Subjective Probability over a Subjective Decision Tree *

Norio TAKEOKA

Department of Economics, University of Rochester, Rochester, NY 14627, USA

May 10, 2005

Abstract

Since Savage's seminal work, a state space has been assumed as a primitive, which requires the analyst to know all the uncertainties a decision maker perceives. Dekel, Lipman and Rustichini (2001) derive a unique subjective state space from preference on a suitable domain.

In a dynamic setting, a state space S and a filtration $\{\mathcal{F}_t\}_{t=0}^T$ over S have been taken as primitives. We derive the triple $(S, \{\mathcal{F}_t\}_{t=0}^T, P)$ from preference, where S is a subjective state space, $\{\mathcal{F}_t\}_{t=0}^T$ is a subjective filtration over S, and P is a subjective probability over S. We also show uniqueness of the representation.

Keywords: preference for flexibility, subjective state space, subjective probability, subjective decision tree.

JEL classification: D81

1 Introduction

1.1 Motivation and Main Results

A state space has been used as the standard tool for modeling uncertainties since Savage [17] and Anscombe and Aumann [1]. In their models, a state space is taken as a primitive. This

*E-mail: ntak@troi.cc.rochester.edu. I would like to thank Larry Epstein for his illuminating guidance and invaluable suggestion and discussion. This paper does not exist without him. I'm thankful also to Luca Anderlini, Alfredo Di Tillio, Takashi Hayashi, Kazuya Hyogo, Jing Li, Massimo Marinacci, Eiichi Miyagawa, Yutaka Nakamura, Jawwad Noor, Alvin Roth, and the audiences at the 2004 Decentralization Conference (Keio, Japan), the 2004 Annual Meeting of the Japanese Economic Association (Meiji Gakuin, Japan), Fall 2004 Midwest Economic Theory and Trade Conference, University of Rochester, Kyoto University, Osaka University, and Kobe University for helpful comments. All remaining errors are mine.

modeling implicitly requires the analyst (observer) to know all the uncertainties a decision maker (DM) perceives. However, the DM may have in mind some "subjective" states other than the objective states, and anticipate that those states are relevant for her decision. Hence, it is questionable whether the state space can be a primitive. Now we are led to ask whether a state space can be derived rather than assumed as a primitive. The derivation of a *subjective state space* is addressed by Kreps [9, 10] and Dekel, Lipman and Rustichini [2] (hereafter DLR).

In a dynamic setting, the standard tool for modeling uncertainties is a decision tree $(S, \{\mathcal{F}_t\}_{t=0}^T)$, where S is a state space and $\{\mathcal{F}_t\}_{t=0}^T$ is a filtration over S. It has been taken as a primitive, which requires the analyst to know not only all the uncertainties the DM perceives, but also how the DM expects those uncertainties to be resolved over time. By the same reason as outlined above, this modeling seems restrictive. We have to ask whether both S and $\{\mathcal{F}_t\}_{t=0}^T$ can be subjective. The derivation of a *subjective decision tree, or a subjective filtration over a subjective state space*, is the focus of this paper.

The following table summarizes our results and the relation to previous literature:

	Subjective		
	S	$\{\mathcal{F}_t\}_{t=0}^T$	P
Savage			✓
Kreps, DLR	✓		
This Paper	✓	✓	✓

In Savage [17], Anscombe and Aumann [1], and their dynamic counterparts, a state space S and a filtration $\{\mathcal{F}_t\}_{t=0}^T$ over S are assumed as primitives. This literature derives a probability measure P over S as a part of the representation. Kreps [9, 10] and DLR derive S without assuming any objective state space. Since their models are static, filtrations over S are not relevant. In this paper, the pair $(S, \{\mathcal{F}_t\}_{t=0}^T)$ is derived from preference on a suitable domain. Moreover, unlike DLR, we provide also a subjective probability measure P over S. Thus, the triple $(S, \{\mathcal{F}_t\}_{t=0}^T, P)$ is subjective in this paper.

1.2 Domain

To derive a subjective state space, DLR consider preference over opportunity sets (called "menus" henceforth) of lotteries. We consider preference over menus of menus of Anscombe-Aumann acts. Precisely, let Ω be a finite objective state space and $\Delta(Z)$ be the set of lotteries over an outcome space Z. Let $\mathcal{H}$ be the set of functions, $h : \Omega \to \Delta(Z)$, called Anscombe-Aumann acts. Thus, the objective state space Ω is assumed to be payoff-relevant in the sense that the DM's payoff is determined once one of those states is realized. [1] We consider preference $\succeq$ on the domain $\mathcal{K}(\mathcal{K}(\mathcal{H}))$, where $\mathcal{K}(\cdot)$ denotes the set of all non-empty

[1] The DM may have in mind some subjective states other than the objective states. We can address the issue of subjective states even when payoff-relevant objective states are taken as primitives.

compact subsets of "$\cdot$". Thus, there are two changes related to DLR: (1) menus of menus rather than menus; and (2) acts rather than lotteries.

We have 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 an act $h \in x_1$

Period 2^+: an objective state is realized and the DM receives the lottery prescribed by h

Notice that the above time line, except period 0, is not a part of the formal model. Especially, subjective signals are not assumed as primitives. However, if the DM has in mind the above timing of decisions and anticipates subjective signals to arrive gradually over time, preference in period 0 should reflect her perception of those subjective uncertainties. That is why $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ is a relevant domain for deriving a subjective decision tree.

1.3 An Example: Asset Choice Problem

What kind of behavior is consistent with the hypothesis that the DM anticipates subjective signals to arrive over time? A key is preference for flexibility. If the DM anticipates that some subjective information is coming later on, she would like to delay a decision until this information arrives.

Imagine a situation where a DM chooses, by the end of period 2, between the two kinds of assets,

$$h_1 = \left[\begin{array}{cc} \$100 & \omega_1 \\ 0 & \omega_2 \end{array} \right], \text{ and } h_2 = \left[\begin{array}{cc} 0 & \omega_1 \\ \$100 & \omega_2 \end{array} \right],$$

where ω_1 and ω_2 are objective states.

First consider the menu of menus $\{\{h_i\}\}$ for $i = 1, 2$. If $\{\{h_i\}\}$ is chosen in period 0, there is no choice afterwards. The DM commits herself right now to choose h_i. For example, this choice object corresponds to the action, such as participating in the futures market and making a long-term contract to buy the asset in the future.

Suppose that the DM is indifferent between h_1 and h_2 if she has to commit herself to one of them. That is,

$$\{\{h_1\}\} \sim \{\{h_2\}\}.$$

This ranking presumably reveals that the DM anticipates ω_1 and ω_2 to be equally likely.

Even though $\{\{h_1\}\}$ and $\{\{h_2\}\}$ are indifferent, the following rankings seem appealing in terms of flexibility:

$$\{\{h_1, h_2\}\} \succeq \{\{h_1\}, \{h_2\}\} \succeq \{\{h_1\}\}. \tag{1}$$

If the DM chooses $\{\{h_1\}, \{h_2\}\}$, she can delay a decision until period 1. If she chooses $\{\{h_1, h_2\}\}$, she can delay a decision until period 2.

Our hypothesis for explaining ranking (1) is as follows: the DM anticipates that subjective signals arrive both in period 1 and in period 2 and that they convey some information about the objective states. She will be able to update her initial belief over $\{\omega_1, \omega_2\}$ in response to those subjective signals. To obtain this new information, the DM would like to delay a decision.

What subjective decision tree can be derived from ranking (1)? There are four cases:

$$\begin{array}{rlclclcl}
\text{(i)} & \{\{h_1, h_2\}\} & \sim & \{\{h_1\}, \{h_2\}\} & \sim & \{\{h_1\}\} & \sim & \{\{h_2\}\}; \\
\text{(ii)} & \{\{h_1, h_2\}\} & \sim & \{\{h_1\}, \{h_2\}\} & \succ & \{\{h_1\}\} & \sim & \{\{h_2\}\}; \\
\text{(iii)} & \{\{h_1, h_2\}\} & \succ & \{\{h_1\}, \{h_2\}\} & \sim & \{\{h_1\}\} & \sim & \{\{h_2\}\}; \\
\text{(iv)} & \{\{h_1, h_2\}\} & \succ & \{\{h_1\}, \{h_2\}\} & \succ & \{\{h_1\}\} & \sim & \{\{h_2\}\}.
\end{array}$$

Ranking (i) says that the DM does not care when she commits herself to choose between h_1 and h_2. In other words, she does not desire flexibility. This ranking reveals that no subjective signals are expected to arrive and that the initial belief about the objective states does not get updated over time.

Ranking (ii) says that the DM strictly desires flexibility in period 1, while she does not in period 2. This ranking can be justified by the following story: the DM anticipates that at least two subjective signals will arrive in period 1. One signal suggests that ω_1 is more likely to happen, while the other signal suggests the opposite. If $\{\{h_1\}, \{h_2\}\}$ has been chosen, the DM can make a decision contingent upon this new information. Hence, $\{\{h_1\}, \{h_2\}\}$ is strictly preferred to $\{\{h_1\}\}$. However, since she does not expect another signal to arrive later on, she is willing to decide in period 1 between h_1 and h_2. Hence, $\{\{h_1, h_2\}\}$ and $\{\{h_1\}, \{h_2\}\}$ are indifferent.

The above reasoning suggests that the DM has a subjective decision tree such as Figure 1.

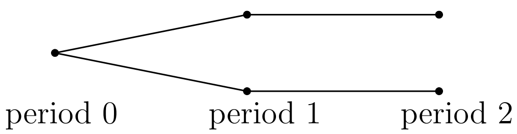

Figure 1: subjective decision tree deduced from ranking (ii)

The opposite explanation works for ranking (iii). That is, the DM presumably anticipates no subjective signal in period 1, while she expects at least two subjective signals to arrive in period 2. That is why she does not desire flexibility in period 1, while she does in period 2. We can deduce a subjective decision tree such as Figure 2.

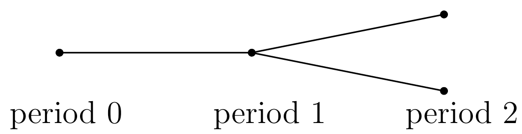

Figure 2: subjective decision tree deduced from ranking (iii)

Finally, ranking (iv) reveals that the DM anticipates at least two subjective signals to arrive in period 1 as well as in period 2. Hence, she strictly desires flexibility both in period 1 and in period 2. Presumably, the DM has a subjective decision tree such as Figure 3.

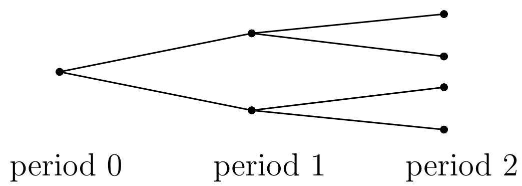

Figure 3: subjective decision tree deduced from ranking (iv)

What is the importance of taking into account menus of menus rather than menus? In other words, what is the main difference between $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ and $\mathcal{K}(\mathcal{H})$ for our purpose? Though $\mathcal{K}(\mathcal{H})$ is relevant for deriving subjective uncertainties, it is too small to distinguish the timing of resolution of those uncertainties.[2] For example, each of ranking (ii), (iii), and (iv) implies

$$\{\{h_1, h_2\}\} \succ \{\{h_1\}\} \sim \{\{h_2\}\},$$

which, in terms of elements of $\mathcal{K}(\mathcal{H})$, translates into the ranking

$$\{h_1, h_2\} \succ \{h_1\} \sim \{h_2\}.$$

Hence, the three rankings (ii), (iii) and (iv) cannot be discriminated on $\mathcal{K}(\mathcal{H})$.

1.4 Functional Form

We axiomatize preference on $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ having the following representation: there exist (i) a "full" state space $S_1 \times S_2 \times \Omega$, where S_1 and S_2 are sets of subjective signals, (ii) the filtration $\{\mathcal{F}_t\}_{t=0}^3$ over $S_1 \times S_2 \times \Omega$ generated by the product structure, (iii) a countably additive probability measure $\mathbb{P}$ on $S_1 \times S_2 \times \Omega$, and (iv) a non-constant mixture linear function $u : \Delta(Z) \to \mathbb{R}$ such that $U_0 : \mathcal{K}(\mathcal{K}(\mathcal{H})) \to \mathbb{R}$ represents preference, where

$$U_t(x_t) \equiv \mathbb{E}_{\mathbb{P}}\left[\max_{x_{t+1} \in x_t} U_{t+1}(x_{t+1}) \,\middle|\, \mathcal{F}_t\right], \ t = 0, 1, 2$$

[2] This difference is analogous to the difference between a static model for lotteries and Kreps and Porteus [11] or Epstein and Zin [5], where the DM may distinguish two compounded lotteries even though these lotteries induce the same distribution on the outcome space; that is, the timing of resolution of risk (objective uncertainties) matters.

with the convention $\max_{x_3 \in x_2} U_3(x_3) \equiv u(x_2)$ for all $x_2 \in \mathcal{H}$. For each x_t, $U_t(x_t) : S_1 \times S_2 \times \Omega \to \mathbb{R}$ is an $\mathcal{F}_t$-measurable function. Thus, this representation is characterized by two components: the filtered probability space $(S_1 \times S_2 \times \Omega, \{\mathcal{F}_t\}_{t=0}^3, \mathbb{P})$ and the mixture linear function u. Moreover, we show uniqueness of the representation.

The above representation is interpreted as follows: the DM behaves as if she has in mind the time line described in Section 1.2 and anticipates subjective signals to arrive gradually over time. The DM certainly knows her future risk preference, while she is not sure of future beliefs about Ω. Hence, subjective uncertainty concerns beliefs about Ω. In response to a subjective signal, the DM will update her initial belief over Ω.

1.5 Related Literature

Kreps [9, 10] provides an axiomatic foundation of a subjective state space. DLR show uniqueness of the representation by imposing a richer structure on the domain. Let Z be finite and $\mathcal{P}(\cdot)$ denote the set of all non-empty subsets of "$\cdot$". DLR consider $\mathcal{P}(\Delta(Z))$ as the domain. Though they have several different models, we focus on the additive representation with a non-negative measure, that is, the functional form $U : \mathcal{P}(\Delta(Z)) \to \mathbb{R}$ defined by

$$U(x) = \int_S \sup_{l \in x} u(l, s)\, \mathrm{d}\mu(s), \tag{2}$$

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. They show that the set of ex-post preferences induced from $\{u(\cdot, s)\}_{s \in S}$ is uniquely determined from preference. This set is called the *subjective state space*.

Rustichini [16] addresses a multi-period extension of DLR. Let C be the set of consumptions and C^∞ be the set of infinite consumption streams. His domain is $\mathcal{P}(C^\infty)$. He fails to derive a subjective decision tree because his model is essentially static in the sense that all subjective uncertainties are resolved in the next period. Modica [12] considers a menu of menus within the Kreps's framework. That is, his domain is $\mathcal{P}(\mathcal{P}(Z))$. This framework, however, cannot pin down the representation as in Kreps [9]. Kraus and Sagi [8] take a different approach to address intertemporal choice consistent with preference for flexibility. In their model, incompleteness of preference is essential.

We may be tempted to interpret μ in (2) as a belief of the DM about the subjective uncertainties. However, μ is not unique because of the state-dependence of the ex-post utility functions. Precisely, there may exist distinct components $(S, u(\cdot, s), \mu)$ and $(S', u'(\cdot, s'), \mu')$ representing the same preference but satisfying $\mu \neq \mu'$ even if $S = S'$. Thus, their model fails to derive a subjective probability over a subjective state space.

The special nature of the domain $\mathcal{H}$ helps us to derive meaningful probabilities. Unlike preference on $\Delta(Z)$, an SEU representation on $\mathcal{H}$ has two components: a risk preference and a belief over Ω. In our model, subjective uncertainties are related to beliefs over Ω, and do not affect the DM's risk preference. This state-independence of risk preference is the reason why we can pin down a subjective probability.

Domains consisting of menus with some objective states are not new. Epstein [4] introduces the domain $\mathcal{P}(\mathcal{H})$ and provides non-Bayesian updating models. In this setup, Ω is assumed as a "complete" description of the world in the sense that the DM's payoff is determined once one of those states is realized. Nevertheless, the issue of subjective uncertainty is still relevant because the DM may expect some subjective signals to arrive prior to the realization of the objective states.

As has been pointed out, when considering $\mathcal{P}(\mathcal{H})$, we have in mind the following timing of decisions: (1) the DM chooses a menu of acts; (2) after receiving a subjective signal, she chooses an act out of the menu; (3) an objective state is realized and she receives the outcome prescribed by the act. Hence, in this domain, subjective uncertainties are resolved first. Presumably the realized subjective state conveys some information about the objective states. Later on, one of the objective states is realized.

Nehring [13], Ghirardato [6] and Ozdenoren [14] are other literature considering menus with objective states. This literature addresses the issue of an "incomplete" state space. When the DM perceives Ω to be a coarse or incomplete description of the world, she should anticipate some subjective uncertainties to remain unresolved even after one of the objective states is realized. To capture the DM's coarse perception of the world, the above three authors take an "opportunity act" as a choice object. That is, their domains are sets of set-valued Savage acts or of set-valued Anscombe-Aumann acts.

