sigma_{x_0}(\mu)\,\mathrm{d}\mu_0^i(\mu),\ \ i = 1, 2. \tag{22}$$

For all $l \in \Delta(Z)$,

$$U_0^i(\{\{l\}\}) = \int\int u(l)\,\mathrm{d}\mu(u)\,\mathrm{d}\mu_0^i(\mu) = \bar{u}^i(l),$$

where $\bar{u}^i$ is the mean vNM-index with respect to μ_0^i. Precisely, for all $j = 1, \cdots, m$,

$$\bar{u}_j \equiv \int\int u_j\,\mathrm{d}\mu(u)\,\mathrm{d}\mu_0^i(\mu).$$

By definition, $\bar{u}^i \in S_2^*$. Since $\bar{u}^1$ and $\bar{u}^2$ represent the identical preference over $\Delta(Z)$, $\bar{u}^1 = \bar{u}^2$.

Since U_0^1 and U_0^2 are mixture linear functions over $\mathcal{D}$ representing the same preference, there exist $\alpha > 0$ and $\beta \in \mathbb{R}$ such that $U_0^2 = \alpha U_0^1 + \beta$. For any lottery l,

$$\begin{aligned} U_0^2(\{\{l\}\}) &= \alpha U_0^1(\{\{l\}\}) + \beta \\ \bar{u}^2(l) &= \alpha \bar{u}^1(l) + \beta. \end{aligned}$$

We must have $\alpha = 1$ and $\beta = 0$, and hence $U_0^1 = U_0^2$. From (22), for all x_0,

$$\int \sigma_{x_0}(\mu)\,\mathrm{d}\mu_0^1(\mu) = \int \sigma_{x_0}(\mu)\,\mathrm{d}\mu_0^2(\mu). \tag{23}$$

Take any $x_0, y_0 \in \mathcal{D}$ and $\alpha, \beta \geq 0$. Equation (23) holds even when σ_{x_0} is replaced with $\alpha\sigma_{x_0} - \beta\sigma_{y_0}$. From Lemma C.6, the set of all such functions is a dense subset of the set of all real-valued continuous functions on $\Delta(S_2^*)$. Hence equation (23) still holds even if σ_{x_0} is replaced with any real-valued continuous function. The Riesz representation theorem (Dunford and Schwartz [2, Theorem 3, p. 265]) implies $\mu_0^1 = \mu_0^2$. □

Since U_0^i and U_0^{*i} represent the same preference $\succeq$ for $i = 1, 2$, so do U_0^{*1} and U_0^{*2}. Since U_0^{*i} has the form of (21), it follows from Lemma C.7 that $\mu_0^{*1} = \mu_0^{*2}$. Especially $\mathrm{supp}(\mu_0^{*1}) = \mathrm{supp}(\mu_0^{*2})$. Since each $\mu \in \Delta(S_2^*)$ induces preference on $\mathcal{K}(\Delta(Z))$ by $U_1(\cdot, \mu)$, we have $\mathbb{S}_1(U_0^{*1}) = \mathbb{S}_1(U_0^{*2})$. This completes the proof.

References

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

[2] N. Dunford and J. T. Schwartz. *Linear Operators, Part I : Gereral Theory*. Wiley, 1958.

[3] F. Gul and W. Pesendorfer. Temptation and self-control. *Econometrica*, 69:1403–1435, 2001.

[4] A. Kraus and J. S. Sagi. Inter-temporal preference for flexibility and risky choice. working paper, September, 2003.

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

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

[7] S. Modica. A note on preference for flexibility. *Decisions in Economics and Finance*, 25:47–63, 2002.

[8] W. Rudin. *Real and Complex Analysis*. McGRAW-HILL, third edition, 1987.

[9] A. Rustichini. Preference for flexibility in infinite horizon problems. *Economic Theory*, 20:677–702, 2002.

[10] L. J. Savage. *The Foundations of Statistics*. Dover, New York, 1954.

[11] H. H. Schaefer. *Topological Vector Spaces*. Number 3 in GTM. Springer, second edition, 1999.

[12] N. Takeoka. Subjective probability over a subjective decision tree. University of Rochester, working paper, May 2005.