CIRJE-F-1083

Preferences for Randomization and Anticipated Utility

Yosuke Hashidate
CIRJE, Faculty of Economics, The University of Tokyo

April 2018

Preferences for Randomization and Anticipated Utility*

Yosuke Hashidate†

First Draft: January 2016; This Draft: April 11, 2018

Abstract

This paper presents a theory of preferences for randomization by using the framework of preferences over menus. In the framework, the decision maker chooses a menu at the first stage; at the second stage, she chooses a probability distribution on the chosen menu at the first stage. The resulting behavior is captured by expected utility theory, but at the choice of the first stage, the decision maker may have non-linear preferences due to cognitive effects. This paper introduces new axioms on preferences for randomization by relaxing the axioms of *Strategic Rationality* and *Independence*. This paper imposes on the axioms of *Randomization* and *Strong Singleton Independence*. The new axioms, along with basic axioms, characterize a *random anticipated utility representation*, in which the subjective belief for the effect of randomization is uniquely identified. Randomization attitudes, captured by probability-weighting functions, and risk attitude are separately identified. By relaxing the two axioms, this paper studies more general cases such as preferences for flexibility, subjective learning, and costly randomization. Moreover, the resulting behaviors are characterized by stochastic choice functions.

Keywords: Preferences for Randomization; Deliberate Randomization; Anticipated Utility; Non-Expected Utility Theory.

JEL Classification Numbers: D01, D81, D91.

*I am indebted to my adviser Akihiko Matsui for his unique guidance, constant support, and encouragement. I would like to thank Mira Frick, Kazuhiro Hara, Atsushi Kajii, Yusuke Kasuya, Nobuo Koida, Bart Lipman, Efe Ok, Kota Saito, Koji Shirai, and Norio Takeoka for their invaluable discussions and suggestions. In particular, I wish to thank Keisuke Yoshihara for his help of numerical analysis, and Ryota Iijima for his insightful comments. I am grateful for the participants at Nagoya (JEA Spring Meeting), Kyoto (4th Summer School of Econometric Society), and Hitotsubashi (ABEF). The earlier version of this paper (Hashidate (2017)) won the honor award for young researchers from the Association of Behavioral Economics and Finance (ABEF) in December 2016. I gratefully acknowledge financial support from the Japan Society for the Promotion of Science (Grant-in-Aid for JSPS Fellows (02605576)). I would like to thank Editage (www.editage.jp) for English language editing. All remaining errors are mine.

†CIRJE, Faculty of Economics, the University of Tokyo: 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033; Email: yosukehashidate@gmail.com

1 Introduction

This paper presents a theory of *preferences for randomization* by studying preferences over menus. The theory is included into a class of non-expected utility theory (non-EUT). In real life, there are some cases such that you delegate decision-making to a coin toss due to indifferent or incomplete preferences (Eliaz and Ok (2006)), information acquisition under uncertainty (Raiffa (1961)), etc. Indeed, experimental or behavioral economics has shown that subjects exhibit a systematic deviation from the EUT, particularly, the violation of the *Independence* axiom.

In decision theory, Machina (1985) argues that the decision maker may make a decision with a *deliberate randomization* if the decision maker has a *non-expected utility preference*, especially if preferences over lotteries are *convex*. The train of thought in the cognitive process does not always follow "probability-theoretic" probabilities, as Ellsberg (1961) shows; that is, people often deviate from the axiom of *Independence*.

Recently, various experimental evidence from laboratory and field settings indicates that preferences for the desire for randomization are supported. Agranov and Ortoleva (2017) demonstrate that, asked the same question with multiple times, subjects deliberately attempt to give stochastic answers.[1] Dwenger et al. (2014) suggest that decision makers randomize between alternatives in attempt to minimize regret.

The difficulty in the study of randomization in the mind is that we generally do not observe subjective randomization, as present in a coin toss in the mind. More importantly, we may not distinguish objective lotteries with subjective randomization. Machina (1984) states that underlying preferences that satisfy EUT will not be chosen over "temporal" risky prospects in a manner which can be modeled as EUT. Since most economically important instances of risk-taking behaviors involve delays as opposed to preferences for the early resolution of risk (Kreps and Porteus (1978)), the standard use of EUT to model such decisions is questionable.

This paper takes preferences over menus as primitives and provides an axiomatic foundation for *preferences for randomization.* The framework of preferences over menus has a two-stage decision problem. In this paper, the decision maker chooses a menu (choice set) at the first stage; at the second stage, she /he chooses a probability distribution over the chosen menu at the first stage. As the decision maker determines a probability at the second stage, the resulting behavior is captured by the structure of EUT. The structure holds with the linearity in probabilities. However, at the first stage, i.e., when the decision maker chooses menus, it may not be reasonable that the linearity in probabilities holds. If the decision

[1]See also Dean and McNeill (2015). They run experimental tests of the link between *preferences for flexibility* (preferring larger menus) and stochastic choice behaviors in a real-effort task. They find that *preference for flexibility* is important: 61% of subjects exhibited a *preference for flexibility* when choosing between contracts to use at a future date.

maker exhibits *preferences for randomization*, then the axiom of *Independence* is violated.[2]

The key axiom is a *monotonic* condition of preferences for randomization, stated as *Randomization*. This new axiom is introduced to study preferences for randomization (at the first stage). The decision maker considers a probability distribution on each menu. This probability distribution is interpreted as "subjective" randomization in one's mind. This axiom states that if there exists a randomization on a menu that *dominates* any randomization on another menu, then the menu is preferred to another one. The effect of randomization is evaluated by the certainty equivalent of randomization. Axiomatically, this axiom is a weaker version of *Strategic Rationality* in Kreps (1979).

This paper also relaxes the axiom of *Independence*. Deliberate randomization mainly stems from non-expected utility preferences, especially Allais-type preferences (Machina (1985), Cerreia-Vioglio et al. (2017)), implementation costs (Fudenberg et al. (2015)), hedging (Saito (2015)[3]), and regret aversion (Dwenger et al. (2014)). These issues are closely related to nonlinear preferences. To allow for this, this paper provides a weaker version of *Independence*. Since there is no effect of randomization in one's mind, this paper requires that the axiom of *Independence* holds under singleton menus.

The key axiom of *Randomization*, along with other basic axioms including *Strong Singleton Independence*, characterizes a *random anticipated utility representation* in which the decision maker's subjective belief for deliberate randomization is uniquely identified. In this representation, the decision maker evaluates a menu by randomizing the alternatives in a menu. The utility of a menu is described as the "optimal" random choice on the menu; that is, the decision maker chooses the optimal probability distribution on the menu. This paper uniquely identifies the effect of randomization in one's mind as an "*anticipation* (*subjective belief*)." This anticipation is mathematically described as a *probability weighting function*.[4] The anticipation reflects the subjective belief for deliberate randomization in one's mind, which affects the optimal random choices on menus.

The contributions of this paper are threefold. First, this paper elicits a subjective belief of

---

[2]If menu-preferences satisfy the axiom of *Independence*, then the resolution of objective lotteries does not matter. The difference between compound lotteries and convex combinations does not matter. See Dekel et al. (2001) in detail.

[3]Saito (2015) is the seminal literature of *preferences for randomization*. He investigates a decision maker's preferences over sets of *Anscombe=Aumann acts* (Anscombe and Aumann (1963)), and presents an axiomatic foundation, in which the subjective belief that the decision maker's randomization in her mind eliminates the effects of uncertainty, is uniquely identified. In his study, the incentive for the desire to randomization stems from *uncertainty-averse preferences* (Gilboa and Schmeidler (1989)). In this sense, his study is limited to only decision-making under uncertainty. See also Raiffa (1961).

[4]The idea of probability weighting functions is reminiscent of the study of non-expected utility theory in Quiggin (1982). However, we need to mention that randomization attitudes captured by g is different from rank-dependent utility. Rank-dependent utility captures a cognitive habit in which lotteries (probabilities) are not perceived in the standard manner.

randomization from *deterministic* preferences. The key is that this paper takes preferences over menus as primitives, instead of using stochastic choice functions as primitives. As preferences for randomization can be interpreted as a type of *taste uncertainty*, this paper studies anticipated utility from the viewpoint of "subjective" uncertainty.[5] In fact, stochastic choice functions may not be reasonable because preferences for randomization include the *aversion* to randomization. Instead, we consider a plausible axiom on menu preferences for preferences for randomization.

Second, this paper identifies a class of *preferences for randomization* ranging from the *desire for randomization* to the *aversion to randomization*. The experimental evidence in Agranov and Ortoleva (2017) and Dwenger et al. (2014) is related to the *desire for randomization*. On the contrary, the decision maker may exhibit preferences for the *aversion to randomization* due to status-quo effects in intertemporal choices. This paper characterizes a class of *preferences for randomization* in terms of the subjective belief of "deliberate randomization." Randomization attitudes are different from risk attitudes; this difference is captured by the third motivation.