The difference between those domains and $\mathcal{P}(\mathcal{H})$ is the timing of decisions. In the above three literature, (1) the DM chooses an opportunity act; (2) an objective state is realized and she receives the menu prescribed by the act; (3) after observing a subjective state, she chooses an outcome out of the menu. Hence, the order in which subjective and objective uncertainties resolve is reversed.

Finally, notice that, on the sub-domain $\mathcal{K}(\mathcal{K}(\Delta(Z))) \subset \mathcal{K}(\mathcal{K}(\mathcal{H}))$, our representation collapses to the following functional form without any subjective states,

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

which does not coincide with the multi-period counterpart of DLR's additive representation. Therefore, our representation is not a generalization of DLR. As shown in Takeoka [19], the generalization can be achieved by the similar argument to our main result if one of our axioms is dropped.

2 Model

2.1 Domain: Formal Definition

Let Ω be a finite objective state space with $\#\Omega = n$. Let Z be a compact metric outcome space. Let $\Delta(Z)$ be the set of all Borel probability measures over Z and $\mathcal{H}$ the set of all functions, $h : \Omega \to \Delta(Z)$, called Anscombe-Aumann acts (henceforth acts). Notice that $\Delta(Z)$ is a compact metric space under the weak convergence topology and $\mathcal{H}$ is also a compact metric space under the product topology.

Let $\mathcal{K}(\mathcal{H})$ be the set of all non-empty compact subsets of $\mathcal{H}$. Generic elements are denoted by x_1, y_1, $\cdots$, and interpreted as menus or opportunity sets of acts. Endow $\mathcal{K}(\mathcal{H})$ with the Hausdorff metric. Details are relegated to Appendix A.

Let $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ be the set of all non-empty compact subsets of $\mathcal{K}(\mathcal{H})$ with the Hausdorff metric. Generic elements are denoted by $x_0, y_0, \cdots$, and interpreted as menus of menus of acts. The assumptions about the timing of decisions are in Section 1.2.

Preference $\succeq$ is defined on $\mathcal{D} \equiv \mathcal{K}(\mathcal{K}(\mathcal{H}))$.

2.2 Axioms

The following five axioms on $\succeq$ are formally identical to those of DLR, but are imposed here on $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ rather than on $\mathcal{K}(\Delta(Z))$.

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

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 (Strong Nondegeneracy): There exist $l, l' \in \Delta(Z)$ such that $\{\{l\}\} \succ \{\{l'\}\}$.

Axiom 3 is stronger than that of DLR, which requires in our setting that there exist $x_0, x_0' \in \mathcal{D}$ such that $x_0 \succ x_0'$.

Define the mixture

$$\lambda x_1 + (1-\lambda)x_1' \equiv \{\lambda h + (1-\lambda)h' | h \in x_1, h' \in x_1'\}, \tag{3}$$

for any $x_1, x_1' \in \mathcal{K}(\mathcal{H})$ 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 as in DLR by two separate steps. Take any $x_1, z_1 \in \mathcal{K}(\mathcal{H})$ and $\lambda \in [0,1]$. As the first step, consider the lottery $\lambda \circ x_1 + (1-\lambda) \circ z_1$, which assigns x_1 with probability λ and z_1 with probability $(1-\lambda)$. vNM independence axiom implies that, for any $\lambda \in (0,1]$, if x_1 is preferred to y_1, then $\lambda \circ x_1 + (1-\lambda) \circ z_1$ is preferred to $\lambda \circ y_1 + (1-\lambda) \circ z_1$. As the second step, consider how the DM ranks $\lambda \circ x_1 + (1-\lambda) \circ z_1$ and $\lambda x_1 + (1-\lambda) z_1$. The difference between these two objects is when the randomization $(\lambda, 1-\lambda)$ gets realized. For the former, this randomization gets realized first, and, subsequently, choice out of the realized menu takes place, while this order is reversed for the latter. As long as the DM surely believes that her future preference on $\mathcal{H}$ satisfies mixture linearity, she does not care about this difference. Therefore, it follows from the above two steps that preference on $\mathcal{K}(\mathcal{H})$ will satisfy Independence. By applying the same argument twice, Independence on $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ can be justified whenever the DM surely anticipates her future preference on $\mathcal{H}$ to satisfy mixture linearity.

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

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

If the DM chooses a bigger menu of menus, she can retain more flexibility until period 1. Hence, this axiom is consistent with preference for flexibility.

The axioms proposed from now on have no counterparts in DLR. The next axiom is relevant only for the multi-period setup and says that the DM always weakly prefers to delay a decision.

Axiom 6 (Aversion to Commitment): For all $x_0' \in \mathcal{D}$ and for all finite $x_0 \in \mathcal{D}$,

$x_0' \cup \{\cup_{x_1 \in x_0} x_1\} \succeq x_0' \cup x_0$.

For example, let $x_0' = \{\{h_0\}\}$ and $x_0 = \{\{h_1\}, \{h_2\}\}$. Then,

$$\begin{aligned} x_0' \cup x_0 &= \{\{h_0\}, \{h_1\}, \{h_2\}\}, \text{ and,} && (4) \\ x_0' \cup \{\cup_{x_1 \in x_0} x_1\} &= \{\{h_0\}, \{h_1, h_2\}\}. && (5) \end{aligned}$$

If the DM chooses (5), she can always choose a weakly bigger menu in period 1 in contrast with (4). That is, (5) leaves more options open until period 2 than does (4). Hence, it is appealing in terms of flexibility that (5) is weakly preferred to (4).

For any $h \in \mathcal{H}$, define

$$O_1(h) \equiv \{h' \in \mathcal{H} \mid \{\{h(\omega)\}\} \succeq \{\{h'(\omega)\}\} \text{ for all } \omega\}.$$

That is, $O_1(h)$ is the set of all acts dominated by h state by state in terms of commitment ranking over lotteries. This dominance notion is applicable to any menu. For each $x_1 \in \mathcal{K}(\mathcal{H})$, let

$$O_1(x_1) \equiv \cup_{h \in x_1} O_1(h). \tag{6}$$

That is, $O_1(x_1)$ is the set of all acts dominated by some act in x_1. Lemma B.4 (i) in Appendix B.2 ensures that $O_1(x_1)$ is a well-defined menu, that is, $O_1(x_1) \in \mathcal{K}(\mathcal{H})$. Notice also that $x_1 \subset O_1(x_1)$. Finally, this dominance notion is extended to any menu of menus by taking the operation O_1 element by element. That is, for each $x_0 \in \mathcal{D}$,

$$O(x_0) \equiv \{O_1(x_1) \mid x_1 \in x_0\}.$$

Lemma B.4 (iii) ensures that $O(x_0) \in \mathcal{D}$.

Axiom 7 (Risk Preference Certainty): For all $x_0 \in \mathcal{D}$, $x_0 \sim O(x_0)$.

This axiom can be justified as follows: suppose that the DM surely knows her future risk preference, that is, the ranking over lotteries. Then commitment ranking $\{\{l\}\} \succeq \{\{l'\}\}$, which is ex ante evaluation of lotteries, coincides with the risk preference in period 2. Since any act in $O_1(h) \setminus \{h\}$ is dominated by h in terms of this risk preference, the DM should be indifferent between $\{h\}$ and $O_1(h)$. Similarly, she should be indifferent between x_1 and $O_1(x_1)$ because the additional part $O_1(x_1) \setminus x_1$ is surely valueless. Consequently, x_0 and $O(x_0)$ should be indifferent.

3 Representations

3.1 Second-Order Additive SEU Representation

Consider the functional form $U_0 : \mathcal{K}(\mathcal{K}(\mathcal{H})) \to \mathbb{R}$ defined by

$$U_0(x_0) \equiv \int_{S_1} \max_{x_1 \in x_0} U_1(x_1, s_1)\, \mathrm{d}\mu_0(s_1), \tag{7}$$

where

$$\begin{aligned} U_1(x_1, s_1) &\equiv \int_{S_2} \max_{h \in x_1} U_2(h, s_1, s_2)\, \mathrm{d}\mu_1(s_2|s_1), \\ U_2(h, s_1, s_2) &\equiv \sum_{\omega \in \Omega} u(h(\omega))\, \mu_2(\omega|s_1, s_2), \end{aligned}$$

S_1 and S_2 are topological state spaces, μ_0 is a countably additive probability measure over S_1, $\mu_1 : S_1 \to \Delta(S_2)$ and $\mu_2 : S_1 \times S_2 \to \Delta(\Omega)$ are conditional probability systems, and $u : \Delta(Z) \to \mathbb{R}$ is a non-constant mixture linear function. [3]

An interpretation of this functional form is as follows: The DM behaves as if she has in mind the time line described in Section 1.2 and anticipates subjective signals to arrive gradually over time. She will update her initial belief over Ω in response to those subjective signals. She is certain about risk preference u, but not sure of future beliefs about Ω. Thus subjective uncertainty concerns beliefs about Ω.

Definition 3.1. *Preference $\succeq$ on $\mathcal{D}$ admits a second-order additive SEU representation if there exists functional form (7) with components $(\{S_t\}_{t=1}^2, \{\mu_t\}_{t=0}^2, u)$ representing $\succeq$.*

[3] Suppose that X and Y are topological spaces. The σ-algebra over X, denoted by $\mathcal{B}(X)$, is assumed to be the Borel σ-algebra. The set of countably additive probability measures on X, denoted by $\Delta(X)$, is endowed with the weak convergence topology. A function $f : X \to \Delta(Y)$ is said to be a conditional probability system if f is measurable with respect to $\mathcal{B}(X)$ and $\mathcal{B}(\Delta(Y))$.

Any second-order additive SEU representation with $(\{S_t\}_{t=1}^2, \{\mu_t\}_{t=0}^2, u)$ determines a filtered probability space $(S, \{\mathcal{F}_t\}_{t=0}^2, P)$ as we now describe. First, a countably additive probability measure P over $S_1 \times S_2$ is defined as the unique measure satisfying

$$P(E_1 \times E_2) \equiv \int_{E_1} \mu_1(E_2|s_1)\, \mathrm{d}\mu_0(s_1),$$

for all $E_1 \in \mathcal{B}(S_1)$ and $E_2 \in \mathcal{B}(S_2)$. The set S is defined as the support of P. Finally, the product structure of $S_1 \times S_2$ induces the filtration $\{\mathcal{F}_t\}_{t=0}^2$ over S as follows:

$$\begin{aligned}
\mathcal{F}_0 &\equiv \{S\}, \\
\mathcal{F}_1 &\equiv \{S_{s_1} \,|\, s_1 \in \mathrm{supp}(\mu_0)\}, \\
\mathcal{F}_2 &\equiv \{\{(s_1, s_2)\} \,|\, (s_1, s_2) \in S\},
\end{aligned}$$

where $S_{s_1} \equiv \{(s_1, s_2) \,|\, (s_1, s_2) \in S\}$.[4] Since the DM has no information in period 0, $\mathcal{F}_0$ is the null partition. She expects to receive a subjective signal s_1 in period 1. The remaining subjective uncertainties conditional on these signals are captured by $\mathcal{F}_1$. Since all subjective uncertainties are resolved in period 2, $\mathcal{F}_2$ is the discrete partition. Notice that the objective uncertainties, that is, Ω, still remain unresolved in period 2.

A second-order additive SEU representation can be rewritten as in Section 1.4. Let $S \times \Omega$ be the "full" state space. Since the conditional probability system μ_2 is regarded as a measurable function from S into $\Delta(\Omega)$, the pair (P, μ_2) determines the probability measure $\mathbb{P}$ over $S \times \Omega$. The filtration $\{\mathcal{F}_t^*\}_{t=0}^3$ over $S \times \Omega$ can be defined by $\mathcal{F}_t^* \equiv \{E \times \Omega | E \in \mathcal{F}_t\}$ for $t = 0, 1, 2$, and by the discrete partition $\mathcal{F}_3^*$ over $S \times \Omega$. Then,

$$U_t(x_t) = \mathbb{E}_{\mathbb{P}} \left[\max_{x_{t+1} \in x_t} U_{t+1}(x_{t+1}) \,\middle|\, \mathcal{F}_t^* \right], \ t = 0, 1, 2$$

with the convention $\max_{x_3 \in x_2} U_3(x_3) \equiv u(x_2)$ for all $x_2 \in \mathcal{H}$. For each x_t, $U_t(x_t) : S \times \Omega \to \mathbb{R}$ is an $\mathcal{F}_t^*$-measurable function.

3.2 Subjective Decision Tree and Canonical Representation

Let $(\{S_t\}_{t=1}^2, \{\mu_t\}_{t=0}^2, u)$ be a second-order additive SEU representation and $(S, \{\mathcal{F}_t\}_{t=0}^2, P)$ be the derived filtered probability space. Notice that a state $(s_1, s_2) \in S$ itself should not matter for the DM. She only cares about the information associated with the state, that is, the belief $\mu_2(s_1, s_2) \in \Delta(\Omega)$. Hence, the set of all $\mu_2(s_1, s_2)$ as (s_1, s_2) varies over S is effectively identified with the subjective state space of the DM. Similarly, $\{\mathcal{F}_t\}_{t=0}^2$ and P are relevant for the DM only because they, together with $\mu_2 : S \to \Delta(\Omega)$, induce conditional distributions over $\Delta(\Omega)$.

More precisely, the filtered probability space $(S, \{\mathcal{F}_t\}_{t=0}^2, P)$ derived from a representation admits a reduced form as we now describe. Let $P \circ \mu_2^{-1} \in \Delta(\Delta(\Omega))$ denote the

[4] For any countably additive measure ν, $\mathrm{supp}(\nu)$ denotes its support.

distribution over $\Delta(\Omega)$ induced by the mapping $\mu_2 : (S, P) \to \Delta(\Omega)$. This distribution is regarded as the initial prior over $\Delta(\Omega)$. We call

$$\mathcal{S} \equiv \text{supp}(P \circ \mu_2^{-1}) \subset \Delta(\Omega)$$

the subjective state space. After receiving a signal s_1 in period 1, the DM updates P by Bayes' Rule, that is, $P(\cdot|s_1) = \mu_1(\cdot|s_1) \in \Delta(S_2)$. This updated belief and the mapping μ_2 induce the conditional distribution over $\mathcal{S} \subset \Delta(\Omega)$, which is denoted by $\mu_1(\cdot|s_1) \circ \mu_2^{-1}$. In period 0, the DM expects s_1 according to μ_0, which is the marginal distribution of P on S_1. Let $\mu_0^* \in \Delta(\Delta(\mathcal{S}))$ be the distribution of μ_0 induced by the mapping $s_1 \mapsto \mu_1(\cdot|s_1) \circ \mu_2^{-1}$. This second-order belief, μ_0^*, is interpreted as a "probability tree" over $\mathcal{S}$. The pair $(\mathcal{S}, \mu_0^*)$ summarizes all the relevant information for the DM, and is called *the subjective decision tree.*

The above argument suggests that the information in period 0 can be summarized by a probability measure $\mu_0 \in \Delta(\Delta(\Delta(\Omega)))$. We are led to the following definition:

Definition 3.2. *A functional $U_0 : \mathcal{D} \to \mathbb{R}$ is canonical if there exist $\mu_0 \in \Delta(\Delta(\Delta(\Omega)))$ and a non-constant mixture linear function $u : \Delta(Z) \to \mathbb{R}$ such that*

$$U_0(x_0) = \int_{\Delta(\Delta(\Omega))} \max_{x_1 \in x_0} U_1(x_1, \mu) \, \mathrm{d}\mu_0(\mu), \tag{8}$$

where

$$\begin{aligned} U_1(x_1, \mu) &= \int_{\Delta(\Omega)} \max_{h \in x_1} U_2(h, p) \, \mathrm{d}\mu(p), \ \ for\ \mu \in \Delta(\Delta(\Omega)), \\ U_2(h, p) &= \sum_{\omega} u(h(\omega)) p(\omega), \ \ for\ p \in \Delta(\Omega). \end{aligned}$$

The next proposition ensures that we can pay attention to canonical representations without loss of generality. A proof can be found in Appendix B.1.

Proposition 3.1.

(i) *Any canonical representation (μ_0, u) is a second-order additive SEU representation.*

(ii) *Let $(\mathcal{S}, \mu_0^*)$ be the subjective decision tree derived from a second-order additive SEU representation $(\{S_t\}_{t=1}^2, \{\mu_t\}_{t=0}^2, u)$. Then, the canonical form (8) with components (μ_0^*, u) represents the identical preference.*

Now we are ready to state the main theorem. See Appendix B.2 for a proof.

Theorem 3.1. *The following statements are equivalent:*

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

(b) *Preference $\succeq$ on $\mathcal{D}$ admits a canonical representation (μ_0, u).*

One might expect that DLR axioms, that is, Axiom 1-5, on $\mathcal{D}$ imply a representation of the form of DLR's additive representation. That is,

$$U_0(x_0) = \int_{S_1} \max_{x_1 \in x_0} U_1(x_1, s_1)\, \mathrm{d}\mu_0(s_1), \tag{9}$$

where S_1 is a state space, μ_0 is a non-negative measure over S_1, and $U_1(\cdot, s_1) : \mathcal{K}(\mathcal{H}) \to \mathbb{R}$ is a state-dependent utility function over menus. In order to obtain (9), we have to deal with functionals $U_1(\cdot, s_1)$ over the infinite dimensional space $\mathcal{K}(\mathcal{H})$, while DLR rely heavily on the fact that their menus are subsets of the finite dimensional space $\Delta(Z)$.[5] Hence, even the "static" result (9) is not straightforward.

Moreover, to go from (9) to a second-order additive SEU representation, $U_1(\cdot, s_1)$ has to be rewritten as

$$U_1(x_1, s_1) = \int_{S_2} \max_{h \in x_1} U_2(h, s_1, s_2)\, \mathrm{d}\mu_1(s_2|s_1). \tag{10}$$

To ensure this, we use Aversion to Commitment. From the argument of DLR, $U_1(\cdot, s_1)$ can be taken to be mixture linear. Hence, we can hope that the ex-post preference induced by $U_1(\cdot, s_1)$ satisfies Order, Continuity and Independence. However, Monotonicity may not be inherited. That is, DLR axioms do not ensure that $U_1(y_1, s_1) \leq U_1(x_1, s_1)$ whenever $y_1 \subset x_1$ for all $s_1 \in S_1$. Moreover, we cannot directly impose Monotonicity on each ex-post preference because $U_1(\cdot, s_1)$ is a part of the representation. Aversion to Commitment indirectly ensures Monotonicity of the ex-post preferences.

Finally, in a second-order additive SEU representation, each $U_2(\cdot, s_1, s_2) : \mathcal{H} \to \mathbb{R}$ differs only in the beliefs over Ω, not in the risk preference. To reduce the differences in the utility functions over acts to the differences in beliefs, we use Risk Preference Certainty. As shown in the next section, this axiom has the added benefit of pinning down a unique subjective probability.

4 Uniqueness

In this section, we discuss uniqueness of representations. We show first that preference on $\mathcal{D}$ admits a unique canonical representation. A proof is relegated to Appendix B.3.

Theorem 4.1. *If two canonical forms (μ_0^i, u^i), $i = 1, 2$, represent the same preference $\succeq$ on $\mathcal{D}$, then:*

[5] When Z is finite, $\Delta(Z)$ can be identified with a compact convex subset of a finite dimensional Euclidean space.

(i) *u^1 and u^2 are cardinally equivalent; and*

(ii) $\mu_0^1 = \mu_0^2$.

Unlike DLR, Theorem 4.1 pins down a unique probability measure. As mentioned in Section 1.5, this is made possible by the state-independence of risk preference.

Next we consider uniqueness of a second-order additive SEU representation. Take two components, $(\{S_t^i\}_{t=1}^2, \{\mu_t^i\}_{t=0}^2, u^i)$, $i = 1, 2$, representing the same preference. To show uniqueness, we have to compare these two representations somehow. However, S_1^1 and S_1^2 need not coincide in the set-theoretic sense because they include subjective signals, which can be anything. Consequently, μ_0^1 and μ_0^2 need not be comparable. To avoid such trivial non-uniqueness, we pay attention to the subjective decision trees derived from those representations. As mentioned in Section 3.2, all the relevant information described by a second-order additive SEU representation can be summarized by its subjective decision tree. The following theorem shows that preference on $\mathcal{D}$ uniquely determines the subjective decision tree:

Theorem 4.2. *If two second-order additive SEU representations, $(\{S_t^i\}_{t=1}^2, \{\mu_t^i\}_{t=0}^2, u^i)$, $i = 1, 2$, represent the same preference $\succeq$ on $\mathcal{D}$, then:*

(i) *u^1 and u^2 are cardinally equivalent; and*

(ii) $(\mathcal{S}^1, \mu_0^{1*}) = (\mathcal{S}^2, \mu_0^{2*})$.

This result follows from Proposition 3.1 (ii) and Theorem 4.1. Details can be found in Appendix B.4.

Finally, we clarify the connection between a canonical representation and its subjective decision tree. As shown in Proposition 3.1 (i), (μ_0, u) can be regarded as a second-order additive SEU representation. Hence, the subjective decision tree $(\mathcal{S}, \mu_0^*)$ can be derived from the representation as in Section 3.2. The following is an immediate consequence of Proposition 3.1 (ii) and Theorem 4.1:

Corollary 4.1. *Let $(\mathcal{S}, \mu_0^*)$ be the subjective decision tree derived from a canonical representation (μ_0, u). Then $\mu_0^* = \mu_0$.*

5 Special Cases

Some readers might think that how many times the DM expects subjective signals is structurally fixed and equal to two by considering preference on $\mathcal{K}(\mathcal{K}(\mathcal{H}))$. This is not necessarily the case because, as illustrated in Section 1.3, the DM may expect subjective signals to arrive just once. This section is devoted to explaining that the number of times of subjective signals depends on preference, not necessarily on the structure of the domain.

Take a canonical representation (μ_0, u). There exist two cases where the DM anticipates subjective signals to arrive just once. At one case, the DM expects no subjective signal in period 1. This case can be captured by $\#\mathrm{supp}(\mu_0) = 1$. At the other case, the DM expects all subjective uncertainties to be resolved by the end of period 1 and hence no subjective uncertainty to remain in period 2. In other words, subjective signals in period 1 are fully informative. This corresponds to the case where $\#\mathrm{supp}(\mu) = 1$ for any $\mu \in \mathrm{supp}(\mu_0)$. Figure 4 shows the subjective decision trees associated with these two cases.

(1) no signal in period 1 (2) no signal in period 2

Figure 4: special cases

As shown below, these special cases are obtained as corollaries of Theorem 3.1 if one of the axioms is replaced with a stronger axiom.

5.1 No Subjective Signal in Period 1

The next axiom states that the DM does not desire flexibility in period 1.

Axiom 5′ (Strategic Rationality): For all $x_0', x_0 \in \mathcal{D}$, $x_0' \succeq x_0 \ \Rightarrow \ x_0' \sim x_0' \cup x_0$.

Kreps [9] shows that $\succeq$ on $\mathcal{P}(Z)$ satisfies Order and Strategic Rationality if and only if there exists a utility function $u : Z \to \mathbb{R}$ such that $U : \mathcal{P}(Z) \to \mathbb{R}$, defined by

$$U(x) \equiv \sup_{z \in x} u(z),$$

represents $\succeq$. That is, the DM surely knows future preference over Z. In other words, she anticipates no subjective signal regarding future preference. Analogously, Strategic Rationality in our model implies that the DM expects no subjective signal in period 1.

Notice that Order and Strategic Rationality imply Monotonicity. Indeed, take any x_0, x_0' with $x_0 \subset x_0'$. Order implies $x_0 \succeq x_0'$ or $x_0' \succeq x_0$. If $x_0 \succeq x_0'$, Strategic Rationality implies $x_0 \sim x_0 \cup x_0' = x_0'$, and hence $x_0' \succeq x_0$.

If Monotonicity is replaced with Strategic Rationality in Theorem 3.1, the following is obtained as a corollary:

Corollary 5.1. *The following statements are equivalent:*

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

(b) *Preference $\succeq$ on $\mathcal{D}$ admits a canonical representation (μ_0, u) with $\#\mathrm{supp}(\mu_0) = 1$.*

See Appendix B.5 for a proof.

5.2 No Subjective Signal in Period 2

The next axiom states that it is not valuable to delay a decision until period 2.

Axiom 6′ (Neutrality to Commitment): For all $x'_0 \in \mathcal{D}$ and for all finite $x_0 \in \mathcal{D}$,

$$x'_0 \cup \{\cup_{x_1 \in x_0} x_1\} \sim x'_0 \cup x_0.$$

This axiom reveals that the DM expects all subjective uncertainties to be resolved by the end of period 1.

Clearly, Neutrality to Commitment implies Aversion to Commitment. If Aversion to Commitment is replaced with Neutrality to Commitment in Theorem 3.1, the following is obtained as a corollary:

Corollary 5.2. *The following statements are equivalent:*

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

(b) *Preference $\succeq$ on $\mathcal{D}$ admits a canonical representation (μ_0, u) such that $\#\mathrm{supp}(\mu) = 1$ for all $\mu \in \mathrm{supp}(\mu_0)$.*

A proof can be found in Appendix B.6.

6 Interpersonal Comparisons

In this section, we compare preferences on $\mathcal{D}$ having canonical representations (μ_0, u) in terms of preference for flexibility, and provide behavioral definitions capturing distinct attitudes toward resolution of subjective uncertainty.

Consider two agents, denoted by $i = 1, 2$. Agent i has preference $\succeq^i$ on $\mathcal{D}$ admitting a canonical representation (μ_0^i, u^i). Let $(\mathcal{S}^i, \mu_0^i)$ be agent i's subjective decision tree and $S_1^i \equiv \mathrm{supp}(\mu_0^i)$ be the set of agent i's subjective signals in period 1. Throughout this section, we assume the following conditions:

Finite Support (FS): μ_0^i and each $\mu^i \in \mathrm{supp}(\mu_0^i)$ have finite supports.

Identical Risk Preference (IR): $\succeq^1$ and $\succeq^2$ are identical on $\Delta(Z)$, that is, for any $l, l' \in \Delta(Z)$, $\{\{l\}\} \succ^1 \{\{l'\}\}$ if and only if $\{\{l\}\} \succ^2 \{\{l'\}\}$.

Under FS, S_1^i is finite. Moreover, $\mathcal{S}^i$ can be written as $\cup_{\mu \in S_1^i} \mathrm{supp}(\mu)$.[6] Hence, $\mathcal{S}^i$ is also finite. Under IR, without loss of generality, we can assume $u^1 = u^2 = u$.

Definition 6.1.

[6] This follows from Corollary 4.1 and Lemma B.8.

(i) $\succeq^2$ *desires more flexibility than* $\succeq^1$ *if, for any* $x_0, x'_0 \in \mathcal{D}$ *with* $x_0 \subset x'_0$,
$x'_0 \succ^1 x_0 \Rightarrow x'_0 \succ^2 x_0$.

(ii) $\succeq^2$ *desires more period 2-flexibility than* $\succeq^1$ *if, for any* $x_1, x'_1 \in \mathcal{K}(\mathcal{H})$ *with* $x_1 \subset x'_1$,
$\{x'_1\} \succ^1 \{x_1\} \Rightarrow \{x'_1\} \succ^2 \{x_1\}$.

(iii) $\succeq^2$ *expects later resolution of subjective uncertainty than* $\succeq^1$ *if:*
(1) for any $x_1, x'_1 \in \mathcal{K}(\mathcal{H})$ *with* $x_1 \subset x'_1$, $\{x'_1\} \succ^1 \{x_1\} \Leftrightarrow \{x'_1\} \succ^2 \{x_1\}$*; and*
(2) for any finite $x_0 \in \mathcal{D}$ *and* $x_1 \equiv \cup_{x'_1 \in x_0} x'_1$, $\{x_1\} \succ^1 x_0 \Rightarrow \{x_1\} \succ^2 x_0$.

Part (i) says that, whenever agent 1 strictly prefers a bigger menu of menus, so does agent 2. As an example, consider the following rankings:

$$\{\{h_1, h_2\}\} \succ^i \{\{h_1\}, \{h_2\}\} \succ^i \{\{h_1\}\} \succeq^i \{\{h_2\}\}.$$

Monotonicity and Aversion to Commitment imply $\{\{h_1, h_2\}, \{h_1\}, \{h_2\}\} \sim^i \{\{h_1, h_2\}\}$. Thus, we have

$$\{\{h_1, h_2\}, \{h_1\}, \{h_2\}\} \succ^i \{\{h_1\}, \{h_2\}\} \succ^i \{\{h_1\}\}. \tag{11}$$

Notice that both strict rankings in (11) are consistent with Monotonicity. The second ranking in (11) seems concerned with preference for flexibility in period 1, while the first ranking in (11) says that the DM prefers to delay a decision until period 2. Presumably, she anticipates some subjective uncertainties to remain unresolved even after seeing a signal in period 1. The first ranking would reflect preference for flexibility in period 2, conditional on signals in period 1. Therefore, Monotonicity captures two kinds of preference for flexibility. Part (i) says that agent 2 is more sensitive to flexibility in period 1 as well as to flexibility in period 2, given the information in period 1.

Part (ii) says that, whenever agent 1 prefers to commit herself to a bigger menu, so does agent 2. Once agent i commits herself to a menu x_1, she faces x_1 in period 2 because she has to choose x_1 in period 1 no matter what signal arrives. Hence, agent i is only concerned with the subjective uncertainties in period 2, but does not care about how those uncertainties are resolved over time. Part (ii) presumably implies that agent 2 perceives more subjective uncertainties than does agent 1.

Imagine two decision trees such that both trees have identical terminal nodes, but one filtration is finer than the other. Part (iii) is intended for such a comparison. Condition (1) says that agent 1 strictly prefers to commit herself to a bigger menu if and only if agent 2 does too. That is, both agents value flexibility in period 2 identically – presumably, they have identical subjective state spaces. Condition (2) says that, whenever agent 1 strictly prefers to delay a decision until period 2, so does agent 2. Presumably, this is because agent 2 anticipates more subjective uncertainties to remain unresolved even after seeing a signal in period 1. In other words, agent 2 expects later resolution of subjective uncertainty than does agent 1.

Notice that part (i) implies part (iii.2). Indeed, Monotonicity and Aversion to Commitment imply $\{x_1\} \sim^i \{x_1\} \cup x_0$.[7] Hence, (iii.2) can be rewritten as: $\{x_1\} \cup x_0 \succ^1 x_0 \Rightarrow \{x_1\} \cup x_0 \succ^2 x_0$, which is a special case of part (i). Since the ranking $\{x_1\} \cup x_0 \succ^i x_0$ corresponds to the first ranking in (11), (iii.2) is concerned with period 2-flexibility, conditional on signals in period 1.

Both of part (ii) and part (iii.2) capture preference for flexibility in period 2. The difference is that the ranking $\{x_1\} \succ^i x_0$ reflects the agent's expectation over the subjective uncertainties conditional on signals in period 1, while the commitment ranking is ex ante, prior to knowing signals in period 1.

The above behavioral definitions admit the following characterizations. Proofs can be found in Appendix B.7.

Theorem 6.1. *Assume FS and IR. Then:*

(i) $\succeq^2$ *desires more flexibility than* $\succeq^1$ *if and only if* $S_1^1 \subset S_1^2$.

(ii) $\succeq^2$ *desires more period 2-flexibility than* $\succeq^1$ *if and only if* $\mathcal{S}^1 \subset \mathcal{S}^2$.

(iii) $\succeq^2$ *expects later resolution of subjective uncertainty than* $\succeq^1$ *if and only if* $\mathcal{S}^1 = \mathcal{S}^2$, *and for any* $\mu^1 \in S_1^1$, *there exists* $\mu^2 \in S_1^2$ *such that* $\mathrm{supp}(\mu^1) \subset \mathrm{supp}(\mu^2)$.

Part (i) says that agent 2 desires more flexibility than agent 1 if and only if agent 2 expects more subjective signals to arrive in period 1 than does agent 1.

Part (ii) says that agent 2 desires more period 2-flexibility than agent 1 if and only if the subjective state space of agent 2 is bigger than that of agent 1. Two remarks are in order. Part (ii) is the counterpart of Theorem 2 (1) of DLR (p. 910).[8] Second, since $S_1^1 \subset S_1^2$ implies $\mathcal{S}^1 \subset \mathcal{S}^2$, part (i) and (ii) imply that $\succeq^2$ desires more period 2-flexibility than $\succeq^1$ whenever $\succeq^2$ desires more flexibility than $\succeq^1$.

The condition $\mathrm{supp}(\mu^1) \subset \mathrm{supp}(\mu^2)$ means that μ^1 assigns probability one to a smaller event than does μ^2. In that sense, μ^1 is more informative than μ^2. Part (iii) says that agent 2 expects later resolution of subjective uncertainty than does agent 1 if and only if they perceive identical subjective states, and for any signal of agent 1, agent 2 expects at least one less informative signal in period 1.

Part (iii) is weaker than saying that one subjective decision tree is finer than the other. Since $\mu \in S_1^i$ is a belief over $\Delta(\Omega)$, it is possible that $\mathrm{supp}(\mu) \cap \mathrm{supp}(\mu') \neq \emptyset$ even if μ and μ' are distinct signals. If S_1^i is "partitional", that is, $\mathrm{supp}(\mu) \cap \mathrm{supp}(\mu') = \emptyset$ for all $\mu, \mu' \in S_1^i$, part (iii) is equivalent to saying that $(\mathcal{S}^2, \mu_0^2)$ is "finer" than $(\mathcal{S}^1, \mu_0^1)$.