Third, this paper shows that the effect of randomization in one's mind stems from several cognitive or psychological effects: (i) complementarities across attributes (Subsection 6.1), (ii) preferences for delay (Subsection 4.2), (iii) costs of thinking (Subsection 4.3), and so on. First, in attribute-based inferences, the comparisons with complementarities between attributes can lead to hard choices. We consider this procedural aspect by relaxing the axiom of *Randomization*. To do so, we allow for menu-dependent probability-weighting functions. Second, in the study of the timing of decision-making, the resolution of subjective uncertainty is related to preferences for randomization. We study a relationship between preferences for randomization and preferences for delay by relaxing the axiom of *Randomization*. Third, deliberate randomization, such as coin toss in the mind, is a procedure in the cognitive process, so it may be costly. The cost of thinking for deliberate randomization is considered by modifying the axiom of *Independence*.

There are some related studies for deliberate randomization. Fudenberg et al. (2015) present a model of deliberate randomization in which the decision maker maximizes the sum of expected utility with a perturbation function (cost function). Although Fudenberg et al. (2015) is closely related to this paper, there are some differences between this paper and Fudenberg et al. (2015). In Fudenberg et al. (2015), it is postulated that the decision maker exhibits the *desire for randomization*. In terms of this paper, only the case where the deci-

[5] In the study of *subjective uncertainty*, see Kreps (1979) and Dekel et al. (2001). The term of "subjectivity" in uncertainty is different from the study of decision-making under uncertainty (Savage (1954)/Anscombe and Aumann (1963)). In Savage (1954) and Anscombe and Aumann (1963), state spaces are *exogenously* given. On the other hand, Kreps (1979) studies an "endogenous" state space, and Dekel et al. (2001) uniquely identifies subjective state spaces. In the framework of *subjective state spaces*, "subjective" uncertainty is captured by an "endogenous" state space.

sion maker exhibits the *desire for randomization* is considered. To compare this paper with Fudenberg et al. (2015), we study a menu-dependent probability-weighting function (*subjective belief*). In this case, this paper corresponds to a menu-invariant additive perturbed utility representation in Fudenberg et al. (2014).

Cerreia-Vioglio et al. (2017) develop an axiomatic model of deliberate randomization by studying a subclass of non-expected utility preferences in Cerreia-Vioglio et al. (2015).[6] One of the objectives of this paper is to uniquely identify a class of preferences for randomization, including the *desire for randomization*, *indifference to randomization*, and *aversion to randomization*. On the other hand, in Cerreia-Vioglio et al. (2017), different menus exhibit different attitudes toward randomization. In a menu, the decision maker may exhibit the desire for randomization, but, in another menu, the same decision maker may exhibit an aversion to randomization. To compare this paper with Cerreia-Vioglio et al. (2017), again we consider a menu-dependent probability-weighting function (*subjective belief*) by introducing *Cautious EU preferences* (Cerreia-Vioglio et al. (2015)).

The rest of this paper is organized as follows. In Section 2, we present the axioms in the main theorem. In Section 3, we present the representation theorem (Theorem 1), the uniqueness result (Proposition 1), and the characterization of the model, respectively. In Section 4, we introduce and relax some axioms to study preferences for flexibility, subjective learning, and thinking aversion for a better and deeper understanding of *preferences for randomization*. In Section 5, we characterize the ex-post stochastic choice of the *random anticipated utility representation*. In Section 6, we state that our model is consistent with several experimental results in *preferences for randomization*. In Section 7, we provide a literature review to compare this paper with other studies. In Section 8, we conclude. All proofs are in the Appendix.

2 Axioms

We introduce notation briefly. Let $X \subseteq \mathbb{R}^n$ be a *convex* and *compact* set of all alternatives. Each $n \in \mathbb{N}$ is interpreted as an attribute of alternatives.[7] The elements in X are denoted by $x, y, z \in X$. Let $\Delta(X)$ be the set of all probability distributions on X with finite support. The elements in $\Delta(X)$ are denoted by $p, q, r \in \Delta(X)$. In this paper, it is postulated that each lottery is an *option*.[8] Let $\mathcal{A}$ denote the set of all non-empty *compact* subsets of $\Delta(X)$,

[6]See also Machina (1985).

[7]This setting is considered to study the effect of subjective randomization. We assume that $|n| > 1$. The case $|n| = 1$ corresponds to *monetary prizes*.

[8]The decision maker may not know about true values of alternatives, and then can evaluate them as expected utility.

endowed with the Hausdorff metric. The Hausdorff metric is defined by

$$d_h(A, B) := \max \left\{ \max_{p\in A} \min_{q\in B} d(p, q), \max_{p\in B} \min_{q\in A} d(p, q) \right\}.$$

The elements in $\mathcal{A}$ are called *menus*, which are denoted by $A, B, C \in \mathcal{A}$. We use the convex combination on menus in the standard manner: for any $A, B \in \mathcal{A}$ and $\lambda \in [0, 1]$, $\lambda A + (1 - \lambda)B := \{ \lambda p + (1 - \lambda)q \mid p \in A,\ q \in B\}$.

The primitive of the model is a binary relation $\succeq$ over $\mathcal{A}$, which describes the decision maker's choice of sets of lotteries. As usual, the asymmetric and symmetric parts of $\succeq$ are denoted by $\succ$ and $\sim$, respectively.

2.1 Standard Preferences

We introduce basic requirements in decision theory.

Axiom (Standard Preferences): $\succeq$ satisfy (i) a *weak order*, (ii) *continuity*, and (iii) *non-degeneracy*:

(i) (Weak Order): $\succeq$ is *complete* and *transitive*.

(ii) (Continuity): The sets $\{A \in \mathcal{A} \mid A \succeq B\}$ and $\{A \in \mathcal{A} \mid B \succeq A\}$ are closed in the Hausdorff metric.

(iii) (Strict Non-Degeneracy): There exists $p, q \in \Delta(X)$ such that $\{p\} \succ \{q\}$.

2.2 Preferences for Randomization

We introduce the key axiom of *Randomization*. To do so, we define the following. Let $\trianglerighteq$ be a binary relation on $\Delta(\Delta(X))$, the set of all probability distributions on $\Delta(X)$. The asymmetric and symmetric parts of $\trianglerighteq$ are denoted by $\triangleright$ and $\simeq$, respectively. Let δ_x be a *degenerate* lottery (Dirac measure at x), which gives x with certainty. For any $p \in \Delta(X)$, let c_p be the *certainty equivalent* of p, i.e., an element in X with $p \simeq \delta_{c_p}$.

Definition 1. For any $\rho, \mu \in \Delta(X)$ such that $\rho = \rho_1 \circ p_1 \oplus \cdots \oplus \rho_m \circ p_m$ and $\mu = \mu_1 \circ q_1 \oplus \cdots \oplus \mu_k \circ q_k$, we say that ρ *dominates* μ if

$$c_{\rho_1 p_1 + \cdots + \rho_m p_m} \geq c_{\mu_1 q_1 + \cdots + \mu_k q_k}.$$

The definition says that if a randomization ρ *dominates* another randomization μ, then the decision maker believes that ρ is better than μ, irrespective of her subjective belief for the effect of randomization (desire/indifference/aversion). Notice that both ρ and μ are *objective* lotteries on $\Delta(X)$. The certainty equivalent of them captures the effect of randomization.

This attitude is a *monotonic* condition for the effect of randomization. The main axiom is stated as follows.

Axiom (Randomization): For any $A, B \in \mathcal{A}$, if for any $\mu \in \Delta(B)$, there exists $\rho \in \Delta(A)$ such that ρ *dominates* μ, then

$$A \succeq B.$$

To illustrate, consider an example in Figure 1. Take two menus: $\{\delta_x, \delta_y\}$ and $\{\delta_z\}$. Suppose that each alternative is not *dominated* each other. If the decision maker randomizes δ_x and δ_y, as Figure 1 shows, a randomization ρ *dominates* the singleton menu $\{\delta_z\}$. Hence, the decision maker prefers $\{\delta_x, \delta_y\}$ to $\{\delta_z\}$ by enjoying the effect of randomization.

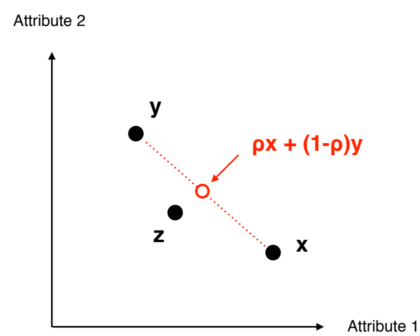

Figure 1: Preferences for Randomization

One remark is that this axiom is a weaker version of the axiom of *Strategic Rationality* (Kreps (1979)): $A \succeq B \Rightarrow A \sim A \cup B$. This axiom is equivalent to the following statement: if for any $q \in B$, there exists $p \in A$ such that $\{p\} \succeq \{q\}$, then $A \succeq B$. Compared with this axiom, the axiom of *Randomization* requires that for any randomization μ on the menu B, there exists a randomization ρ on the menu A such that a randomization ρ on A *dominates* any randomization μ on B, then the menu A is preferred to the menu B.