[7] By Monotonicity, $\{x_1\} \cup x_0 \succeq^i \{x_1\}$. The other direction follows from Lemma B.1 in Appendix B.2.

[8] In Theorem 2 (1) of DLR, the implication only works in one direction. If preference satisfies Monotonicity, the other direction is also true.

A Hausdorff Metric

Let d be a metric on $\mathcal{H}$. Let

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

For each $x_1, y_1 \in \mathcal{K}(\mathcal{H})$, define

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

Then, d_h is called the *Hausdorff metric* on $\mathcal{K}(\mathcal{H})$. It is known that $\mathcal{K}(\mathcal{H})$ is a compact metric space under this metric.

The Hausdorff metric on $\mathcal{D}$ is similarly constructed from d_h. Let

$$D(x_1, x_0) \equiv \min_{x_1' \in x_0} d_h(x_1, x_1'), \text{ and } E(x_0', x_0) \equiv \max_{x_1' \in x_0'} D(x_1', x_0).$$

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 $\mathcal{K}(\mathcal{H})$ is compact under the Hausdorff metric d_h, so is $\mathcal{K}(\mathcal{K}(\mathcal{H}))$ under d_H.

B Proofs

B.1 Proof of Proposition 3.1

(i) Let (μ_0, u) be a canonical representation. Let $S_1 \equiv \text{supp}(\mu_0)$ and $S_2 \equiv \Delta(\Omega)$. Define $\mu_1 : S_1 \to \Delta(S_2)$ as the identity mapping, that is, $\mu_1(s_1) = s_1$. Define $\mu_2 : S_1 \times S_2 \to \Delta(\Omega)$ by $\mu_2(s_1, s_2) = s_2$. Since they are continuous, μ_1 and μ_2 are measurable with respect to the Borel σ-algebra.

(ii) Let $(\mathcal{S}, \mu_0^*)$ be the subjective decision tree derived from $(\{S_t\}_{t=1}^2, \{\mu_t\}_{t=0}^2, u)$. Let $\varphi(s_1) \equiv \mu_1(\cdot|s_1) \circ \mu_2^{-1} \in \Delta(\Delta(\Omega))$. That is, $\varphi(s_1)$ is the distribution of $\mu_1(s_1)$ induced by $\mu_2 : S \to \Delta(\Omega)$. By definition, μ_0^* is the distribution of μ_0 induced by the mapping $\varphi : S_1 \to \Delta(\Delta(\Omega))$. By the change-of-variable formulas, for all x_1,

$$\begin{aligned} U_1(x_1, s_1) &= \int_{S_2} \max_{h \in x_1} \Big(\sum u(h(\omega)) \mu_2(\omega | s_1, s_2) \Big) \mathrm{d}\mu_1(s_2 | s_1) \\ &= \int_{\Delta(\Omega)} \max_{h \in x_1} \Big(\sum u(h(\omega)) p(\omega) \Big) \mathrm{d}\varphi(s_1)(p) \\ &\equiv U_1^*(x_1, \varphi(s_1)), \end{aligned}$$

and, for all x_0,

$$\begin{aligned}
U_0(x_0) &= \int_{S_1} \max_{x_1 \in x_0} U_1(x_1, s_1)\, \mathrm{d}\mu_0(s_1) \\
&= \int_{S_1} \max_{x_1 \in x_0} U_1^*(x_1, \varphi(s_1))\, \mathrm{d}\mu_0(s_1) \\
&= \int_{\Delta(\Delta(\Omega))} \max_{x_1 \in x_0} U_1^*(x_1, \mu)\, \mathrm{d}\mu_0 \circ \varphi^{-1}(\mu) \\
&= \int_{\Delta(\Delta(\Omega))} \max_{x_1 \in x_0} U_1^*(x_1, \mu)\, \mathrm{d}\mu_0^*(\mu) \\
&\equiv U_0^*(x_0).
\end{aligned}$$

Since $U_0 = U_0^*$, the canonical form U_0^* with components (μ_0^*, u) represents the identical preference.

B.2 Proof of Theorem 3.1

B.2.1 Outline of the Proof

Necessity of the axioms is routine. We show sufficiency. Recall the key lemma for the DLR's representation (Lemma 3.1, p.915) about characterization of a compact convex menu via its support function. We analogously identify a menu of menus x_0 with a (suitably defined) support function. For any $p \in S_2 \equiv \Delta(\Omega)$, let

$$U_2(h, p) \equiv \sum_{\omega} u(h(\omega))p(\omega),$$

where u is a mixture linear function representing commitment ranking over $\Delta(Z)$. For any $\mu \in S_1 \equiv \Delta(\Delta(\Omega))$, let

$$U_1(x_1, \mu) \equiv \int_{S_2} \max_{h \in x_1} U_2(h, p)\, \mathrm{d}\mu(p).$$

Now, for any $x_0 \in \mathcal{D} \equiv \mathcal{K}(\mathcal{K}(\mathcal{H}))$, define the "support function" $\sigma_{x_0} : S_1 \to \mathbb{R}$ by

$$\sigma_{x_0}(\mu) \equiv \max_{x_1 \in x_0} U_1(x_1, \mu). \tag{12}$$

The key step in the proof is to show that support functions identify menus of menus up to indifference, that is,

$$\sigma_{x_0} = \sigma_{y_0} \;\Leftrightarrow\; x_0 \sim y_0. \tag{13}$$

This does not follow from convexity theory: in particular, since support functions defined by (12) are not standard, it is not the case that $\sigma_{x_0} = \sigma_{y_0} \Rightarrow x_0 = y_0$, even when x_0 and y_0 are convex. Indeed, notice the following:

(i) Any x_0 and its closed convex hull $\overline{\mathrm{co}}(x_0)$ have the same support function because $U_1(\cdot, \mu)$ is mixture linear with respect to mixture operation (3).

(ii) Let $\overline{\text{co}}_1(x_0)$ be the set of all closed convex hulls $\overline{\text{co}}(x_1)$ as x_1 varies over x_0. Then x_0 and $\overline{\text{co}}_1(x_0)$ have the same support function because $U_2(\cdot, p)$ is mixture linear.

(iii) Recall that $O_1(x_1) \in \mathcal{K}(\mathcal{H})$ is the menu of all acts dominated by some act in x_1. (See (6) in Section 2.2 for details.) Let $O(x_0) \in \mathcal{D}$ be the set of all $O_1(x_1)$ as x_1 varies over x_0. Any x_0 and $O(x_0)$ have the same support function because $U_2(h, p) \geq U_2(h', p)$ whenever h "dominates" h'.

(iv) Let $I(x_0)$ be the set of all non-empty compact menus y_1, where $y_1 \subset x_1$ for some $x_1 \in x_0$. Then x_0 and $I(x_0)$ have the same support function because $U_1(x_1, \mu) \geq U_1(y_1, \mu)$ whenever $y_1 \subset x_1$.

These observations imply that x_0 and x_0' have identical support functions if x_0' can be derived from x_0 by a finite sequence of the above steps.

Let $\text{CO}(x_0)$ be the set of all menus $O_1(\overline{\text{co}}(y_1))$, where y_1 is a non-empty compact subset of $O_1(\overline{\text{co}}(x_1))$ and x_1 varies over $\overline{\text{co}}(x_0)$. We show that: (1) the axioms imply $x_0 \sim \text{CO}(x_0)$; and (2)

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

These two steps and the above observations (i)-(iv) imply (13).

The remaining part of the proof is to show that there exists a unique probability measure μ_0 over S_1 such that

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

represents preference over the set of $\text{CO}(x_0)$'s. For this step, we adapt the argument in DLR.

B.2.2 Proof of Sufficiency

As a preliminary result, we 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 B.1. *If x_0 covers y_0, then $x_0 \succeq y_0$.*

Proof. We first show the statement when y_0 is finite.

Step 1: If $\succeq$ satisfies Monotonicity and Aversion to Commitment and if y_0 is finite, then $x_0 \succeq y_0$ whenever x_0 covers y_0.

Denote y_0 by $\{y_1^i | i = 1, \cdots, I\}$. 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 $z_0^i \equiv \{x_1^i \setminus y_1^i, y_1^i\}$. 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, $x_0 \succeq y_0$.

Now we turn to the general case. 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 [7], there exists a finite subset $y_0^* \subset y_0$ such that $y_0^* \succ x_0$. Since x_0 covers y_0^*, Step 1 implies $x_0 \succeq y_0^*$. 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.

Order, Continuity and Independence ensure a mixture linear representation $U : \mathcal{D}_1 \to \mathbb{R}$ because $\mathcal{D}_1$ is a mixture space. Let $u : \Delta(Z) \to \mathbb{R}$ be the restriction of U on $\Delta(Z)$, that is, $u(l) \equiv U(\{\{l\}\})$. Since $\Delta(Z)$ is compact, there exist a maximal element $\bar{l}$ and a minimal element $\underline{l}$ with respect to u. Since Strong Nondegeneracy implies that u is not constant, we can assume $u(\underline{l}) = 0$ and $u(\bar{l}) = 1$ without loss of generality.

Say that a metric ρ on $\Delta(Z)$ is the Prohorv metric if, for any $\mu, \nu \in \Delta(Z)$,

$$\rho(\mu, \nu) \equiv \inf\{\varepsilon > 0 | \mu(E) \leq \nu(U_\varepsilon(E)) + \varepsilon,\ \nu(E) \leq \mu(U_\varepsilon(E)) + \varepsilon,\ \forall E \in \mathcal{B}(Z)\},$$

where $U_\varepsilon(E) \equiv \{z \in Z | \inf_{z' \in E} \pi(z, z') \leq \varepsilon\}$ and π is a metric on Z. It is well-known that the weak convergence topology is metrizable by this metric. Furthermore, the product metric on $\mathcal{H} \equiv \Delta(Z)^n$ is equivalent to the topology generated by $d(h, h') \equiv (\sum_\omega \rho(h(\omega), h'(\omega))^p)^{1/p}$ for some $p \geq 1$. Hence, without loss of generality, we can choose these metrics.

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

$$\overline{\text{co}}_1(x_0) \equiv \{\overline{\text{co}}(x_1) | x_1 \in x_0\}. \tag{14}$$

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

Lemma B.2.

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

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

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

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

Proof. (i) Consider the closed convex hull operator $\overline{\text{co}}(\cdot) : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$. First of all, since $\mathcal{K}(\mathcal{H})$ is compact, $\overline{\text{co}}(x_1) \in \mathcal{K}(\mathcal{H})$. Hence, this operator is well-defined. To show $\overline{\text{co}}_1(x_0) \in \mathcal{D}$, it suffices to show that $\overline{\text{co}}(\cdot) : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ is Hausdorff continuous.

It can be shown that the Prohorov metric ρ has the following properties: (1) $\rho(\alpha\mu, \alpha\nu) = \alpha\rho(\mu, \nu)$ for any $\alpha > 0$ and $\mu, \nu \in \mathcal{M}(Z)$, where $\mathcal{M}(Z)$ is the set of non-negative countably additive measures over Z; and (2) $\rho(\mu + \nu, \mu' + \nu') \leq \rho(\mu, \nu) + \rho(\mu', \nu')$ for all μ, μ', ν and $\nu' \in \mathcal{M}(Z)$. Hence, the metric d on $\mathcal{H}$ has the similar properties.

Recall the notation in Appendix A,

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

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

Take any $h_1, h_2 \in \mathcal{H}$, and $\lambda \in [0,1]$. Let $h_i' \equiv \operatorname{argmin}_{h' \in x_1'} d(h_i, h')$, $i = 1, 2$. Then,

$$\begin{aligned}
\lambda d(h_1, x_1') + (1-\lambda) d(h_2, x_1') &= d(\lambda h_1, \lambda h_1') + d((1-\lambda) h_2, (1-\lambda) h_2') \\
&\geq d(\lambda h_1 + (1-\lambda) h_2, \lambda h_1' + (1-\lambda) h_2') \\
&\geq \min_{h' \in x_1'} d(\lambda h_1 + (1-\lambda) h_2, h') \\
&= d(\lambda h_1 + (1-\lambda) h_2, x_1').
\end{aligned}$$

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

To show that $\overline{\text{co}}(\cdot)$ is Hausdorff continuous, the next step is sufficient.

Step 2: $d_h(\overline{\text{co}}(x_1), \overline{\text{co}}(y_1)) \leq d_h(x_1, y_1)$ for all $x_1, y_1 \in \mathcal{K}(\mathcal{H})$.

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

$$d(h, \overline{\text{co}}(y_1)) \leq e(x_1, \overline{\text{co}}(y_1)). \tag{15}$$

Since $d(\cdot, \overline{\text{co}}(y_1))$ is a convex function by Step 1, (15) holds for any $h \in \overline{\text{co}}(x_1)$. Thus,

$$e(\overline{\text{co}}(x_1), \overline{\text{co}}(y_1)) \leq e(x_1, \overline{\text{co}}(y_1)). \tag{16}$$

On the other hand, since $y_1 \subset \overline{\text{co}}(y_1)$,

$$e(x_1, \overline{\text{co}}(y_1)) \leq e(x_1, y_1). \tag{17}$$

Taking (16) and (17) together,

$$e(\overline{\text{co}}(x_1), \overline{\text{co}}(y_1)) \leq e(x_1, y_1). \tag{18}$$

By the same argument, (18) holds when x_1 and y_1 are reversed. Hence,

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

(ii) From Dunford and Schwartz [3, Lemma 4 (iii) and (iv), p.415], $\overline{\text{co}}(\cdot) : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ is mixture linear, that is, for any $x_1, y_1 \in \mathcal{K}(\mathcal{H})$ and $\lambda \in [0,1]$,

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

Since a mixture linear operator preserves convexity, $\overline{\text{co}}_1(x_0)$ is convex whenever $x_0 \subset \mathcal{K}(\mathcal{H})$ is convex.

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

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

By Step 2 in the proof of part (i),

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

By assumption, $d_H(x_0^n, x_0)$ converges to zero. Hence, $\overline{\text{co}}_1(x_0^n) \to \overline{\text{co}}_1(x_0)$.

(iv) Since $x_1 \subset \overline{\text{co}}(x_1)$, Lemma B.1 implies $\overline{\text{co}}_1(x_0) \succeq x_0$. We will show the converse direction step by step.

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 $\text{co}(\overline{\text{co}}_1(x_0)) \subset \lambda \text{co}(x_0) + (1 - \lambda)\text{co}(\overline{\text{co}}_1(x_0))$.

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

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

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

Step 2: For any $x_0 \in \mathcal{D}$ satisfying the condition of Step 1, $x_0 \succeq \overline{\text{co}}_1(x_0)$.

Suppose otherwise, that is, $\overline{\text{co}}_1(x_0) \succ x_0$. Since $\text{co}(\overline{\text{co}}_1(x_0)) \sim \overline{\text{co}}_1(x_0)$ and $\text{co}(x_0) \sim x_0$, $\text{co}(\overline{\text{co}}_1(x_0)) \succ \text{co}(x_0)$. Independence requires that, for any $\lambda \in (0, 1]$,

$$\lambda \text{co}(\overline{\text{co}}_1(x_0)) + (1 - \lambda)\text{co}(\overline{\text{co}}_1(x_0)) \succ \lambda \text{co}(x_0) + (1 - \lambda)\text{co}(\overline{\text{co}}_1(x_0)).$$

Monotonicity and Step 1 imply that $\lambda \text{co}(x_0) + (1-\lambda)\text{co}(\overline{\text{co}}_1(x_0)) \succeq \text{co}(\overline{\text{co}}_1(x_0))$ for some $\lambda \in (0, 1]$. Since $\lambda \text{co}(\overline{\text{co}}_1(x_0)) + (1 - \lambda)\text{co}(\overline{\text{co}}_1(x_0)) = \text{co}(\overline{\text{co}}_1(x_0))$, we have a contradiction.

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

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, $x_0^n \succeq \overline{\text{co}}_1(x_0^n)$. Part (iii) implies $x_0 \succeq \overline{\text{co}}_1(x_0)$. $\square$

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

$$I(x_1) \equiv \{y_1 \in \mathcal{K}(\mathcal{H}) | y_1 \subset x_1\}. \tag{19}$$

Let $I(x_0) \equiv \cup_{x_1 \in x_0} I(x_1)$.

Lemma B.3.

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

(ii) *For all $x_0 \in \mathcal{D}_1$, $I(x_0) \in \mathcal{D}_1$.*

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

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

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

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

(iii) 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, we assume without loss of generality that $I(x_0^n) \to z_0$ for some point $z_0 \in \mathcal{D}$. We want to show $I(x_0) = z_0$.

Step 1: $z_0 \subset I(x_0)$.

Let $z_1 \in z_0$. Since $I(x_0^n) \to z_0$, we can find a sequence $z_1^n \to z_1$ with $z_1^n \in I(x_0^n)$. There exists a sequence $x_1^n \in x_0^n$ with $z_1^n \subset x_1^n$. Since $\{x_1^n\}$ is a sequence in $\mathcal{K}(\mathcal{H})$, we can assume $x_1^n \to x_1$ for some $x_1 \in \mathcal{K}(\mathcal{H})$. 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 $z_1 \in I(x_0)$ because $z_1 \subset x_1$.

Step 2: $I(x_0) \subset z_0$.

Let $y_1 \in I(x_0)$. There exists $x_1 \in x_0$ such that $y_1 \subset x_1$. 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 $z_1^m \to y_1$ such that z_1^m is a finite subset of y_1. Take the open $1/m$-neighborhood of y_1, denoted by $B(y_1, 1/m)$. We can assume without loss of generality that $z_1^m \in B(y_1, 1/m)$ 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 $y_1^{n_m} \in B(y_1, 1/m)$. Since $\mathcal{K}(\mathcal{H})$ is compact, the subsequence $\{y_1^{n_m}\}_{m=0}^{\infty}$ converges to y_1. Since $I(x_0^{n_m}) \to z_0$ and $y_1^{n_m} \to y_1$ with $y_1^{n_m} \in I(x_0^{n_m})$, we have $y_1 \in z_0$.

(iv) Since $x_0 \subset I(x_0)$, Monotonicity implies $I(x_0) \succeq x_0$. By definition of I, for any $y_1 \in I(x_0)$, there exists $x_1 \in x_0$ such that $y_1 \subset x_1$. That is, x_0 covers $I(x_0)$. Lemma B.1 implies $x_0 \succeq I(x_0)$, and hence $x_0 \sim I(x_0)$. □

Lemma B.4.

(i) *For all $x_1 \in \mathcal{K}(\mathcal{H})$, $O_1(x_1) \in \mathcal{K}(\mathcal{H})$.*

(ii) *If $x_1 \in \mathcal{K}(\mathcal{H})$ is convex, $O_1(x_1)$ is also convex.*

(iii) *For all $x_0 \in \mathcal{D}$, $O(x_0) \in \mathcal{D}$.*

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

(v) *$O : \mathcal{D} \to \mathcal{D}$ is Hausdorff continuous.*

Proof. (i) We want to show that $O_1(x_1)$ is compact. Since $\mathcal{K}(\mathcal{H})$ is a compact metric space, it suffices to show that $O_1(x_1)$ is closed. Let $h^n \to h$ with $h^n \in O_1(x_1)$. Then there exists a sequence $\{k^n\}_{n=0}^{\infty}$ in x_1 satisfying $\{k^n(\omega)\} \succeq \{h^n(\omega)\}$ for all ω. Since $\Delta(Z)$ is compact, for each ω, the sequence $\{k^n(\omega)\}_{n=0}^{\infty}$ has a convergent subsequence $\{k^{n_i}(\omega)\}_{i=0}^{\infty}$ with the limit point $l_\omega \in \Delta(Z)$. Define $k^* \in \mathcal{H}$ by $k^*(\omega) \equiv l_\omega$. Since Ω is finite, we can find a subsequence $\{k^m\}_{m=0}^{\infty}$ of $\{k^n\}_{n=0}^{\infty}$ satisfying $k^m \to k^*$. Notice that $k^* \in x_1$. Since $\{k^m(\omega)\} \succeq \{h^m(\omega)\}$ for all ω, Continuity implies $\{k^*(\omega)\} \succeq \{h(\omega)\}$. Thus, $h \in O_1(k^*) \subset O_1(x_1)$.

(ii) Take any $h, h' \in O_1(x_1)$. There exist $k, k' \in x_1$ such that $\{k(\omega)\} \succeq \{h(\omega)\}$ and $\{k'(\omega)\} \succeq \{h'(\omega)\}$ for all ω. Since x_1 is convex, $\lambda k + (1-\lambda)k' \in x_1$ for any $\lambda \in [0, 1]$. Order and Independence imply that, for all ω,

$$\lambda\{k(\omega)\} + (1-\lambda)\{k'(\omega)\} \succeq \lambda\{h(\omega)\} + (1-\lambda)\{h'(\omega)\},$$

equivalently,

$$\{\lambda k(\omega) + (1-\lambda)k'(\omega)\} \succeq \{\lambda h(\omega) + (1-\lambda)h'(\omega)\}.$$

Hence, $\lambda h + (1-\lambda)h' \in O_1(x_1)$.

(iii) It suffices to show that $O_1(\cdot) : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ is Hausdorff continuous. Let $x_1^n \to x_1$. Since $\mathcal{K}(\mathcal{H})$ is compact, the sequence $\{O_1(x_1^n)\}_{n=1}^{\infty}$ has a convergent subsequence $\{O_1(x_1^m)\}_{m=1}^{\infty}$ with the limit $y_1 \in \mathcal{K}(\mathcal{H})$. We want to show $O_1(x_1) = y_1$.

Step 1: $O_1(x_1) \subset y_1$.

Let $h \in O_1(x_1)$. There exists $\bar{h} \in x_1$ such that $\{\{\bar{h}(\omega)\}\} \succeq \{\{h(\omega)\}\}$ for all ω. Since $x_1^m \to x_1$, we can find a sequence $\{\bar{h}^m\}_{m=1}^{\infty}$ in $\mathcal{H}$ satisfying $\bar{h}^m \in x_1^m$ and $\bar{h}^m \to \bar{h}$ in the sense of the metric on $\mathcal{H}$, equivalently, for all ω, $\bar{h}^m(\omega) \to \bar{h}(\omega)$ in the sense of the metric on $\Delta(Z)$. To show $h \in y_1$, it suffices to find a sequence $\{h^m\}_{m=1}^{\infty}$ such that $h^m \to h$ with $h^m \in O_1(x_1^m)$. Indeed, $O_1(x_1^m) \to y_1$ and $h^m \to h$ with $h^m \in O_1(x_1^n)$ imply $h \in y_1$.

Fix an arbitrary ω. There exist two cases: (1) $\{\{\bar{h}(\omega)\}\} \succ \{\{h(\omega)\}\}$; and (2) $\{\{\bar{h}(\omega)\}\} \sim \{\{h(\omega)\}\}$. If case (1) holds, Continuity implies $\{\{\bar{h}^m(\omega)\}\} \succ \{\{h(\omega)\}\}$ for all m sufficiently large. Hence, define $h^m(\omega) \equiv h(\omega)$ for all m sufficiently large and, otherwise, $h^m(\omega) \equiv l^*$ for some lottery l^*. If case (2) holds, define $h^m(\omega) \equiv h(\omega)$ as long as $\{\{\bar{h}^m(\omega)\}\} \succeq \{\{\bar{h}(\omega)\}\} \sim \{\{h(\omega)\}\}$. Otherwise, take the first natural number $k \geq 1$ satisfying $\{\{\bar{h}(\omega)\}\} \sim \{\{h(\omega)\}\} \succ \{\{\bar{h}^k(\omega)\}\}$. Let $l_\omega(\lambda) \equiv \lambda \bar{h}^k(\omega) + (1-\lambda)h(\omega)$. Continuity ensures that there exists λ^m such that $\{\{l_\omega(\lambda^m)\}\} \sim \{\{\bar{h}^m(\omega)\}\}$ for all sufficiently large m satisfying $\{\{\bar{h}(\omega)\}\} \succ \{\{\bar{h}^m(\omega)\}\}$. Since $\bar{h}^m(\omega) \to \bar{h}(\omega)$,

$\lambda^m \to 0$ as $m \to \infty$. Define $h^m(\omega) \equiv l_\omega(\lambda^m)$. Case (1) and (2) imply that there exists a sequence $\{h^m\}_{m=1}^\infty$ such that $h^m \to h$ and $\{\{\bar{h}^m(\omega)\}\} \succeq \{\{h^m(\omega)\}\}$ for all ω, and hence $h^m \in O_1(x_1^m)$.

Step 2: $y_1 \subset O_1(x_1)$.

Take $h \in y_1$. Since $O_1(x_1^m) \to y_1$, there exists a sequence $h^m \in O_1(x_1^m)$ with $h^m \to h$ in the sense of the metric on $\mathcal{H}$. By definition, there exists $\bar{h}^m \in x_1^m$ such that $\{\{\bar{h}^m(\omega)\}\} \succeq \{\{h^m(\omega)\}\}$ for all ω. Since $\mathcal{H}$ is compact, we can assume that $\{\bar{h}^m\}$ converges to the limit $\bar{h} \in \mathcal{H}$. Since $\bar{h}^m \to \bar{h}$ and $x_1^m \to x_1$ with $\bar{h}^m \in x_1^m$, $\bar{h} \in x_1$. Furthermore, Continuity implies $\{\{\bar{h}(\omega)\}\} \succeq \{\{h(\omega)\}\}$ for all ω. Hence, $h \in O_1(x_1)$.

(iv) It suffices to show that $O_1(\cdot) : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ is mixture linear, that is, for any $x_1, x_1' \in \mathcal{K}(\mathcal{H})$ and $\lambda \in [0,1]$,

$$\lambda O_1(x_1) + (1-\lambda) O_1(x_1') = O_1(\lambda x_1 + (1-\lambda) x_1').$$

Step 1: $\lambda O_1(x_1) + (1-\lambda) O_1(x_1') \subset O_1(\lambda x_1 + (1-\lambda) x_1')$.

Take any $h'' \in \lambda O_1(x_1) + (1-\lambda) O_1(x_1')$. There exist $h \in O_1(x_1)$ and $h' \in O_1(x_1')$ satisfying $h'' = \lambda h + (1-\lambda) h'$. By definition, there exist $\bar{h} \in x_1$ and $\bar{h}' \in x_1'$ such that

$$\{\{\bar{h}(\omega)\}\} \succeq \{\{h(\omega)\}\}, \text{ and } \{\{\bar{h}'(\omega)\}\} \succeq \{\{h'(\omega)\}\},$$

for all ω. Take $\lambda\bar{h} + (1-\lambda)\bar{h}' \in \lambda x_1 + (1-\lambda) x_1'$. By Independence,

$$\begin{aligned}
\{\{\lambda\bar{h}(\omega) + (1-\lambda)\bar{h}'(\omega)\}\} &= \lambda\{\{\bar{h}(\omega)\}\} + (1-\lambda)\{\{\bar{h}'(\omega)\}\}, \\
&\succeq \lambda\{\{h(\omega)\}\} + (1-\lambda)\{\{h'(\omega)\}\}, \\
&= \{\{\lambda h(\omega) + (1-\lambda) h'(\omega)\}\}.
\end{aligned}$$

Hence, $h'' \in O_1(\lambda x_1 + (1-\lambda) x_1')$.

Step 2: $O_1(\lambda x_1 + (1-\lambda) x_1') \subset \lambda O_1(x_1) + (1-\lambda) O_1(x_1')$.

Take any $h'' \in O_1(\lambda x_1 + (1-\lambda) x_1')$. There exist $\bar{h} \in x_1$ and $\bar{h}' \in x_1'$ satisfying $\{\{\lambda\bar{h}(\omega) + (1-\lambda)\bar{h}'(\omega)\}\} \succeq \{\{h''(\omega)\}\}$ for all ω. We will find $h \in O_1(x_1)$ and $h' \in O_1(x_1')$ satisfying $h'' = \lambda h + (1-\lambda) h'$. Fix an arbitrarily ω. Assume first that $\{\{\bar{h}(\omega)\}\} \succeq \{\{\bar{h}'(\omega)\}\}$. By Independence,

$$\{\{\bar{h}(\omega)\}\} \succeq \{\{\lambda\bar{h}(\omega) + (1-\lambda)\bar{h}'(\omega)\}\} \succeq \{\{\bar{h}'(\omega)\}\}.$$

We have the following two cases: (1) $\{\{\bar{h}'(\omega)\}\} \succeq \{\{h''(\omega)\}\}$; and (2) $\{\{\bar{h}(\omega)\}\} \succeq \{\{h''(\omega)\}\} \succ \{\{\bar{h}'(\omega)\}\}$.

If case (1) holds, define $h(\omega) = h'(\omega) \equiv h''(\omega)$. Since $\{\{\bar{h}(\omega)\}\} \succeq \{\{\bar{h}'(\omega)\}\}$,

$$\{\{\bar{h}(\omega)\}\} \succeq \{\{h(\omega)\}\} \text{ and } \{\{\bar{h}'(\omega)\}\} \succeq \{\{h'(\omega)\}\}.$$

Moreover, $h''(\omega) = \lambda h(\omega) + (1-\lambda) h'(\omega)$.

If case (2) holds, take two lotteries l_ω and l'_ω such that $l_\omega \sim \bar{h}(\omega)$, $l'_\omega \sim \bar{h}'(\omega)$ and $h''(\omega) = \alpha l_\omega + (1-\alpha)l'_\omega$ for some $\alpha \in (0,1]$. From Independence, $\lambda \geq \alpha$. Define

$$h(\omega) \equiv \frac{\alpha}{\lambda} l_\omega + \left(1 - \frac{\alpha}{\lambda}\right) l'_\omega, \text{ and } h'(\omega) \equiv l'_\omega.$$

Then, we have

$$\{\{\bar{h}(\omega)\}\} \succeq \{\{h(\omega)\}\} \text{ and } \{\{\bar{h}'(\omega)\}\} \succeq \{\{h'(\omega)\}\},$$

and $h''(\omega) = \lambda h(\omega) + (1-\lambda)h'(\omega)$. By the symmetric argument, we can find such $h(\omega)$ and $h'(\omega)$ even when $\{\{\bar{h}'(\omega)\}\} \succeq \{\{\bar{h}(\omega)\}\}$.

Now h and h' constructed as above satisfy $h \in O_1(\bar{h}) \subset O_1(x_1)$, $h' \in O_1(\bar{h}') \subset O_1(x'_1)$, and $h'' = \lambda h + (1-\lambda)h'$. Therefore, $h'' \in \lambda O_1(x_1) + (1-\lambda)O_1(x'_1)$.

(v) Let $x_0^n \to x_0$ with $x_0^n, x_0 \in \mathcal{D}$. We want to show $O(x_0^n) \to O(x_0)$. By definition,

$$d_H(O(x_0^n), O(x_0)) = \max\left[\max_{x_1 \in x_0^n} \min_{y_1 \in x_0} d_h(O_1(x_1), O_1(y_1)),\ \max_{y_1 \in x_0} \min_{x_1 \in x_0^n} d_h(O_1(x_1), O_1(y_1))\right].$$

We have to show that, for any $\varepsilon > 0$, there exists N such that, for all $n \geq N$,

$$\max_{x_1 \in x_0^n} \min_{y_1 \in x_0} d_h(O_1(x_1), O_1(y_1)) < \varepsilon, \text{ and } \max_{y_1 \in x_0} \min_{x_1 \in x_0^n} d_h(O_1(x_1), O_1(y_1)) < \varepsilon. \tag{20}$$

Step 1: For any $\varepsilon > 0$, there exists N_1 such that, for all $n \geq N_1$,

$$\max_{x_1 \in x_0^n} \min_{y_1 \in x_0} d_h(O_1(x_1), O_1(y_1)) < \varepsilon.$$

Suppose otherwise. Then there exists $\varepsilon > 0$ such that

$$\max_{x_1 \in x_0^n} \min_{y_1 \in x_0} d_h(O_1(x_1), O_1(y_1)) \geq \varepsilon$$

for any n. There exists $x_1^n \in x_0^n$ satisfying $\min_{y_1 \in x_0} d_h(O_1(x_1^n), O_1(y_1)) \geq \varepsilon$. Since $\mathcal{K}(\mathcal{H})$ is compact, we can assume that the sequence $\{x_1^n\}$ converges to the limit x_1^*. Since $x_0^n \to x_0$ and $x_1^n \to x_1^*$ with $x_1^n \in x_0^n$, $x_1^* \in x^0$. From Hausdorff continuity of $O_1 : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ (Lemma B.4 (iii)), $O_1(x_1^n) \to O_1(x_1^*)$. Since $d_h(O_1(x_1^n), O_1(y_1)) \geq \varepsilon$ for all n and $y_1 \in x_0$, we have $d_h(O_1(x_1^*), O_1(y_1)) \geq \varepsilon$ for all $y_1 \in x_0$, which contradicts $x_1^* \in x_0$.

Step 2: For any $\varepsilon > 0$, there exists N_2 such that, for all $n \geq N_2$,

$$\max_{y_1 \in x_0} \min_{x_1 \in x_0^n} d_h(O_1(x_1), O_1(y_1)) < \varepsilon.$$

Suppose otherwise. Then there exists $\varepsilon > 0$ such that

$$\max_{y_1 \in x_0} \min_{x_1 \in x_0^n} d_h(O_1(x_1), O_1(y_1)) \geq \varepsilon$$

for any n. There exists $y_1^n \in x_0$ satisfying $\min_{x_1 \in x_0^n} d_h(O_1(x_1), O_1(y_1^n)) \geq \varepsilon$ for all n. Since x_0 is compact, we can assume that the sequence $\{y_1^n\}$ converges to the limit $y_1^* \in x_0$. Since $x_0^n \to x_0$, we can find a sequence $x_1^n \in x_0^n$ such that $x_1^n \to y_1^*$. By Hausdorff continuity of O_1, $O_1(x_1^n) \to O_1(y_1^*)$ and $O_1(y_1^n) \to O_1(y_1^*)$. Since $d_h(O_1(x_1^n), O_1(y_1^n)) \geq \varepsilon$ for all n, we have $0 = d_h(O_1(y_1^*), O_1(y_1^*)) \geq \varepsilon$, which is a contradiction.