2.3 Relaxing Independence

We provide a weaker version of the axiom of *Independence.* The axiom of *Independence* is stated as follows.

Axiom (Independence): For any $A, B, C \in \mathcal{A}$ and $\lambda \in [0, 1]$,

$$A \succeq B \Rightarrow \lambda A + (1 - \lambda)C \succeq \lambda B + (1 - \lambda)C.$$

Preferences for randomization may deviate from the axiom of *Independence*. The mixture on menus may change the ranking, due to the effect of randomization. To elicit and study the effect of randomization in one's mind, we provide the following weaker axiom of *Independence*.[9]

Axiom (Strong Singleton Independence): For any $p, q \in \Delta(X)$, any $r \in \Delta(X) \setminus \{p, q\}$, and any $\lambda \in [0, 1]$,

$$\{p\} \succeq \{q\} \Leftrightarrow \lambda\{p\} + (1 - \lambda)\{r\} \succeq \lambda\{q\} + (1 - \lambda)\{r\}.$$

This axiom requires that, when mixing with a singleton menu not included in both $\{p\}$ and $\{q\}$, the ranking on the mixture does not change the original ranking on the two singleton menus. The *linearity* holds since there is no effect of randomization under singleton menus.

3 Results

3.1 Representation Theorem

We state the main result. Given a menu $A \in \mathcal{A}$, let $\Delta(A)$ be the set of all probability measures on A, i.e., $\sum_{q \in A} \rho(q) = 1$, $\rho(q) \in [0, 1]$ for any $q \in A$.

Theorem 1. *The following statements are equivalent:*
(a) $\succeq$ *satisfies Standard Preferences, Randomization, and Strong Singleton Independence.*
(b) There exists a pair $\langle u, g \rangle$ *where* u *is a non-constant function* $u : \Delta(X) \to \mathbb{R}$ *and* g *is a continuous and strictly increasing function* $g : [0, 1] \to [0, 1]$ *where* $g(0) = 0$ *and* $g(1) = 1$ *such that* $\succeq$ *is represented by* $V : \mathcal{A} \to \mathbb{R}$ *defined by*

$$V(A) = \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q)).$$

We call the value function V a *random anticipated utility representation* (RAU) if $\succeq$ satisfies the axioms in Theorem 1.

The interpretation of RAU is as follows. The utility of a menu A states that the decision maker chooses the *optimal* probability distribution on the menu A, by randomizing options (*lotteries*) in the menu A. The utility captures how frequently the decision maker chooses each alternative. The non-constant function u is an expected utility, and evaluates alternatives. The probability-weighting function g captures a subjective attitude toward the effect of randomization in the mind.

One remark is that RAU is different from rank-dependent utility (RDU) (Quiggin (1982)). In RDU, objective lotteries are distorted. In RAU, objective lotteries correspond to options.

[9] In the framework of preferences over menus, various weaker versions of *Independence* are introduced. For example, see Ergin and Sarver (2010) and Noor and Takeoka (2015).

In the general model of RAU, options themselves are evaluated in the standard manner of EUT.

To get an intuition, we provide a simple numerical example. For simplicity, assume that $u : X \to \mathbb{R}$ where $X \subset \mathbb{R}^2$. Consider a menu $A = \{x, y\}$ where $x = (7, 3)$ and $y = (8, 2)$. Suppose that a utility function is given by, for each $x \in X$, $u(x) = \log x_1 + \log x_2$. In addition, a probability weighting function is given by $g(\rho(\cdot)) = \sqrt{\cdot}$, for each $\rho(\cdot) \in [0, 1]$. Then, the optimal random choice as the ex-post stochastic choice is described by $\rho^*(A) = (\rho^*(x), \rho^*(y)) = (0.5466, 0.4534)$. The decision maker chooses an alternative x with probability 0.5466, and chooses another alternative y with probability 0.4534. Consequently, in the ex-ante utility, we have $V(A) \approx 4.118 \geq 3.045 = V(\{x\})$ and $V(A) \approx 4.118 \geq 2.773 = V(\{y\})$. Since g is concave, the resulting behavior reflects her belief for effect of randomization. That is, the decision maker prefers randomizing them on the menu A to choose each singleton menu.

Claim 1. *In the framework of subjective state spaces, random anticipated utility representations correspond to the case that it has one positive state, in which the sign on the state is positive. In Section 4.1, we consider a general case where there exists a subjective state space, by relaxing the axiom of Randomization. We obtain the representation in which different subjective states have different (state-dependent) probability-weighting functions.*

Claim 2. *Random anticipated utility representations (RAU) are not consistent with FOSD (First Order Stochastic Dominance). The representation $\sum_{q \in A} u(q)g(\rho(q))$ can deviate from the monotonicity with respect to FOSD even if both u and g are strictly increasing. The random anticipated utility representation satisfies FOSD if $g(\rho(\cdot)) = \rho(\cdot)$ for all $\rho(\cdot) \in [0, 1]$ (see also Proposition 3).*

Claim 3. *We reconsider the axiom of Randomization to elicit menu-dependent probability-weighting functions.*

Consider the following axiom. The axiom is an alternative version of Randomization. The attitude toward randomization, i.e., the randomization attitude depends on choice sets generally.[10]

Axiom (Weak Randomization): For any $A, B \in \mathcal{A}$, if for any $\mu \in \Delta(B)$, there exists $\rho \in \Delta(A)$ such that $\rho \trianglerighteq \mu$, then $A \succeq B$.

The axiom says that if for any randomization μ on B, there exists a "better" randomization ρ on A, then A is preferred to B. We have the following corollary.

[10] The setting $X \subseteq \mathbb{R}^n$ does not matter in this axiom.

Corollary 1. *$\succeq$ satisfies Standard Preferences, Weak Randomization, and Singleton Independence if and only if there exists a pair $\langle u, (g_A)_{A\in\mathcal{A}}\rangle$ where u is a non-constant function $u : \Delta(X) \to \mathbb{R}$ and g_A is a continuous and strictly increasing function $g : \mathcal{A} \times [0, 1] \to [0, 1]$ where $g_A(0) = 0$ and $g_A(1) = 1$ such that $\succeq$ is represented by $V : \mathcal{A} \to \mathbb{R}$ defined by*

$$V(A) = \max_{\rho\in\Delta(A)} \sum_{q\in A} u(q)g_A(\rho(q)).$$

3.2 Uniqueness Result

We present the uniqueness result for *random anticipated utility representations.*

Proposition 1. *If two random anticipated utility representations with $\langle u, g\rangle$ and $\langle u', g'\rangle$ represent the same preference relation $\succeq$, then the following holds:*

(i) u is unique up to a positive affine transformation;

(ii) $g = g'$.

3.3 Proof Overview

We provide the proof outline of the sufficiency part of Theorem 1. The key point of the *random anticipated utility representation* is a separation of utility and subjective belief for deliberate randomization, i.e., $\langle u, g\rangle$. We obtain a *decomposition* of non-additive belief for subjective randomization to elicit the probability-weighting function g.

In Step 1, we construct a value function $V : \mathcal{A} \to \mathbb{R}$ such that, for any $A, B \in \mathcal{A}$, $A \succeq B \Leftrightarrow V(A) \geq V(B)$ (Lemma 1). Define, for any $p \in \Delta(X)$, $V(\{p\}) = u(p)$ for some $u : \Delta(X) \to \mathbb{R}$. Of course, this case corresponds to singleton menus.

In Step 2, we consider the case of doubleton menus. By applying the bi-separability of $\succeq$ in Ghirardato and Marinacci (2001), we elicit a cardinal utility function $u : \Delta(X) \to \mathbb{R}$ and a capacity θ^* on the set of doubleton menus, denoted by $\mathcal{A}^* \subset \mathcal{A}$. First, we show that $\succeq$ that satisfies the axioms in Theorem 1 satisfies the axiom of $\frac{1}{2}$-*Average*, by mainly using the axioms of *Continuity* and *Strong Singleton Independence* (Lemma 2).

Axiom ($\frac{1}{2}$-Average): For any $p, q \in \Delta(X)$ such that $\{\delta_{c_p}\} \succeq \{\delta_{c_q}\}$, there exists $r \in \Delta(X)$ where $\{\delta_{c_p}\} \succeq \{\delta_{c_r}\} \succeq \{\delta_{c_q}\}$ such that

$$\frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \frac{1}{2}\Big\{\frac{1}{2}\delta_{c_p} + \frac{1}{2}\delta_{c_r}\Big\} + \frac{1}{2}\Big\{\frac{1}{2}\delta_{c_q} + \frac{1}{2}\delta_{c_r}\Big\}.$$

This axiom states that *preference average* holds under "objective" lotteries for singleton menus with average options.

Next, by using the axiom of $\frac{1}{2}$-*Average*, we show that V is mixture-linear with respect to singleton menus (Lemma 3). To show that V is mixture-linear with respect to singleton menus, consider *dyadic rational*, i.e., a number $\gamma \in (0,1)$ such that for some finite integer $M \in \mathbb{N}$,

$$\gamma = \sum_{i=1}^{M} \frac{a_i}{2^i},$$

where $a_i \in \{0,1\}$ for every $i \in \{1, \cdots, M\}$ and $a_M = 1$. By the axiom of $\frac{1}{2}$-*Average*, we can consider the following *preference average*:

$$\frac{1}{2}\{r_1\} \oplus \frac{1}{2}\Big(\cdots \Big(\frac{1}{2}\{r_{M-1}\} \oplus \frac{1}{2}\Big(\frac{1}{2}\{r_M\} \oplus \frac{1}{2}\{q\}\Big)\Big) \cdots \Big),$$

where for any $i \in \{1, 2, \cdots, M\}$, $r_i = p$ if $a_i = 1$, $r_i = q$ otherwise. Thus, we can show that V is mixture-linear with respect to singleton menus (Lemma 4).

Finally, in the case of *doubletons*, we show that $\succeq$ is represented by $V : \mathcal{A}^* \to \mathbb{R}$ for some pair $\langle u, \theta^* \rangle$ where $u : \Delta(X) \to \mathbb{R}$ is a cardinal utility function, and θ^* is a capacity on $\mathcal{A}^*$ (Lemma 5). The non-additive measures on menus are different from probability measures. Notice that, by the axiom of *Strong Singleton Independence*, various attitudes toward subjective randomization are allowed.

In Step 3, we extend the result in Step 2 into the whole domain $\mathcal{A}$, and show that, for any $A \in \mathcal{A}$, a capacity θ is *decomposable*; that is, for any $A \in \mathcal{A}$, $\theta(A) = g \circ \rho(A)$ for some continuous and strictly increasing function $g : [0,1] \to [0,1]$. To do so, first, we induce a binary relation $\trianglerighteq^*$ on $\mathcal{A}$, a ranking of *deliberate randomization*, by using the axiom of *Randomization*. This binary relation states that how effective the randomization on a menu is (*subjective belief*). The binary relation is a qualitative probability measure. Then, we show that for all $A, B \in \mathcal{A}$, $A \trianglerighteq^* B \Leftrightarrow \theta(A) \geq \theta(B)$. Next, we also show that θ has a property, *weak additivity*, which is a necessary and sufficient condition to elicit a probability-weighting function g, which is continuous and strictly increasing. Finally, we show that for all $A \in \mathcal{A}$, θ is *decomposable*, i.e., for any $A \in \mathcal{A}$, $\theta(A) = g \circ \rho(A)$ (Lemma 6).

In Step 4, we consider the general case, and complete the desired utility representation. First, we consider the case of *doubletons*; that is, we show that $\succeq$ on $\mathcal{A}^*$ is represented by V defined by a pair $\langle u, g \rangle$ (Lemma 7). Next, we show that the representation of $\succeq$ on $\mathcal{A}^*$ is extended into the whole domain $\mathcal{A}$ by induction. To do so, by using mainly the axioms of *Randomization* and *Strong Singleton Independence*, we show that $\succeq$ satisfies a weaker version of *Strong Singleton Independence*.

Axiom (Singleton Independence): For any $\lambda \in [0,1]$ and $r \in \Delta(X)$,

$$A \succeq B \Rightarrow \lambda A + (1-\lambda)\{r\} \succeq \lambda B + (1-\lambda)\{r\}.$$

By using the weaker version of *Strong Singleton Independence*, along with the result in Step 3, we show that the representation of $\succeq$ is extended into the whole domain $\mathcal{A}$. Thus, we obtain the desired utility representation.

3.4 Characterization of the Attitude toward Randomization

We characterize the probability weighting function g. In Theorem 1, the property of the probability weighting function g is strictly increasing and continuous. We show that the additional axioms characterize randomization attitudes.

We characterize the attitude toward the *desire for randomization*, and the *aversion to randomization*, respectively.

Axiom (Desire for Randomization): If $\{p\} \sim \{q\}$, then, for any $\lambda \in [0, 1]$,

$$\lambda \circ \{p\} \oplus (1 - \lambda) \circ \{q\} \succeq \{p\}.$$

This axiom states that if two lotteries p and q are indifferent, then randomizing between p and q is more preferred. On the other hand, the following axiom, the axiom of *Aversion to Randomization* requires that the decision maker dislikes randomizing between p and q.

Axiom (Aversion to Randomization): If $\{p\} \sim \{q\}$, then, for any $\lambda \in [0, 1]$,

$$\{p\} \succeq \lambda \circ \{p\} \oplus (1 - \lambda) \circ \{q\}.$$

We apply the following definitions for the concavity and the convexity of g in Wakker (2010) (p.174).

Definition 2. Let $g : [0, 1] \to [0, 1]$ be a continuous and strictly increasing function satisfying $g(0) = 0$ and $g(1) = 1$.

(i) We say that a probability weighting function g is *concave* if $g(a + b') - g(b') \leq g(a + b) - g(b)$ whenever $1 \geq b' \geq b \geq 0$.

(ii) We say that a probability weighting function g is *convex* if $g(a+b')-g(b') \leq g(a+b)-g(b)$ whenever $1 \geq b' \geq b \geq 0$.

Proposition 2. *Suppose that a random anticipated utility representation is represented by a pair $\langle u, g\rangle$. Then, the following statements hold:*

(i) g is concave if and only if $\succeq$ exhibits Desire to Randomization;

(ii) g is convex if and only if $\succeq$ exhibits Aversion to Randomization.

3.5 No Effect of Randomization

We define the following. We say that a probability weighting function g is *linear* if for all $\rho \in [0, 1]$, $g(\rho) = \rho$.

Proposition 3. *Suppose that a random anticipated utility representation is represented by a pair $\langle u, g\rangle$. Then, $\succeq$ satisfies Independence if and only if g is linear.*

Suppose that $\succeq$ is represented by a pair $\langle u, g\rangle$ satisfying the axiom of *Independence*. Then, $V : \mathcal{A} \to \mathbb{R}$ is *mixture-linear*: for any $A, B \in \mathcal{A}$ and $\lambda \in [0, 1]$,

$$V(\lambda A + (1 - \lambda)B) = \lambda V(A) + (1 - \lambda)V(B).$$

The binary relation $\trianglerighteq$ on $\Delta(\Delta(X))$ satisfies the axiom of *Independence* of the vNM-type EUT. Hence, g must be *linear*.

Moreover, we obtain the following corollary.

Corollary 2. *Suppose that a random anticipated utility representation is represented by a pair $\langle u, g\rangle$. Then, $\succeq$ satisfies Independence if and only if $\succeq$ satisfies Strategic Rationality.*

In this paper, if $\succeq$ satisfies the axiom of *Strategic Rationality* ($A \succeq B \Rightarrow A \sim A \cup B$), then we cannot capture any beliefs for the effect of randomization. Suppose that $\succeq$ satisfies the axiom of *Strategic Rationality.* Since g is *linear*, the RAU representation is as follows. For any $A \in \mathcal{A}$,

$$V(A) = \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q)\rho(q).$$

If for any $r \in A$, there exists $q \in A$ such that $u(q) > u(r)$, then the decision maker puts a probability one on the alternative q, i.e., $\rho(q) = 1$. This reduces to "deterministic" choices, which is an expected utility.

We mention that the form of g is related to a difference between "indecisiveness" and "indifference."[11] If g is linear, then the interpretation of choice correspondences by a random anticipated utility representation corresponds to Kreps (1988), which is related to "indifference." There is no desire or aversion to randomization. On the other hand, the interpretation of choice correspondence by a random anticipated utility representation corresponds to Sen (1993), which is related to "indecisiveness." Indecisiveness can lead to desire or aversion to randomization.

3.6 Comparative Statics on Randomization

We provide a comparative attitude toward probability weighting functions. First, we say that a decision maker 1 has a stronger preference for randomization than another decision maker 2 if the decision maker 1 prefers a menu A to a singleton menu $\{q\}$ in the set A whenever the decision maker 2 also does. Formally, the statement is summarized as follows.

[11] See also Eliaz and Ok (2006), who take a deterministic choice-theoretic approach.