Let $N = \max[N_1, N_2]$. By Step 1 and 2, we have (20). □

For all $x_0 \in \mathcal{D}_1$, define

$$\mathrm{CO}(x_0) \equiv O(\overline{\mathrm{co}}_1(I(O(\overline{\mathrm{co}}_1(x_0))))). \tag{21}$$

That is, $\mathrm{CO}(x_0)$ is the set of all menus $O_1(\overline{\mathrm{co}}(y_1))$, where y_1 is a non-empty compact subset of $O_1(\overline{\mathrm{co}}(x_1))$ and x_1 varies over x_0. Lemma B.2 (ii), (iii), B.3 (ii), (iii), B.4 (iv) and (v) imply that $\mathrm{CO} : \mathcal{D}_1 \to \mathcal{D}_1$ is Hausdorff continuous.

From Lemma B.2 (iv), B.3 (iv) and Risk Preference Certainty, we can pay attention to the sub-domain

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

Since CO is Hausdorff continuous, $\mathcal{D}_2$ is compact. For any $x_1 \in x_0 \in \mathcal{D}_2$, $O_1(x_1) = x_1$. Moreover, if $y_1 \subset x_1$ and $O_1(y_1) = y_1$, then $y_1 \in x_0$.

Let

$$\mathcal{K}^*(\mathcal{H}) \equiv \{x_1 \in \mathcal{K}(\mathcal{H}) | x_1 = O_1(\overline{\mathrm{co}}(x_1))\}.$$

From Lemma B.4 (i), $\mathcal{K}^*(\mathcal{H})$ is well-defined. From Lemma B.4 (ii), any $x_1 \in \mathcal{K}^*(\mathcal{H})$ is convex. From Lemma B.2 (i), (ii), Lemma B.4 (iii), and (iv), both $\overline{\mathrm{co}} : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ and $O_1 : \mathcal{K}(\mathcal{H}) \to \mathcal{K}(\mathcal{H})$ are Hausdorff continuous and mixture linear. Hence, $\mathcal{K}^*(\mathcal{H})$ is compact and convex.

Let $S_2 \equiv \Delta(\Omega)$. Since $\#\Omega = n$, S_2 is identified with the $(n-1)$-dimensional unit simplex. For all $x_1 \in \mathcal{K}^*(\mathcal{H})$ and $p \in S_2$, let

$$U_2(h, p) \equiv \sum_{\omega \in \Omega} u(h(\omega)) p(\omega)$$

and

$$\zeta_{x_1}(p) \equiv \max_{h \in x_1} U_2(h, p). \tag{22}$$

(22) defines the function $\zeta : \mathcal{K}^*(\mathcal{H}) \to \mathcal{C}(S_2)$, where $\mathcal{C}(S_2)$ is the set of all real-valued continuous functions on S_2 with the sup-norm. As shown in Takeoka [20, Lemma B.2, p. 21-22], ζ is injective, continuous and mixture linear.

Let $S_1 \equiv \Delta(S_2)$. Since S_2 is a compact metric space, S_1 is also a compact metric space under the weak convergence topology. For all $x_1 \in \mathcal{K}^*(\mathcal{H})$ and $\mu \in S_1$, define

$$U_1(x_1, \mu) \equiv \int_{S_2} \zeta_{x_1}(p)\, \mathrm{d}\mu(p) = \int_{S_2} \max_{h \in x_1} U_2(h, p)\, \mathrm{d}\mu(p).$$

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

$$\sigma_{x_0}(\mu) \equiv \max_{x_1 \in x_0} U_1(x_1, \mu). \tag{23}$$

Since any $x_1 \in x_0 \in \mathcal{D}_2$ belongs to $\mathcal{K}^*(\mathcal{H})$, $\sigma_{x_0}(\mu)$ is well-defined. Now (23) defines the function $\sigma : \mathcal{D}_2 \to \mathcal{C}(S_1)$, where $\mathcal{C}(S_1)$ is the set of all real-valued continuous functions on S_1 with the sup-norm.

Lemma B.5.

(i) *σ is continuous.*

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

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

Proof. (i) For each $x_0 \in \mathcal{D}_2$, define $u(x_0) \equiv \{u(x_1) | x_1 \in x_0\}$, where

$$u(x_1) \equiv \{(u(h(\omega)))_{\omega \in \Omega} \in \mathbb{R}^n | h \in x_1\}. \tag{24}$$

Since $u : \Delta(Z) \to [0,1]$ is continuous and mixture linear, $u(x_1) \subset [0,1]^n$ is a compact and convex set. From Step 1 in Lemma B.2 of Takeoka [20, p. 22], the function $u : \mathcal{K}^*(\mathcal{H}) \to \mathcal{K}([0,1]^n)$, defined by (24), is Hausdorff continuous, where $\mathcal{K}([0,1]^n)$ is the set of all non-empty compact subsets of $[0,1]^n$ with the Hausdorff metric. Hence, $u(x_0)$ is compact. Let $\mathcal{K}(\mathcal{K}([0,1]^n))$ be the set of non-empty compact subsets of $\mathcal{K}([0,1]^n)$ with the Hausdorff metric.

Step 1: The map $\Psi : \mathcal{D}_2 \ni x_0 \mapsto u(x_0) \in \mathcal{K}(\mathcal{K}([0,1]^n))$ is Hausdorff continuous.

Take a sequence $x_0^n \to x_0$ with $x_0^n, x_0 \in \mathcal{D}_2$. Since $\mathcal{K}(\mathcal{K}([0,1]^n))$ is compact, we can assume that $\{\Psi(x_0^n)\}_{n=1}^\infty$ converges to the limit $w_0 \in \mathcal{K}(\mathcal{K}([0,1]^n))$. We want to show $\Psi(x_0) = w_0$.

To show $\Psi(x_0) \subset w_0$, take any $u(\bar{x}_1) \in \Psi(x_0)$, where $\bar{x}_1 \in x_0$. Since $x_0^n \to x_0$, we can find $\{x_1^n\}_{n=1}^\infty$ such that $x_1^n \to \bar{x}_1$ with $x_1^n \in x_0^n$. Hausdorff continuity of u implies $u(x_1^n) \to u(\bar{x}_1)$. Since $u(x_1^n) \to u(\bar{x}_1)$ and $\Psi(x_0^n) \to w_0$ with $u(x_1^n) \in \Psi(x_0^n)$, we have $u(\bar{x}_1) \in w_0$. Hence, $\Psi(x_0) \subset w_0$.

For the other direction, take any $w_1 \in w_0$. Since $\Psi(x_0^n) \to w_0$, we can find $\{x_1^n\}_{n=1}^\infty$ such that $u(x_1^n) \to w_1$ with $x_1^n \in x_0^n$. Since $\mathcal{K}^*(\mathcal{H})$ is compact, assume $x_1^n \to \bar{x}_1$ for some $\bar{x}_1 \in \mathcal{K}^*(\mathcal{H})$. The conditions, $x_1^n \to \bar{x}_1$ and $x_0^n \to x_0$ with $x_1^n \in x_0^n$, imply $\bar{x}_1 \in x_0$. Hausdorff continuity of u implies $w_1 = u(\bar{x}_1) \in \Psi(x_0)$, and hence $w_0 \subset \Psi(x_0)$.

Step 2: For any $x_0, y_0 \in \mathcal{D}_2$, $d_{\mathrm{supnorm}}(\sigma_{x_0}, \sigma_{y_0}) \leq d_{\mathrm{Hausdorff}}(u(x_0), u(y_0))$.

For any fixed $\mu \in S_1$,

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

where $x_1^* \in x_0$ and $y_1^* \in y_0$ are maximizers. Let $\bar{y}_1$ be a minimizer of the following problem:

$$\min_{y_1 \in A} d_{\text{supnorm}}(\zeta_{y_1}, \zeta_{x_1^*}), \text{ where}$$
$$A \equiv \left\{ y_1 \in \mathcal{K}^*(\mathcal{H}) \Big| \int_{S_2} \zeta_{y_1} \, \mathrm{d}\mu \leq \int_{S_2} \zeta_{y_1^*} \, \mathrm{d}\mu \right\}.$$

Notice $y_0 \subset A$. Since ζ is continuous, A is compact, and hence the minimizer indeed exists. Since

$$d_{\text{supnorm}}(\zeta_{x_1}, \zeta_{y_1}) \leq d_{\text{Hausdorff}}(u(x_1), u(y_1))$$

holds for all $x_1, y_1 \in \mathcal{K}^*(\mathcal{H})$ by Step 2 in Lemma B.2 (i) of Takeoka [20],

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

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

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

From Step 1 and Step 2, σ is continuous.

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

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

(iii) Let $x_0 \neq x_0'$. Then there exists $\bar{x}_1 \in x_0' \setminus x_0$.[9] Since $\zeta : \mathcal{K}^*(\mathcal{H}) \to \mathcal{C}(S_2)$ is injective, $\zeta_{\bar{x}_1} \in \zeta(x_0') \setminus \zeta(x_0)$, where $\zeta(x_0)$ and $\zeta(x_0')$ are the images of x_0 and of x_0' under ζ, respectively.

Step 1: $\zeta(x_0) \cap (\mathcal{C}_+(S_2) + \{\zeta_{\bar{x}_1}\}) = \emptyset$, where $\mathcal{C}_+(S_2) \subset \mathcal{C}(S_2)$ is the set of non-negative continuous functions on S_2.

[9] The symmetric argument works when $\bar{x}_1 \in x_0 \setminus x_0'$.

Suppose otherwise. There exist $\hat{x}_1 \in x_0$ and $f \in \mathcal{C}_+(S_2)$ such that $\zeta_{\hat{x}_1} = f + \zeta_{\bar{x}_1}$. Since $f \geq 0$, $\zeta_{\hat{x}_1}(p) \geq \zeta_{\bar{x}_1}(p)$ for all $p \in S_2$. Moreover, since $\hat{x}_1 \neq \bar{x}_1$, f is non-zero, and hence there exists some $\underline{p} \in S_2$ satisfying $\zeta_{\hat{x}_1}(\underline{p}) > \zeta_{\bar{x}_1}(\underline{p})$. Thus, we have $\bar{x}_1 \subset \hat{x}_1$. From the property of $\mathcal{D}_2$, $\bar{x}_1 \in x_0$. 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 $\Lambda(f) > c > \Lambda(f')$ for all $f \in \mathcal{C}_+(S_2) + \{\zeta_{\bar{x}_1}\}$ and $f' \in \zeta(x_0)$.

Since ζ is continuous and mixture linear and since x_0 is compact and convex, $\zeta(x_0)$ is a compact convex subset of $\mathcal{C}(S_2)$. Furthermore, since $\mathcal{C}_+(S_2)$ is closed and convex, so is $\mathcal{C}_+(S_2)+\{\zeta_{\bar{x}_1}\}$. From Step 1, we know that $\zeta(x_0)$ and $\mathcal{C}_+(S_2) + \{\zeta_{\bar{x}_1}\}$ are disjoint. The separation hyperplane theorem (Dunford and Schwartz [3, Theorem 10, p.417]) ensures that there exist a linear functional Λ on $\mathcal{C}(S_2)$ and a constant $c \in \mathbb{R}$ such that $\Lambda(f) > c > \Lambda(f')$ for all $f \in \mathcal{C}_+(S_2) + \{\zeta_{\bar{x}_1}\}$ and $f' \in \zeta(x_0)$.

Since $0 \in \mathcal{C}_+(S_2)$, we know that $\Lambda(\zeta_{\bar{x}_1}) > c > \Lambda(\zeta_{x_1})$ for all $\zeta_{x_1} \in \zeta(x_0)$. Thus,

$$\Lambda(\zeta_{\bar{x}_1}) > \max_{x_1 \in x_0} \Lambda(\zeta_{x_1}). \tag{25}$$

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

From Step 2, $\Lambda(f_+) > \Lambda(\zeta_{x_1} - \zeta_{\bar{x}_1})$ for all $f_+ \in \mathcal{C}_+(S_2)$ and $\zeta_{x_1} \in \zeta(x_0)$. 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 $f_+ \in \mathcal{C}_+(S_2)$.

Since Λ is a positive linear functional on $\mathcal{C}(S_2)$, the Riesz Representation theorem (Rudin [15, Theorem 2.14, p. 40]) ensures the existence of a positive measure $\bar{\nu}$ on S_2 satisfying

$$\Lambda(f) = \int_{S_2} f(p)\,\mathrm{d}\bar{\nu}(p), \tag{26}$$

for all $f \in \mathcal{C}(S_2)$. Let $\bar{\mu} \in \Delta(S_2)$ be the normalization of $\bar{\nu}$. Taking (25) and (26) together, we have

$$\sigma_{x_0'}(\bar{\mu}) = \max_{x_1 \in x_0'} U_1(x_1, \bar{\mu}) \geq \int_{S_2} \zeta_{\bar{x}_1}\,\mathrm{d}\bar{\mu} > \max_{x_1 \in x_0} \int_{S_2} \zeta_{x_1}\,\mathrm{d}\bar{\mu} = \sigma_{x_0}(\bar{\mu}).$$

Since $\sigma_{x_0'} \neq \sigma_{x_0}$, we have shown that σ is injective. □

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

Lemma B.6.

(i) *C is convex.*

(ii) *The constant function equal to zero belongs to C.*

(iii) *The constant function equal to one belongs to C.*

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

(v) $f \geq 0$ *for all* $f \in C$.

Proof. (i) Take any $f, f' \in C$ and $\lambda \in [0,1]$. There exist $x_0', x_0 \in \mathcal{D}_2$ satisfying $f' = \sigma_{x_0'}$ and $f = \sigma_{x_0}$. From Lemma B.5 (ii),

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

Since $\mathrm{CO}(\lambda x_0' + (1-\lambda)x_0) \in \mathcal{D}_2$, C is convex.

(ii) Let $x_0 \equiv \mathrm{CO}(\{\{\underline{l}\}\}) \in \mathcal{D}_2$. Since $u(\underline{l}) = 0$, $\sigma_{x_0}(\mu) = 0$ for all $\mu \in S_1$.

(iii) Let $x_0 \equiv \mathrm{CO}(\{\{\bar{l}\}\}) \in \mathcal{D}_2$. Since $u(\bar{l}) = 1$, $\sigma_{x_0}(\mu) = 1$ 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 $x_0'' \equiv \mathrm{CO}(\overline{\mathrm{co}}(x_0 \cup x_0')) \in \mathcal{D}_2$ and $f'' \equiv \sigma_{x_0''} \in C$. Then, $f''(\mu) = \max[\sigma_{x_0}(\mu), \sigma_{x_0'}(\mu)]$.

(v) Since each $x_1 \in x_0 \in \mathcal{D}_2$ contains the constant act $\underline{l}$,

$$\zeta_{x_1}(p) = \max_{h \in x_1} \left(\sum_{\omega \in \Omega} u(h(\omega)) p(\omega) \right) \geq \sum_{\omega \in \Omega} u(\underline{l}) p(\omega) = 0,$$

for any $p \in S_2$. Hence, for any $\mu \in S_1$, $\sigma_{x_0}(\mu) \geq 0$. $\square$

Since σ is injective, we can define $W : C \to \mathbb{R}$ by $W(f) \equiv U(\sigma^{-1}(f))$. Notice that $W(\mathbf{0}) = 0$ and $W(\mathbf{1}) = 1$, where $\mathbf{0}$ and $\mathbf{1}$ are identified with the zero function and with the unit function, respectively. Since U and σ are continuous, so is W with respect to the sup-norm. Furthermore, $W : C \to \mathbb{R}$ is mixture linear. Indeed, take any $f, f' \in C$ and $\lambda \in [0,1]$. There exist $x_0, x_0' \in \mathcal{D}_2$ such that $f = \sigma_{x_0}$ and $f' = \sigma_{x_0'}$. Since U is mixture linear,

$$\begin{aligned}
W(\lambda f' + (1-\lambda) f) &= W(\lambda \sigma_{x_0'} + (1-\lambda)\sigma_{x_0}) \\
&= U(\sigma^{-1}(\lambda \sigma_{x_0'} + (1-\lambda)\sigma_{x_0})) \\
&= U(\sigma^{-1}(\sigma_{\mathrm{CO}(\lambda x_0' + (1-\lambda)x_0)})) \\
&= U(\mathrm{CO}(\lambda x_0' + (1-\lambda)x_0)) \\
&= U(\lambda x_0' + (1-\lambda)x_0) \\
&= \lambda U(x_0') + (1-\lambda) U(x_0) \\
&= \lambda W(f') + (1-\lambda) W(f).
\end{aligned}$$

By adapting Lemma 10 (p. 928) of DLR, we can show that W is linear in the sense that $W(\alpha f + \beta f') = \alpha W(f) + \beta W(f')$ for any $f, f', \alpha f + \beta f' \in C$, where $\alpha, \beta \in \mathbb{R}_+$.