Definition 3. $\succeq_1$ exhibits a *stronger preference for randomization* than $\succeq_2$ if, for any $A \in \mathcal{A}$ and $q \in A$,

$$A \succeq_2 \{q\} \Rightarrow A \succeq_1 \{q\}.$$

Next, in the following, we define the *more* concavity of g.

Definition 4. Given two probability weighting functions g_1, g_2, we say that a probability weighting function g_1 is *more concave than* g_2 if $[g_1(a+b) - g_1(b)] - [g_1(a+b') - g_1(b')] \geq [g_2(a+b) - g_2(b)] - [g_2(a+b') - g_2(b')]$ whenever $1 \geq b' \geq b \geq 0$.

Proposition 4. *Suppose that $\succeq_j$ ($j \in \{1,2\}$) are represented by $\langle u, g_j \rangle$. The following statements are equivalent.*

(i) $\succeq_1$ exhibits a stronger preference for randomization than $\succeq_2$.

(ii) g_1 is more concave than g_2.

Finally, we consider the ex-post stochastic choice in a *random anticipated utility representation.* Suppose that $\succeq$ is represented by a pair $\langle u, g \rangle$. Let

$$\rho^*(A) \in \arg \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q)).$$

Definition 5. Let $\rho^*(A) \in \arg\max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q))$. ρ_1^* and ρ_2^* have the *same mean* if $\sum_{q \in A} u(q) \rho_1^*(q) = \sum_{q \in A} u(q) \rho_2^*(q)$.

Corollary 3. *Suppose that $\succeq_j$ ($j \in \{1,2\}$) are represented by $\langle u, g_j \rangle$, and that the ex-post stochastic choice is represented by ρ_j^* ($j \in \{1,2\}$) where ρ_1^* and ρ_2^* have the same mean. Then, $\succeq_1$ exhibits a stronger preference for randomization than $\succeq_2$ if and only if ρ_2^* is a mean-preserving spread of ρ_1^*.*

4 Analysis

4.1 Preferences for Flexibility

We consider a weaker form of *random anticipated utility representations*, by incorporating *preferences for flexibility* (Kreps (1979)) into our framework. Suppose that a decision maker has a *preference for randomization.* However, before choosing a menu at the first stage, the decision maker may not know about the attitude toward her true randomization. There exists a type of *taste uncertainty.* Before the second stage, her subjective state is realized, and then she makes "optimal" random choices with her *state-dependent* probability-weighting function.

Since decision analysts cannot observe a subjective state space directly, we may consider the following axiom to capture subjective states.

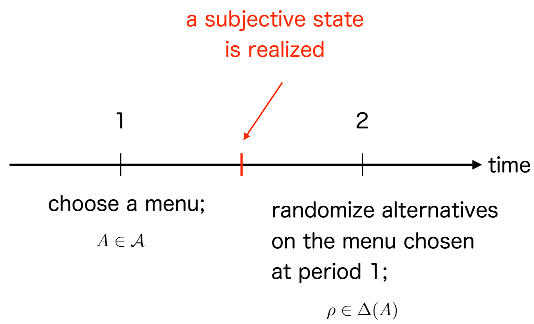

Figure 2: Timing of Decision-Making

Axiom (Monotonicity): For any $A, B \in \mathcal{A}$, if $B \subseteq A$, then $A \succeq B$.

Simply speaking, this axiom states that the decision maker prefers larger menus.[12] The reason for imposing on this axiom can be interpreted as follows. Suppose that the decision maker has a preference for randomization, but she is still uncertain about the true subjective state. To cope with such a subjective uncertainty, at the first stage, she prefers larger menus.

In Theorem 1, the *random anticipated utility representation* has a unique subjective state, as mentioned in Remark 1. To allow for multiple subjective states, we need to relax the axiom of *Randomization*. For each $A \in \mathcal{A}$, let $\text{co}(A)$ be a convex hull of a menu A.

Axiom (Monotonic Randomization): For any $A, B \in \mathcal{A}$, if

(i) for any $\mu \in \Delta(B)$, there exists $\rho \in \Delta(A)$ such that ρ *dominates* μ; and

(ii) $\text{co}(B) \subseteq \text{co}(A)$,

then $A \succeq B$.

The first condition (i) is the same as that in the axiom of *Randomization*. The second condition (ii) is new. The condition requires that $\text{co}(B)$ is a subset of $\text{co}(A)$. Each $\text{co}(A)$ reflects randomization in the mind. If $\text{co}(B)$ is a subset of $\text{co}(A)$, then the randomization in the menu A is at least beneficial than that in the menu B. In addition to the condition (i), if the condition (ii) holds, then the menu A is preferred to the menu B. Note that, since the condition (ii) is added into the axiom of *Randomization*, the condition is getting stronger. Consequently, the axiom itself is getting weaker.

[12] This axiom is introduced by Kreps (1979). In a recent development, for example, see Piermont et al. (2016).

We also consider the following axiom of *Independence.*

Axiom (Weak Independence): If for any $A, B, C \in \mathcal{A}$ and $\lambda \in (0,1)$, there exist $A', B' \in \mathcal{A}$ such that $A' \sim \lambda A + (1-\lambda)C$ and $B' \sim \lambda B + (1-\lambda)C$, then

$$A \succ B \Leftrightarrow A' \succ B'.$$

This axiom requires that $\succeq$ is *mixture linear* if there exist $A', B' \in \mathcal{A}$ such that $A' \sim \lambda A + (1-\lambda)C$ and $B' \sim \lambda B + (1-\lambda)C$. Such menus A', B' exist only if $\succeq$ is *state-wise linear.* There is no effect of *hedging* between subjective states, because the decision maker chooses an alternative after her subjective state is realized.

Proposition 5. *The following statements are equivalent:*
(a) $\succeq$ satisfies Standard Preferences, Monotonic Randomization, Strong Singleton Independence, and Weak Independence.
(b) There exists a four-tuple $\langle u, S, \mathcal{G}, \mu \rangle$ where u is a non-constant function $u : \Delta(X) \to \mathbb{R}$, S is a state space, and $\mathcal{G}$ is a set of state-dependent probability-weighting functions, i.e., $g : [0,1] \times S \to \mathbb{R}$ is a continuous and strictly increasing function where $g(0) = 0$ and $g(1) = 1$, and μ is a probability distribution over S, such that $\succeq$ is represented by $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \int_S \Big(\max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q), s) \Big) d\mu(s),$$

where μ is a probability distribution over S.

$g(\cdot, s)$ is a "state-dependent" probability-weighting function. The attitude toward randomization is changeable based on subjective states. As a result, observable stochastic choice data including the ratio of probabilities depends on u and $\mathcal{G}$. Especially, it also depends on how much each subjective state is realized.

In each singleton menu, we have $V(\{q\}) = u(q)$. This implies that in singleton menus, the subjective state space is not meaningful.

One remark is that the random anticipated utility representation with preferences for flexibility is not uniquely identified.

Claim 4. *(Identification) By the axiom of Monotonic Randomization, $\succeq$ exhibits preferences for flexibility. And, by the axiom of Weak Independence, V is mixture-linear. Thus, we can apply the result of Ahn and Sarver (2013). We consider the following axioms in Ahn and Sarver (2013). We consider $\langle \succeq, \rho \rangle$ as a primitive, where $\succeq$ is a binary relation on $\mathcal{A}$ and $\rho : \mathcal{A} \to \Delta(\Delta(X))$ is a probability distribution. For simplicity, throughout this remark, we assume that S is finite.*

We provide two axioms for $\langle \succeq, \rho \rangle$.

Axiom 1. If $A \cup \{p\} \succ A$, then $\rho(p|A \cup \{p\}) > 0$.

This axiom says that if adding p into A is beneficial, i.e., $A \cup \{p\}$ is (strictly) preferred to A, then the ex-post stochastic choice produces a positive probability on p.[13]

Axiom 2. For any $A \in \mathcal{A}$ and $p \notin A$, if there exists $\varepsilon > 0$ such that $\rho(q|B \cup \{q\}) > 0$ whenever $d(p,q) < \varepsilon$ and $d_h(A,B) < \varepsilon$, then $A \cup \{p\} \succ A$.

We say that $\succeq$ has a random anticipated utilty with subjective states if $\succeq$ has a utility representation in Proposition 5. And, we say that ρ has an ex-post stochastic choice of $\succeq$ if for all $A \in \mathcal{A}$ and $p \in A$,

$$\rho(p|A) = \sum_{s \in S} \Big(u(p) g(\rho^*(p), s) \Big) \mu(s).$$

We say that a pair $\langle \succeq, \rho \rangle$ has a general random anticipated utility representation if $\succeq$ has a random anticipated utility with subjective states, and ρ has an ex-post stochastic choice of $\succeq$.