By the same argument as in DLR, we will extend W to $\mathcal{C}(S_1)$ step by step. For any $r \geq 0$, let $rC \equiv \{rf | f \in C\}$. Let $H \equiv \cup_{r \geq 0} rC$. As the first step, we will extend $W : C \to \mathbb{R}$ to H. For any $f \in H \setminus \{\mathbf{0}\}$, there exists $r > 0$ satisfying $(1/r)f \in C$. Define $W(f) \equiv rW((1/r)f)$. As shown in DLR, $W : H \to \mathbb{R}$ is well-defined, monotonic, and linear.

Let

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

Since $\mathbf{0} \in H$, $H \subset H^*$. For any $f \in H^*$, there exist $f_1, f_2 \in H$ satisfying $f = f_1 - f_2$. Define $W(f) \equiv W(f_1) - W(f_2)$. As in DLR, we can show that $W : H^* \to \mathbb{R}$ is well-defined and linear.

Finally, we will extend W to $\mathcal{C}(S_1)$.

Lemma B.7. *H^* is dense in $\mathcal{C}(S_1)$.*

Proof. By the Stone-Weierstrass theorem (Schaefer [18, Theorem 8.1, p. 243]), 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 $f(\mu) \neq f(\mu')$; and (iii) H^* contains the constant function equal to one. By the same argument as in Lemma 11 (p. 928) of DLR, condition (i) holds. Condition (iii) directly follows from Lemma B.6 (iii) and the definition of H^*.

To show condition (ii), take any 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 [3, 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 B.5 in Takeoka [20, p. 25], there exist $x_1, y_1 \in \mathcal{K}^*(\mathcal{H})$ 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,

$$\int_{S_2} \zeta_{x_1} \, \mathrm{d}\mu > \int_{S_2} \zeta_{y_1} \, \mathrm{d}\mu, \text{ and } \int_{S_2} \zeta_{y_1} \, \mathrm{d}\mu' > \int_{S_2} \zeta_{x_1} \, \mathrm{d}\mu'. \tag{27}$$

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 $\{h \in \mathcal{H} | d(x_1, h) \leq \varepsilon\}$ for some small $\varepsilon > 0$. Then,

$$\int_{S_2} \zeta_{x_1} \, \mathrm{d}\mu > \int_{S_2} \zeta_{y_1} \, \mathrm{d}\mu'. \tag{28}$$

Moreover, as long as $\varepsilon > 0$ is small enough, (27) still holds after this modification. Let $x_0 \equiv \mathrm{CO}(\overline{\mathrm{co}}(\{x_1, y_1\}))$. Taking (27) and (28) 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 $\sigma_{x_0} \in C \subset H^*$, condition (ii) holds. □

Since $\mathcal{D}_2$ is compact, the same argument as in Lemma 12 (p. 929) of DLR implies that there exists a constant $K > 0$ such that $W(f) \leq K\|f\|$ for any $f \in H^*$. By the Hahn-Banach theorem, we can extend W to $\overline{W} : \mathcal{C}(S_1) \to \mathbb{R}$ so that $\overline{W}$ is linear, continuous and increasing. Lemma B.7 ensures uniqueness of this extension. Now we have the following commutative diagram:

$$\begin{tikzcd} \mathcal{D}_2 \arrow[r, "U"] \arrow[d, "\sigma"'] & \mathbb{R} \\ \mathcal{C}(S_1) \arrow[ur, "\overline{W}"'] & \end{tikzcd}$$

Since $\overline{W}$ is a positive linear functional on $\mathcal{C}(S_1)$, the Riesz representation theorem (Dunford and Schwartz [3, Theorem 3, p. 265]) ensures that there exists 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 $x_0 \in \mathcal{D}_2$,

$$U(x_0) = \overline{W}(\sigma(x_0)) = \int_{S_1} \sigma(x_0) \, \mathrm{d}\mu_0(\mu) = \int_{S_1} \max_{x_1 \in x_0} U_1(x_1, \mu) \, \mathrm{d}\mu_0(\mu). \tag{29}$$

B.3 Proof of Theorem 4.1

(i) Since u and u' are mixture linear representations of the same preference on $\Delta(Z)$, they are cardinally equivalent by the standard argument.

(ii) As shown above, u and u' are cardinally equivalent. Thus, (μ_0', u) also represents the same preference. Let U_0 and U_0' be the canonical representations associated with (μ_0, u) and with (μ_0', u), respectively. For all $x_0 \in \mathcal{D}$ and $\mu \in \Delta(\Delta(\Omega))$, define

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

where

$$U_1(x_1, \mu) \equiv \int_{\Delta(\Omega)} \max_{h \in x_1} \left(\sum_{\omega} u(h(\omega)) p(\omega) \right) \mathrm{d}\mu(p).$$

Then,

$$U_0(x_0) = \int \sigma_{x_0}(\mu) \, \mathrm{d}\mu_0(\mu), \text{ and } U_0'(x_0) = \int \sigma_{x_0}(\mu) \, \mathrm{d}\mu_0'(\mu). \tag{30}$$

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

$$\begin{aligned} U_0'(\{\{l\}\}) &= \alpha U_0(\{\{l\}\}) + \beta \\ u(l) &= \alpha u(l) + \beta. \end{aligned}$$

We must have $\alpha = 1$ and $\beta = 0$, and hence $U_0 = U_0'$. From (30), for all x_0,

$$\int \sigma_{x_0}(\mu)\,\mathrm{d}\mu_0(\mu) = \int \sigma_{x_0}(\mu)\,\mathrm{d}\mu_0'(\mu). \tag{31}$$

Take any $x_0, y_0 \in \mathcal{D}$ and $\alpha, \beta \geq 0$. Equation (31) holds even when σ_{x_0} is replaced with $\alpha\sigma_{x_0} - \beta\sigma_{y_0}$. From Lemma B.7, the set of all such functions is a dense subset of the set of all real-valued continuous functions on $\Delta(\Delta(\Omega))$. Hence, equation (31) still holds even if σ_{x_0} is replaced with any real-valued continuous function. The Riesz representation theorem (Dunford and Schwartz [3, Theorem 3, p. 265]) implies $\mu_0 = \mu_0'$.

B.4 Proof of Theorem 4.2

By the same argument as in Theorem 4.1 (i), u^1 and u^2 are cardinally equivalent. We show the second result. Let $(\mathcal{S}^i, \mu_0^{i*})$ be the subjective decision tree derived from a second-order additive SEU representation $(\{S_t^i\}_{t=1}^2, \{\mu_t^i\}_{t=0}^2, u^i)$ for $i = 1, 2$.

For any $\eta \in \Delta(\Delta(\Delta(\Omega)))$, let $E\eta \in \Delta(\Delta(\Omega))$ denote the "mean" of η. Formally, $E\eta$ is defined as the unique probability measure satisfying

$$\int_{\Delta(\Omega)} f(p)\,\mathrm{d}(E\eta)(p) = \int_{\Delta(\Delta(\Omega))}\int_{\Delta(\Omega)} f(p)\,\mathrm{d}\mu(p)\,\mathrm{d}\eta(\mu), \tag{32}$$

for all real-valued continuous functions f on $\Delta(\Omega)$. Indeed, $E\eta$ is well-defined by the Riesz representation theorem.

Lemma B.8. $\mathcal{S}^i = \mathrm{supp}(E\mu_0^{i*})$.

Proof. Since $\mathcal{S}^i = \mathrm{supp}(P^i \circ (\mu_2^i)^{-1})$, it suffices to show that $P^i \circ (\mu_2^i)^{-1} = E\mu_0^{i*}$. For any real-valued continuous function f on $\Delta(\Omega)$,

$$\begin{aligned}
\int_{\Delta(\Omega)} f(p)\,\mathrm{d}P^i \circ (\mu_2^i)^{-1}(p) &= \int_{S_1^i \times S_2^i} f(\mu_2^i(s_1, s_2))\,\mathrm{d}P^i(s_1, s_2) \\
&= \int_{S_1^i}\int_{S_2^i} f(\mu_2^i(s_1, s_2))\,\mathrm{d}\mu_1^i(s_2|s_1)\,\mathrm{d}\mu_0^i(s_1) \\
&= \int_{S_1^i}\int_{\Delta(\Omega)} f(p)\,\mathrm{d}\left(\mu_1^i(\cdot|s_1) \circ (\mu_2^i(s_1, \cdot))^{-1}\right)\,\mathrm{d}\mu_0^i(s_1) \\
&= \int_{\Delta(\Delta(\Omega))}\int_{\Delta(\Omega)} f(p)\,\mathrm{d}\mu(p)\,\mathrm{d}\mu_0^{i*}(\mu).
\end{aligned}$$

From uniqueness of $E\mu_0^{i*}$, $P^i \circ (\mu_2^i)^{-1} = E\mu_0^{i*}$. $\square$

From Lemma B.8, if $\mu_0^{1*} = \mu_0^{2*}$, then $\mathcal{S}^1 = \mathcal{S}^2$. Hence, it is enough to see that $\mu_0^{1*} = \mu_0^{2*}$. From Proposition 3.1 (ii), the original representation $(\{S_t^i\}_{t=1}^2, \{\mu_t^i\}_{t=0}^2, u^i)$ and its canonical form (μ_0^{i*}, u^i) represent the same preference. Moreover, by assumption, $(\{S_t^i\}_{t=1}^2, \{\mu_t^i\}_{t=0}^2, u)$, $i = 1, 2$, represent the same preference. Thus, (μ_0^{i*}, u^i), $i = 1, 2$, also represent the same preference. It follows from Theorem 4.1 (ii) that $\mu_0^{1*} = \mu_0^{2*}$.

B.5 Proof of Corollary 5.1

$((b) \Rightarrow (a))$ Assume that there exists $\mu \in \Delta(\Delta(\Omega))$ such that $U_0(x_0) = \max_{x_1 \in x_0} U_1(x_1, \mu)$ represents preference. We will show that $\succeq$ satisfies Strategic Rationality.

Let $x_0 \succeq y_0$. The representation implies that

$$\max_{x_1 \in x_0} U_1(x_1, \mu) \geq \max_{x_1 \in y_0} U_1(x_1, \mu). \tag{33}$$

Since $x_0 \subset x_0 \cup y_0$, we have $\max_{x_1 \in x_0 \cup y_0} U_1(x_1, \mu) \geq \max_{x_1 \in x_0} U_1(x_1, \mu)$. Suppose that this weak inequality holds with strict inequality. Then there exists $y_1^* \in y_0$ such that $U_1(y_1^*, \mu) > \max_{x_1 \in x_0} U_1(x_1, \mu)$. This contradicts (33). Thus, $\max_{x_1 \in x_0 \cup y_0} U_1(x_1, \mu) = \max_{x_1 \in x_0} U_1(x_1, \mu)$, and hence $x_0 \cup y_0 \sim x_0$.

$((a) \Rightarrow (b))$ From Theorem 3.1, there exists a canonical representation (μ_0, u). Suppose $\#\text{supp}(\mu_0) \neq 1$. There exist $\mu, \mu' \in \text{supp}(\mu_0)$ with $\mu \neq \mu'$. As in Lemma B.7, we can find $\bar{x}_1, \bar{x}'_1 \in \mathcal{K}(\mathcal{H})$ such that

$$U_1(\bar{x}_1, \mu) > U_1(\bar{x}'_1, \mu) \text{ and } U_1(\bar{x}'_1, \mu') > U_1(\bar{x}_1, \mu').$$

(See condition (27).) From Continuity, there exists $\varepsilon > 0$ such that $\bar{x}_1(\varepsilon) \equiv \{h | d(h, \bar{x}_1) \leq \varepsilon\}$ and $\bar{x}'_1(\varepsilon) \equiv \{h | d(h, \bar{x}'_1) \leq \varepsilon\}$ satisfy

$$\begin{aligned} U_1(\bar{x}_1(\varepsilon), \mu) > U_1(\bar{x}_1, \mu) > U_1(\bar{x}'_1(\varepsilon), \mu) > U_1(\bar{x}'_1, \mu), \text{ and} \\ U_1(\bar{x}'_1(\varepsilon), \mu') > U_1(\bar{x}'_1, \mu') > U_1(\bar{x}_1(\varepsilon), \mu') > U_1(\bar{x}_1, \mu'). \end{aligned} \tag{34}$$

Let

$$\begin{aligned} x_0 &\equiv \{x_1 | U_1(x_1, \mu) \leq U_1(\bar{x}_1(\varepsilon), \mu)\} \cap \{x_1 | U_1(x_1, \mu') \leq U_1(\bar{x}'_1, \mu')\}, \\ x'_0 &\equiv \{x_1 | U_1(x_1, \mu) \leq U_1(\bar{x}_1, \mu)\} \cap \{x_1 | U_1(x_1, \mu') \leq U_1(\bar{x}'_1(\varepsilon), \mu')\}. \end{aligned}$$

From (34), the representation implies

$$x_0 \cup x'_0 \succ x_0 \text{ and } x_0 \cup x'_0 \succ x'_0. \tag{35}$$

Since Order implies $x_0 \succeq x'_0$ or $x'_0 \succeq x_0$, we have, by Strategic Rationality,

$$x_0 \cup x'_0 \sim x_0 \text{ or } x_0 \cup x'_0 \sim x'_0.$$

This contradicts (35).

B.6 Proof of Corollary 5.2

$((b) \Rightarrow (a))$ We will show that $\succeq$ satisfies Neutrality to Commitment. Take any x'_0 and finite x_0. Let $\bar{x}_1 \equiv \cup_{x_1 \in x_0} x_1$.

Take any $p \in \Delta(\Omega)$. Since $x_1 \subset \bar{x}_1$ for any $x_1 \in x_0$, we have $\max_{h \in \bar{x}_1} U_2(h, p) \geq \max_{h \in x_1} U_2(h, p)$ for all $x_1 \in x_0$. Thus,

$$\max_{h \in \bar{x}_1} U_2(h, p) \geq \max_{x_1 \in x_0} \max_{h \in x_1} U_2(h, p). \tag{36}$$

Suppose that this weak inequality holds with strict inequality. Then there exists $h^* \in \bar{x}_1$ such that $U_2(h^*, p) > \max_{x_1 \in x_0} \max_{h \in x_1} U_2(h, p) \geq U_2(h^*, p)$. This is a contradiction. Hence, (36) holds with equality. For all p, we have

$$\max_{x_1 \in x_0' \cup \{\bar{x}_1\}} \max_{h \in x_1} U_2(h, p) = \max_{x_1 \in x_0' \cup x_0} \max_{h \in x_1} U_2(h, p). \tag{37}$$

Let δ_p denote the degenerate probability measure at p, which assigns p to probability one. Then, (37) implies

$$\begin{aligned} U_0(x_0' \cup \{\bar{x}_1\}) &= \int \max_{x_1 \in x_0' \cup \{\bar{x}_1\}} U_1(x_1, \delta_p) \, \mathrm{d}\mu_0(\delta_p) \\ &= \int \max_{x_1 \in x_0' \cup \{\bar{x}_1\}} \max_{h \in x_1} U_2(h, p) \, \mathrm{d}\mu_0(\delta_p) \\ &= \int \max_{x_1 \in x_0' \cup x_0} \max_{h \in x_1} U_2(h, p) \, \mathrm{d}\mu_0(\delta_p) \\ &= \int \max_{x_1 \in x_0' \cup x_0} U_1(x_1, \delta_p) \, \mathrm{d}\mu_0(\delta_p) \\ &= U_0(x_0' \cup x_0). \end{aligned}$$

($(a) \Rightarrow (b)$) From Theorem 3.1, there exists a canonical representation (μ_0, u).

Step 1: If x_0 is finite, $x_0 \sim \{\cup_{x_1 \in x_0} x_1\}$.

Neutrality to Commitment implies

$$x_0 = x_0 \cup x_0 \sim x_0 \cup \{\cup_{x_1 \in x_0} x_1\} \sim \{\cup_{x_1 \in x_0} x_1\} \cup \{\cup_{x_1 \in x_0} x_1\} = \{\cup_{x_1 \in x_0} x_1\}.$$

Step 2: For all x_0, $x_0 \sim \{\mathrm{cl}(\cup_{x_1 \in x_0} x_1)\}$, where $\mathrm{cl}(x_1)$ is the closure of x_1.