Corollary 4. *Suppose that $\succeq$ has a random anticipated utility with subjective states and ρ has an ex-post stochastic choice of $\succeq$. Then, the pair $\langle \succeq, \rho \rangle$ satisfies Axiom 1 and Axiom 2 if and only if it has a general random anticipated utility representation.*

By considering a pair $\langle \succeq, \rho \rangle$, we identify a unique belief on a subjective state space, and a unique set of state-dependent probability-weighting functions.

Corollary 5. *The two general random anticipated utility representations $\langle u^1, S^1, \mathcal{G}^1, \mu^1 \rangle$ and $\langle u^2, S^2, \mathcal{G}^2, \mu^2 \rangle$ represent the same pair $\langle \succeq, \rho \rangle$ if and only if there exists a bijection $\pi : S^1 \to S^2$, and a positive real number $\alpha > 0$, and a real number $\beta \in \mathbb{R}$ such that the following equalities hold:*

(i) $u^1 = \alpha u^2 + \beta$.

(ii) $g_{s^1} = g_{\pi(s^1)}$, *for all* $s^1 \in S^1$.

(iii) $\mu^1(s^1) = \mu^2(\pi(s^1))$, *for all* $s^1 \in S^1$.

---

[13] Ahn and Sarver (2013) considers the following condition.
Axiom (Consequentialism): If $\rho(A) = \rho(B)$, then $A \sim B$.
In Proposition 1 in Ahn and Sarver (2013), if $\succeq$ satisfies the axiom of *Monotonicity* (preferences for flexibility), and ρ satisfies the axiom of *Regularity* ($\rho(p|A) \leq \rho(p|B)$ if $p \in B \subset A$), then Axiom 1 and *Consequentialism* are equivalent. Let us state another condition.

4.2 Preferences for Delay and Subjective Partitional Learning

We study how the timing of decision-making is related to preferences for randomization by studying subjective learning (Dillenberger et al. (2014)). In a *random anticipated utility representation* described by a pair $\langle u, g \rangle$, u is a non-constant function, and g is a probability-weighting function. To study the relationship between the timing of decision-making and preferences for randomization, we explicitly introduce an *objective* state space Ω, which is different from a subjective state space S. Formally, let Ω be a *finite* state space. The elements of the state space Ω is denoted by $\omega, \omega' \in \Omega$.

To illustrate, consider a simple story (see Figure 3). It is the time for lunch. A Ph.D. student is in front of the school gate to go for a lunch. If he turns right, there is a restaurant r. Conversely, if he turns left, there are two restaurants, p and q. For simplicity, suppose that the three restaurants are almost *indifferent.*[14] If he turns right, he has a lunch at the restaurant r with certainty. If he turns left, he chooses a restaurant from the two restaurants p and q. The key axiom, *Randomization*, requires that if there exists a subjective randomization on the two restaurants p and q that *dominates* any randomization on the restaurant r, then he should turn left.[15]

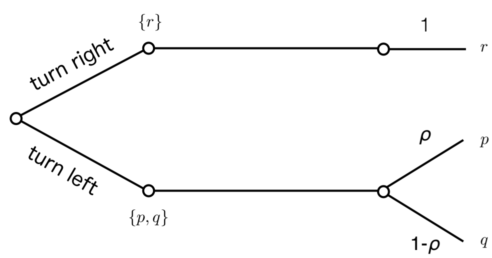

Figure 3: An Example

Turning left gives the Ph.D. student the opportunity to go to the two restaurants: p or q. However, the Ph.D. student has not known whether the restaurants are crowded or not.

[14]Let us assume that $\{p\} \sim \{q\} \sim \{r\}$.

[15]In attribute-based inferences, complementarities across attributes may exhibit a preference for randomization. For example, in the choice of restaurants in the motivating example, prices, locations, or service can be clue for decision-making. Gul et al. (2014) develops an attribute-based random choice rule by relaxing the axiom of Luce's IIA. In addition, preferences for randomization can be related to information acquisition. When he decides to turn right or left, the Ph.D. student may not know about how crowded each restaurant is at lunch. Obtaining such information can determine his attitude toward randomization in mind. Moreover, if he cares about costs of thinking, he might turn right to go to the restaurant r. Costs of thinking is also related to preferences for randomization.

We can consider the objective state space of the motivating example in the following way:

(i) state ω_1: the restaurant p is crowded, but the restaurant q is not;

(ii) state ω_2: the restaurant p is not crowded, but the restaurant q is;

(iii) state ω_3: the both the restaurants p and q are crowded;

and so on. Consequently, the decision maker can have "state-contingent" attitudes toward randomization.

Let $\Delta(X)^\Omega$ be the set of all state-contingent alternatives.[16] The elements of $\Delta(X)^\Omega$ is denoted by $\mathbf{p}, \mathbf{q} \in \Delta(X)^\Omega$. Let us denote a state-contingent alternative by $(p|\omega)$, which means that the decision maker chooses p if $\omega \in \Omega$ happens.

Definition 6. We say that an alternative $\mathbf{p}$ *state-wise dominates* another alternative $\mathbf{q}$ if for all $\omega \in \Omega$, $\{(p|\omega)\} \succeq \{(q|\omega)\}$.

If for any $\mathbf{q} \in B$ there exists $\mathbf{p} \in A$ such that $\mathbf{p}$ state-wise dominates $\mathbf{q}$, then $A \succeq B$. We provide the following axiom.

Axiom (Dominance): For any $\mathbf{p} \in A$ and $\mathbf{q} \in \Delta(X)^\Omega$ such that $\mathbf{p}$ *state-wise dominates* $\mathbf{q}$, if $\{\mathbf{p}\} \sim \{\mathbf{p}, \mathbf{q}\}$, then, $A \sim A \cup \{\mathbf{q}\}$.

Suppose that for any $\mathbf{p} \in A$, there exists $\mathbf{q} \in \Delta(X)^\Omega$ such that $\mathbf{p}$ *state-wise dominates* $\mathbf{q}$. If $\{\mathbf{p}\} \sim \{\mathbf{p}, \mathbf{q}\}$, then the decision maker chooses $\mathbf{p}$ from $\{\mathbf{p}, \mathbf{q}\}$ with certainty. The axiom of *Dominance* requires that if each $\mathbf{p}$ *state-wise dominates* $\mathbf{q}$ and, for each $\mathbf{p} \in A$, $\{\mathbf{p}\} \sim \{\mathbf{p}, \mathbf{q}\}$, then the decision maker does not put positive probabilities on $\mathbf{q}$. Hence, $A \sim A \cup \{\mathbf{q}\}$.

We introduce another new axiom. Let $I \in 2^\Omega$ be an *event*. Following Dillenberger et al. (2014), for any event $I \in 2^\Omega$ and $p, q \in \Delta(X)$, define a *composite* alternative pIq by

$$pIq(\omega) := \begin{cases} p & \text{if } \omega \in I \\ q & \text{otherwise.} \end{cases}$$

We provide the following axiom.

Axiom (State-Contingent Indifference): For any $\mathbf{p}, \mathbf{q} \in \Delta(X)^\Omega$, there exist $I \in 2^\Omega$ such that $\{\mathbf{p}I\mathbf{q}\} \sim \{\mathbf{p}, \mathbf{q}\}$.

This axiom says that for any two alternatives $\mathbf{p}$ and $\mathbf{q}$, there exists an event I such that the decision maker is *indifferent* between committing the *composite* alternative $\mathbf{p}I\mathbf{q}$ ex-ante, and choosing one of them ex-post.

[16]This setting is the same as that of Anscombe and Aumann (1963).

Proposition 6. *The following statements are equivalent:*
(a) $\succeq$ satisfies Standard Preferences, Randomization, Singleton Independence, Dominance, and State-Contingent Indifference.
(b) There exists a four-tuple $\langle u, (g_\omega)_{\omega \in \Omega}, \mu, \mathcal{P} \rangle$ where u is a non-constant function $u : \Delta(X) \to \mathbb{R}$ and g_ω is a state-contingent probability-weighting function, i.e., $g_\omega : [0,1] \to [0,1]$ is a continuous and strictly increasing function satisfying $g(0) = 0$ and $g(1) = 1$ for each $\omega \in \Omega$, μ is a probability distribution on Ω, and $\mathcal{P}$ is a partition of $\sigma(\mu)$, i.e., the support of μ, such that $\succeq$ is represented by $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \sum_{I \in \mathcal{P}} \max_{\rho \in \Delta(A)} \sum_{\omega \in I} \Big(\sum_{q \in A} u(x) g_\omega(\rho(q)) \Big) \mu(\omega).$$

Since we analyze *preferences for randomization* by explicitly imposing on an objective state space, the utility representation has some additional structures, compared with the *random anticipated utility representation* in Theorem 1.

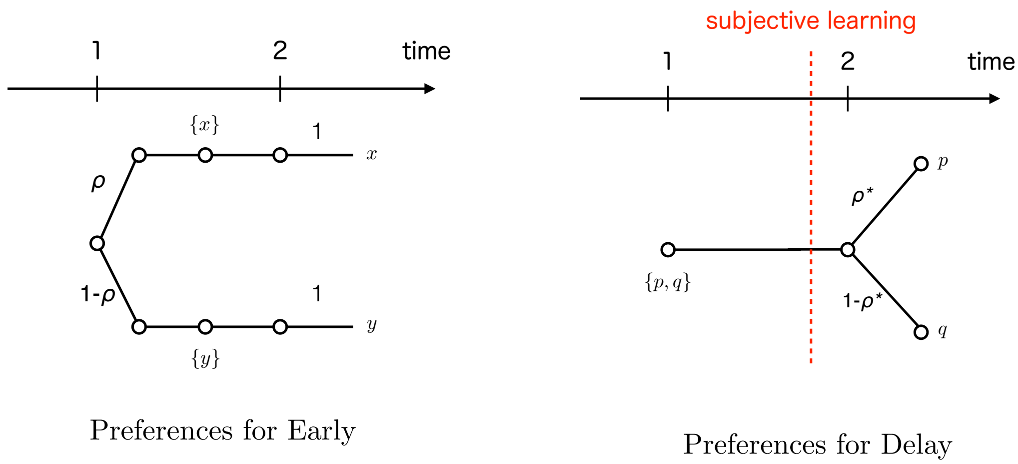

Figure 4: Preferences for Randomization and Subjective Learning

We provide an intuition of the *random anticipated utility representation with subjective partitional learning* in Figure 4. In the left hand side (LHS) of Figure 4, the decision maker chooses a menu $\rho \circ \{\mathbf{p}\} \oplus (1 - \rho) \circ \{\mathbf{q}\}$. On the other hand, in the right hand side (RHS) of Figure 4, the decision maker choose $\{\mathbf{p}, \mathbf{q}\}$. In the *random anticipated utility representation with subjective learning*, since the partition is *endogenous*, this partition captures information structures of the state space. The decision maker has in the mind a partition of the state space, and the partition describes what the decision maker expects to learn before the choice at the second stage. For example, in the motivating example, the Ph.D. student may obtain information about restaurants by using his social network services and his friends. This learning leads to *preferences for delay*. If the decision maker chooses a menu $\rho \circ \{\mathbf{p}\} \oplus (1 -$

$\rho) \circ \{\mathbf{q}\}$, then, irrespective of subjective learning, it is equivalent that the decision maker chooses $\mathbf{p}$ with probability ρ, and chooses $\mathbf{q}$ with probability $1-\rho$. In this sense, "objective" randomization is not consistent with subjective learning. On the other hand, in the RHS, the decision maker can postpone the timing of decision-making based on her own subjective learning. The ex-post stochastic choice behaviors can capture the decision maker's subjective randomization.

4.3 Thinking Aversion: A Costly Randomization

We extend the theory of *preferences for randomization* into *thinking aversion* in Ortoleva (2013). We apply the result of Ortoleva (2013) into *preferences for randomization*, since the size of menus has an effect on the decision maker's randomization in the mind. The mental cost of randomization in one's mind is crucially related to *deliberate randomization*. We show that *thinking aversion* is closely related to *preferences for randomization*, and the attitude toward randomization with thinking aversion is uniquely identified.

We study a relationship between the two binary relations: $\trianglerighteq$ on $\Delta(\Delta(X)$ and $\succeq$ on $\mathcal{A}$. Let ρ_A be a probability distribution on the menu A, for each $A \in \mathcal{A}$.

Definition 7. For any $\mu_B \in \Delta(B)$, there exists $\rho_A \in \Delta(X)$ such that $\rho_A \trianglerighteq \rho_B$ if $\lambda A + (1-\lambda)B \succeq \lambda' A + (1-\lambda')B$ for some $\lambda, \lambda' \in (0,1)$ with $\lambda > \lambda'$.

This definition states that the decision maker prefers a randomization on the menu A. We can infer that the decision maker prefers a menu A to another menu B. By using this definition, we consider the following axiom.

Axiom (Weak Thinking Aversion): For any $A \in \mathcal{A}$, $p, r \in \Delta(X)$, and $\lambda \in [0,1]$,

$$\delta_p \trianglerighteq \rho_A \Rightarrow \lambda\{p\} + (1-\lambda)\{r\} \succeq \lambda A + (1-\lambda)\{r\}.$$

This axiom states that if the decision maker prefers obtaining a singleton menu $\{p\}$ with higher probability, then the mixture with singleton menus does not change the ranking on the primitive of the model, i.e., $\succeq$ on $\mathcal{A}$. Since any singleton menus do not require *contemplation*, there is no need for thinking. The taste that the decision maker prefers obtaining a singleton menu with higher probability reflects the notion of "thinking aversion."

One remark is that the axiom of *Weak Thinking Aversion* is a weaker version of the axiom of *Strong Singleton Independence*; that is, by relaxing the axiom of *Strong Singleton Independence*, we can elicit costs of randomization in one's mind.

Proposition 7. *The following statements are equivalent:*
(a) $\succeq$ *satisfies Standard Preferences, Weak Randomization, and Weak Thinking Aversion.*
(b) There exists a tuple $\langle u, (g_A)_{A\in\mathcal{A}}, c\rangle$ *where* u *is a non-constant function* $u : \Delta(X) \to \mathbb{R}$

and g is a continuous and strictly increasing function $g : [0,1] \to [0,1]$ satisfying $g(0) = 0$ and $g(1) = 1$, and $c : \mathcal{A} \to \mathbb{R}$ is a monotonic function, i.e., for all $A, B \in \mathcal{A}$ with $B \subseteq A$, $c(\text{supp}(\rho)) \geq c(\text{supp}(\mu))$ where $\rho \in \Delta(A)$ and $\mu \in \Delta(B)$, such that $\succeq$ is represented by $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g_A(\rho(q)) - c(\text{supp}(\rho)),$$

where $c(\text{supp}(\rho)) \geq 0$ for all $A \in \mathcal{A}$ and $c(\varnothing) = 0$.

In the *random anticipated utility representation with thinking aversion*, "thinking aversion" is related to *preferences for randomization* in the sense that the cost of "subjective" randomization in one's mind is considered. The cost function c is monotonic. If the size of menus is large, then the cost also increases.

Proposition 8. *If two random anticipated utility representations with thinking aversion by $\langle u, (g_A)_{A \in \mathcal{A}}, c \rangle$ and $\langle u', (g'_A)_{A \in \mathcal{A}}, c' \rangle$ represent the same preference relation $\succsim$, then the following holds:*

(i) u is unique up to a positive affine transformation; there exist $a > 0$ and $b \in \mathbb{R}$ such that $u = au' + b$.

(ii) $g_A = g'_A$ for each $A \in \mathcal{A}$.

(iii) $c = ac'$.

This uniqueness result gives us a deeper understanding for the attitude toward randomization. Consider the following example. Take three alternatives $p, q, r \in \Delta(X)$ such that $\{p\} \sim \{q\} \sim \{r\}$. Suppose that $\succeq$ is represented by a three-tuple (u, g, c), i.e., a *random anticipated utility representation with thinking aversion*. Consider the following two menus, $\{p, q\}$ and $\{r\}$. Notice that $c : \mathcal{A} \to \mathbb{R}$ is a monotonic function, i.e., for all $A, B \in \mathcal{A}$ with $B \subseteq A$, $c(\text{supp}(A)) \geq c(\text{supp}(B))$. Then, if $\succeq$ exhibits the desire for randomization, then $\{p, q\} \succ \{r\}$ due to the effect of randomization in the mind. If $\succeq$ exhibits the indifference to randomization, then $\{p, q\} \sim \{r\}$. If $\succeq$ exhibits the aversion to randomization, then $\{r\} \succ \{p, q\}$ due to the cost of randomization in the mind.

5 Ex-Post Choices: Characterization of Stochastic Choices

We study the ex-post stochastic choice of RAU: $\langle u, g \rangle$. Let

$$\rho^* \in \arg\max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q))$$

be the maximizer of the *random anticipated utility representation*. In the random/stochastic choice theory, Luce (1959) characterizes random choice functions by the the axiom of *Luce's IIA* (*Independence of Irrelevant Alternatives*). For any $q \in A$, let $\rho(q|A)$ denote the probability of choosing an alternative q from a menu A. The axiom of *Luce's IIA* is stated in the following way:

Axiom (Luce's IIA): For any $p, q \in \Delta(X)$ and $A \in \mathcal{A}$ with $\{p, q\} \subseteq A$,

$$\frac{\rho(p|\{p,q\})}{\rho(q|\{p,p\})} = \frac{\rho(p|A)}{\rho(q|A)}.$$

This axiom states that the ratio of probabilities between p and q is independent from any choice sets. In general, however, the optimal probability distribution ρ^* of the *random anticipated utility representation* violates the axiom of *Luce's IIA* due to the probability-weighting function g.