Since $\{\mathrm{cl}(\cup_{x_1 \in x_0} x_1)\}$ covers x_0, Lemma B.1 implies $\{\mathrm{cl}(\cup_{x_1 \in x_0} x_1)\} \succeq x_0$. Suppose $\{\mathrm{cl}(\cup_{x_1 \in x_0} x_1)\} \succ x_0$. From Continuity and the property of the Hausdorff metric, there exists a finite subset $y_1 \subset \mathrm{cl}(\cup_{x_1 \in x_0} x_1)$ such that $\{y_1\} \succ x_0$. Since any $h \in y_1$ is an accumulation point of $\cup_{x_1 \in x_0} x_1$, there exists $h' \in \cup_{x_1 \in x_0} x_1$ close to h in the sense of the metric on $\mathcal{H}$. Thus, by Continuity, we can assume that $y_1 \subset \cup_{x_1 \in x_0} x_1$. Denote y_1 by $\{h^i | i = 1, \cdots, I\}$. For all i, there exists $x_1^i \in x_0$ such that $h^i \in x_1^i$. Let $x_0^* = \{x_1^i | i = 1, \cdots, I\}$. Step 1 and Lemma B.1 imply that

$$x_0^* \sim \{\cup_{x_1 \in x_0^*} x_1\} \succeq \{y_1\} \succ x_0. \tag{38}$$

Since $x_0^* \subset x_0$, (38) violates Monotonicity.

From Step 2, for any x_0,

$$\begin{aligned} U_0(x_0) &= U_0(\{\mathrm{cl}(\cup_{x_1 \in x_0} x_1)\}) \\ &= \int_{\Delta(\Delta(\Omega))} \int_{\Delta(\Omega)} \max_{h \in \mathrm{cl}(\cup_{x_1 \in x_0} x_1)} U_2(h, p) \, \mathrm{d}\mu(p) \, \mathrm{d}\mu_0(\mu) \\ &= \int_{\Delta(\Omega)} \max_{h \in \mathrm{cl}(\cup_{x_1 \in x_0} x_1)} U_2(h, p) \, \mathrm{d}\bar{\mu}(p), \end{aligned}$$

where $\bar{\mu}$ over $\Delta(\Omega)$ is the "mean" probability with respect to μ_0. (See (32) for details.) Redefine $\mu_0 \in \Delta(\Delta(\Delta(\Omega)))$ by $\mu_0(\delta_p) \equiv \bar{\mu}(p)$. For each $p \in \mathrm{supp}(\bar{\mu})$, let $U_1(x_1, \delta_p) \equiv \max_{h \in x_1} U_2(h, p)$. Then,

$$\begin{aligned} U_0(x_0) &= \int_{\Delta(\Omega)} \sup_{h \in (\cup_{x_1 \in x_0} x_1)} U_2(h, p) \, \mathrm{d}\bar{\mu}(p) \\ &= \int_{\Delta(\Omega)} \max_{x_1 \in x_0} \max_{h \in x_1} U_2(h, p) \, \mathrm{d}\bar{\mu}(p) \\ &= \int_{\Delta(\Delta(\Omega))} \max_{x_1 \in x_0} U_1(x_1, \delta_p) \, \mathrm{d}\mu_0(\delta_p). \end{aligned}$$

This is the required result.

B.7 Proof of Theorem 6.1

First of all, we prepare the following lemma:

Lemma B.9. *For any finite subset $P \equiv \{p^1, \cdots, p^i, \cdots, p^I\} \subset \Delta(\Omega)$, there exist positive numbers $v^1, \cdots, v^I$ such that*

$$x_1 \equiv \bigcap_{i=1}^{I} \left\{ h \in \mathcal{H} \,\middle|\, U_2(h, p^i) \leq v^i \right\} \tag{39}$$

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

Proof. Since $u : \Delta(Z) \to \mathbb{R}$ is continuous, there exist a maximal and a minimal lottery, $\bar{l}$ and $\underline{l}$, with respect to u. Since u is non-constant and mixture linear, without loss of generality, we can assume $u(\bar{l}) = 1$ and $u(\underline{l}) = 0$. Consequently, u is regarded as a mixture linear function from $\Delta(Z)$ into $[0, 1]$.

Consider the set

$$W \equiv \left\{ w \in [0,1]^n \,\middle|\, \|w\| \leq \frac{1}{2} \right\},$$

where $\|\cdot\|$ is the square norm of $\mathbb{R}^n$. For any $p^i \in P$, there exists a unique point $w^i \in W$ such that $w^i \cdot p^i > w \cdot p^i$ for all $w \in W$ with $w \neq w^i$. That is, $w \cdot p^i = w^i \cdot p^i$ is the supporting hyperplane of W at w^i. Now let $v^i \equiv w^i \cdot p^i$ and define

$$x_1 \equiv \bigcap_{i=1}^{I} \left\{ h \in \mathcal{H} \,\middle|\, U_2(h, p^i) \leq v^i \right\}.$$

Since $U_2(\underline{l}, p^i) \leq v^i$ for all i, x_1 is non-empty. Since $U_2(\cdot, p^i)$ is continuous and mixture linear, x_1 is a compact and convex menu. Finally, notice that there exists an act h^i such that $u(h^i) = w^i$. Since $w^i \cdot p^i > w^j \cdot p^i$ for all i and $j \neq i$, we have $U_2(h^i, p^i) > U_2(h^j, p^i)$ for all i and $j \neq i$. Thus, the last property holds. $\square$

Lemma B.10. *For any finite subset* $\{\mu^1, \cdots, \mu^j, \cdots, \mu^J\} \subset \Delta(\Delta(\Omega))$ *with finite* $\mathrm{supp}(\mu^j)$*, there exist* $x_1^1, \cdots, x_1^j, \cdots, x_1^J \in \mathcal{K}(\mathcal{H})$ *such that, for any* j*,* $U_1(x_1^j, \mu^j) > U_1(x_1^k, \mu^j)$ *for all* $k \neq j$*.*

Proof. Let

$$S_2^* \equiv \cup_{j=1}^J \mathrm{supp}(\mu^j) \subset \Delta(\Omega).$$

Let $I \equiv \# S_2^*$. Since S_2^* is a finite set, S_2^* can be denoted by $\{p^1, \cdots, p^i, \cdots, p^I\}$.

Each μ^j is regarded as an element of the $(I-1)$-dimensional unit simplex. We can find $\{a^1, \cdots, a^J\} \subset \mathbb{R}^I$ such that $\mu^j a^j > \mu^j a^k$ for any j and $k \neq j$. Without loss of generality, assume that the absolute value of the i-th coordinate of a^j is sufficiently small.

From Lemma B.9, there exist positive numbers $v^1, \cdots, v^I$ such that the menu $x_1 \in \mathcal{K}(\mathcal{H})$, defined by

$$x_1 \equiv \bigcap_{p^i \in S_2^*} \left\{ h \in \mathcal{H} \,\middle|\, U_2(h, p^i) \leq v^i \right\},$$

satisfies all the properties in the lemma. For each j, define

$$x_1^j \equiv \bigcap_{p^i \in S_2^*} \left\{ h \in \mathcal{H} \,\middle|\, U_2(h, p^i) \leq v^i + a_i^j \right\} \in \mathcal{K}(\mathcal{H}).$$

Since each a_i^j is sufficiently small, the boundary of each lower contour set partly coincides with a non-trivial part of the boundary of x_1^j. Thus, for any $k \neq j$,

$$\begin{aligned}
U_1(x_1^j, \mu^j) &= \sum_i \max_{h \in x_1^j} U_2(h, p^i) \mu^j(p^i) \\
&= \sum_i \left(v^i + a_i^j \right) \mu^j(p^i) \\
&= \sum_i v^i \mu^j(p^i) + \sum_i a_i^j \mu^j(p^i) \\
&> \sum_i v^i \mu^j(p^i) + \sum_p a_i^k \mu^j(p^i) \\
&= U_1(x_1^k, \mu^j).
\end{aligned}$$

□

(i) (only if part) Suppose $S_1^1 \not\subset S_1^2$. Then there exists $\mu^0 \in S_1^1 \setminus S_1^2$. Denote S_1^2 by $\{\mu^j | j = 1, \cdots, J\}$. Apply Lemma B.10 to the set $S_1^2 \cup \{\mu^0\}$. Then there exist menus $\{x_1^j\}_{j=0}^J$ such that, for any $j = 0, \cdots, J$, $U_1(x_1^j, \mu^j) > U_1(x_1^k, \mu^j)$ for all $k \neq j$. Define

$$x_0 \equiv \bigcap_{j=0}^J \left\{ x_1 \in \mathcal{K}(\mathcal{H}) \,\middle|\, U_1(x_1, \mu^j) \leq U_1(x_1^j, \mu^j) \right\}.$$

Then x_0 is a non-empty compact convex menu of menus, and the boundary of each lower contour set partly coincides with a non-trivial part of the boundary of x_0. Let

$$\hat{x}_0 \equiv \bigcap_{j=1}^{J} \left\{ x_1 \in \mathcal{K}(\mathcal{H}) \,\middle|\, U_1(x_1, \mu^j) \leq U_1(x_1^j, \mu^j) \right\} \in \mathcal{D}.$$

Since $x_0 \subset \hat{x}_0$, for all $\mu \in S_1^1$,

$$\max_{x_1 \in \hat{x}_0} U_1(x_1, \mu) \geq \max_{x_1 \in x_0} U_1(x_1, \mu),$$

and this weak inequality holds with strict inequality for μ^0. The associated representation implies $\hat{x}_0 \succ^1 x_0$. On the other hand, for all $\mu \in S_1^2$,

$$\max_{x_1 \in \hat{x}_0} U_1(x_1, \mu) = \max_{x_1 \in x_0} U_1(x_1, \mu),$$

and hence the associated representation implies $\hat{x}_0 \sim^2 x_0$. This is a contradiction.

(if part) Assume $S_1^1 \subset S_1^2$. For any x_0, y_0 with $y_0 \subset x_0$, let $x_0 \succ^1 y_0$. Since $y_0 \subset x_0$, for all $\mu \in S_1^1$,

$$\max_{x_1 \in x_0} U_1(x_1, \mu) \geq \max_{x_1 \in y_0} U_1(x_1, \mu). \tag{40}$$

Furthermore, since $x_0 \succ^1 y_0$, there must exist at least one signal $\mu^* \in S_1^1$ such that (40) holds with strict inequality. Thus, the assumption $S_1^1 \subset S_1^2$ implies $x_0 \succ^2 y_0$.

(ii) (only if part) Suppose $\mathcal{S}^1 \not\subset \mathcal{S}^2$. Then there exists $p^0 \in \mathcal{S}^1 \setminus \mathcal{S}^2$. Denote $\mathcal{S}^2$ by $\{p^i | i = 1, \cdots, I\}$. By applying Lemma B.9 to $\{p^0\} \cup \mathcal{S}^2$, there exist positive numbers $\{v^i\}_{i=0}^I$ such that the menu x_1 defined as in Lemma B.9 satisfies all the properties in the lemma. Let

$$\hat{x}_1 \equiv \bigcap_{i=1}^{I} \left\{ h \in \mathcal{H} \,\middle|\, U_2(h, p) \leq v^i \right\}.$$

Since $x_1 \subset \hat{x}_1$, for all $p \in \mathcal{S}^1$,

$$\max_{h \in \hat{x}_1} U_2(h, p) \geq \max_{h \in x_1} U_2(h, p),$$

and this weak inequality holds with strict inequality for p^0. Thus, the associated representation implies $\{\hat{x}_1\} \succ^1 \{x_1\}$. However, by construction, for all $p \in \mathcal{S}^2$,

$$\max_{h \in \hat{x}_1} U_2(h, p) = \max_{h \in x_1} U_2(h, p).$$

Hence, the associated representation implies $\{\hat{x}_1\} \sim^2 \{x_1\}$. This is a contradiction.

(if part) Assume $\mathcal{S}^1 \subset \mathcal{S}^2$. For any x_1, y_1 with $y_1 \subset x_1$, let $\{x_1\} \succ^1 \{y_1\}$. Since $y_1 \subset x_1$, for all $p \in \mathcal{S}^1$,

$$\max_{h \in x_1} U_2(h, p) \geq \max_{h \in y_1} U_2(h, p). \tag{41}$$

Furthermore, since $\{x_1\} \succ^1 \{y_1\}$, there must exist at least one state $p^* \in \mathcal{S}_2$ such that (41) holds with strict inequality. Thus, the assumption $\mathcal{S}^1 \subset \mathcal{S}^2$ implies $\{x_1\} \succ^2 \{y_1\}$.

(iii) (only if part) The first result is a direct consequence of part (ii). We will show the second claim.

Suppose otherwise. There exists $\mu^{1*} \in S_1^1$ such that $\mathrm{supp}(\mu^{1*}) \not\subset \mathrm{supp}(\mu^2)$ for all $\mu^2 \in S_1^2$. Since $\mathcal{S}^1 = \mathcal{S}^2$, $\mathrm{supp}(\mu^{1*}) \subset \mathcal{S}^2$. By applying Lemma B.9 to $\mathcal{S}^2$, we have positive numbers $\{v^p\}_{p \in \mathcal{S}^2}$ such that the menu x_1, defined as (39), satisfies all the properties in the lemma. For each $\mu^2 \in S_1^2$, choose h^p satisfying $U_2(h^p, p) = v^p$ for each $p \in \mathrm{supp}(\mu^2)$. Let $x_1(\mu^2)$ be the convex hull of $\{h^p | p \in \mathrm{supp}(\mu^2)\}$. Let $x_0 \equiv \{x_1(\mu^2) | \mu^2 \in S_1^2\}$ and $\bar{x}_1 \equiv \cup_{x_1 \in x_0} x_1$. For each $\mu^2 \in S_1^2$,

$$\int_{\mathcal{S}^2} \max_{h \in \bar{x}_1} U_2(h, p)\, \mathrm{d}\mu^2(p) = \max_{x_1 \in x_0} \int_{\mathcal{S}^2} \max_{h \in x_1} U_2(h, p)\, \mathrm{d}\mu^2(p),$$

and hence $\{\bar{x}_1\} \sim^2 x_0$. On the other hand, taking into account $\mathcal{S}^1 = \mathcal{S}^2$, for each $\mu^1 \in S_1^1$,

$$U_1(\bar{x}_1, \mu^1) = \int_{\mathcal{S}^2} \max_{h \in \bar{x}_1} U_2(h, p)\, \mathrm{d}\mu^1(p) \geq \max_{x_1 \in x_0} \int_{\mathcal{S}^2} \max_{h \in x_1} U_2(h, p)\, \mathrm{d}\mu^1(p) = \max_{x_1 \in x_0} U_1(x_1, \mu^1),$$

and this weak inequality holds with strict inequality for μ^{1*} because $\mathrm{supp}(\mu^{1*}) \not\subset \mathrm{supp}(\mu^2)$ for all $\mu^2 \in S_1^2$. Thus, we have $\{\bar{x}_1\} \succ^1 x_0$. This is a contradiction.

(if part) Assume that, for any $\mu^1 \in S_1^1$, there exists $\mu^2 \in S_1^2$ such that $\mathrm{supp}(\mu^1) \subset \mathrm{supp}(\mu^2)$. Take any finite x_0 and assume $\{x_1\} \succ^1 x_0$, where $x_1 \equiv \cup_{x'_1 \in x_0} x'_1$. Since $x'_1 \subset x_1$ for all $x'_1 \in x_0$, we have

$$\int \max_{h \in x_1} U_2(h, p)\, \mathrm{d}\mu^1(p) \geq \max_{x'_1 \in x_0} \int \max_{h \in x'_1} U_2(h, p)\, \mathrm{d}\mu^1(p), \tag{42}$$

for all $\mu^1 \in S_1^1$. Furthermore, since $\{x_1\} \succ^1 x_0$, there must exist at least one signal $\mu^{1*} \in S_1^1$ such that (42) holds with strict inequality. By assumption, there exists $\mu^{2*} \in S_1^2$ such that $\mathrm{supp}(\mu^{1*}) \subset \mathrm{supp}(\mu^{2*})$. Thus, the associated representation implies $\{x_1\} \succ^2 x_0$.

References

[1] F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *The Annals of Mathematical Statistics*, 34:199–205, 1963.

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

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

[4] L. G. Epstein. An axiomatic model of non-Bayesian updating. University of Rochester, working paper, March 2004.

[5] L. G. Epstein and S. Zin. Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica*, 57:937–969, 1989.

[6] P. Ghirardato. Coping with ignorance: unforeseen contingencies and non-additive uncertainty. *Economic Theory*, 17:247–276, 2001.

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

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

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

[10] 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.

[11] D. M. Kreps and E. L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46:185–200, 1978.

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

[13] K. Nehring. Preference for flexibility in a Savage framework. *Econometrica*, 67:101–119, 1999.

[14] E. Ozdenoren. Completing the state space with subjective states. *Journal of Economic Theory*, 105:531–539, 2002.

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

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

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

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

[19] N. Takeoka. Representation with a unique subjective decision tree. University of Rochester, working paper, November 2004.

[20] N. Takeoka. Subjective prior over subjective states, stochastic choice, and updating. University of Rochester, working paper, November 2004.