5.1 Random Utility

Random Anticipated Utility Representations are not consistent with the theory *Random Utility* because it does not satisfies the axiom of *Luce's IIA*. To characterize the ex-post stochastic choice ρ^* of the *random anticipated utility representation*, we present a weaker version of the axiom of *Luce's IIA*. This axiom follows from Tversky and Russo (1969):

Axiom (IIA): For any $p, q, r, r' \in \Delta(X)$,

$$\rho(p|\{p,r\}) \geq \rho(q|\{q,r\}) \Leftrightarrow \rho(p|\{p,r'\}) \geq \rho(q|\{q,r'\}).$$

This axiom states that, from the observed "stochastic" choice data, since the probability of choosing p is larger than that of q, we can infer that the decision maker prefers an alternative p to another alternative q. In other words, in our framework, we have $\{p\} \succeq \{q\}$, and $\{p,r\} \succeq \{q,r\}$ for any $r \in \Delta(X)$.

Proposition 9. *Suppose that $\succeq$ is represented by a random anticipated utility representation by a pair $\langle u, g \rangle$. Then, the optimal probability distribution ρ^* satisfies IIA.*

The intuition of Proposition 9 is as follows. The optimal random choice is determined by a pair $\langle u, g \rangle$. By IIA, $\rho(p|\{p,r\}) \geq \rho(q|\{q,r\})$ for any $r \in \Delta(X)$ implies that $u(p) \geq u(q)$. Moreover, since g is strictly increasing, the resulting probability of choosing p is always superior to that of q.

Tversky and Russo (1969) show that the axiom of *IIA* is equivalent to the axiom of *Strong Stochastic Transitivity*:

Axiom (Strong Stochastic Transitivity): For any $p, q, r \in \Delta(X)$, $\rho(p|\{p, q\}) \geq 0.5$ and $\rho(q|\{q, r\}) \geq 0.5$ implies $\rho(x|\{x, z\}) \geq \max\{\rho(x|\{x, y\}), \rho(y|\{y, z\})\}$.

Since the primitive of the model $\succeq$ satisfies the axiom of *Transitivity*, the resulting choice of probability distributions can also satisfy the axiom of *Strong Stochastic Transitivity.* There is no menu effect related to preference reversals.

5.2 Deliberate Randomization

We mention that there is another condition in Fudenberg et al. (2015), who provide a weaker version of the axiom of *Luce's IIA*.

Axiom (Ordinal Luce's IIA): For any $p, q, r, r' \in \Delta(X)$, there exists a continuous and monotone function $f : [0, 1] \to \mathbb{R}_+ \cup \{\infty\}$ with $f(0) = 0$ such that

$$\frac{f(\rho(p|\{p, q\}))}{f(\rho(q|\{p, q\}))} = \frac{f(\rho(p|A))}{f(\rho(q|A))}.$$

This axiom states that the ratios of "re-scaled" choice probabilities are the same for any menus. If $f(\rho) = \rho$, then this axiom reduces to the axiom of *Luce's IIA*. The ex-post *optimal* stochastic choice ρ^* can satisfy a weaker version of Ordinal Luce's IIA when $f : [0, 1] \to [0, 1]$ with $f(1) = 1$.

We mention that there is another condition in Fudenberg et al. (2015) (p. 2385), which is a weaker version of *acyclic* condition, called *Weak Acyclicity.* Let $\geq^*$ be the relation on $[0, 1] \times [0, 1]$. The definition is as follows.

Definition 8. For any $\rho, \mu \in [0, 1]$,

$$\rho \geq^* \mu \Leftrightarrow \rho > \mu \text{ or } \rho = \mu \in (0, 1).$$

Definition 9. We say that a finite sequence of quadruples $\{(p_k, A_k), (q_k, B_k)\}_{k=1}^n$ is *admissible* if

(i) $p_k \in A_k$ and $q_k \in B_k$ for all k;

(ii) $(q_k)_{k=1}^n$ is a permutation of $(p_k)_{k=1}^n$; and

(iii) $(B_k)_{k=1}^n$ is a permutation of $(A_k)_{k=1}^n$.

Formally, suppose that there exists a permutation $f : \{1, \cdots, n\} \to \{1, \cdots, n\}$ such that $q_k = x_{f(k)}$. In the same way, suppose that there exists a permutation $f' : \{1, \cdots, n\} \to \{1, \cdots, n\}$ such that $B_k = A_{f'(k)}$.

Axiom (Weak Acyclicity): ρ satisfies *Weak Acyclicity* if there is no admissible sequence such that

$$\rho(p_1|A_1) > \rho(q_1|B_1), \rho(p_k|A_k) \geq^* \rho(q_k|A_k), \text{ and } \rho(p_n|A_n) \geq^* \rho(q_n|B_n).$$

We may show that the ex-post *optimal* stochastic choice ρ^* satisfies a weaker version of the axiom of *Weak Acyclicity.*

6 Consistency with Experiments

We provide numerical examples for the *random anticipated utility representation* $\langle u, g \rangle$. The model is consistent with experimental evidence of deliberate randomization by subjects.

6.1 Complements between Attributes: Easy and Hard Choices

In attribute-based inferences, if an alternative *dominates* another alternative, decision-making is not hard. We call this case an *easy* choice. On the other hand, if both alternatives are not dominated each other, decision-making can be hard. We call this case a *hard* choice (see Figure 5).

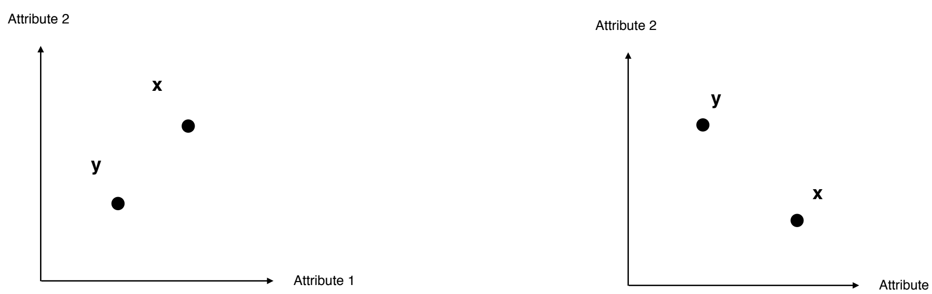

Figure 5: Easy and Hard Choices

We consider a numerical example. Suppose $X = \mathbb{R}^2$. In the left hand side (LHS), let $x = (x_1, x_2) = (10, 10)$ and $y = (y_1, y_2) = (4, 4)$. In the right hand side (RHS), let $x = (7, 3)$ and $y = (2, 8)$. Assume that a utility function is given by, for any $x \in X$, $u(x) = \log x_1 + \log x_2$. The probability-weighting function g is given by $g(\rho(x)) = \rho(x)^\alpha$, for each $x \in X$.

In the LHS of Figure 5, the singleton menu of utility is given by $V(\{x\}) = u(x) = 2$, and $V(\{y\}) = u(y) \risingdotseq 1.2$. The following table shows the optimal probability distribution and the utilities of the menu when α changes from 0.1 to 1.

In the RHS of Figure 5, the singleton menu of utility is given by $V(\{x\}) = u(x) \risingdotseq 1.32$, and $V(\{y\}) = u(y) \risingdotseq 1.20$. in the similar way, the following table shows the optimal probability distribution and the menu of utility when α changes from 0.1 to 1.

Table 1: Easy Choices

α	$\rho^*(x)$	$\rho^*(y)$	$V(\{x,y\})$
0.1	0.6373	0.3627	2.995
0.2	0.6535	0.3465	2.806
0.3	0.6737	0.3263	2.633
0.4	0.6997	0.3003	2.474
0.5	0.7340	0.2660	2.331
0.6	0.7805	0.2195	2.206
0.7	0.8444	0.1556	2.102
0.8	0.9267	0.0733	2.02
0.9	0.9938	0.0062	2.0003
1	1	0	2

Table 2: Hard Choices

α	$\rho^*(x)$	$\rho^*(y)$	$V(\{x,y\})$
0.1	0.5260	0.4740	2.350
0.2	0.5292	0.4708	2.270
0.3	0.5334	0.4666	2.04
0.4	0.5389	0.4611	1.910
0.5	0.5466	0.4534	1.783
0.6	0.5582	0.4418	1.665
0.7	0.5773	0.4227	1.554
0.8	0.6149	0.3851	1.453
0.9	0.7182	0.2818	1.362
1	1	0	1.32

As the two tables shows, compared with easy choices, deliberate randomization can occur in hard choices. The attitude toward randomization is described by α. As α decreases, the optimal random choice becomes more random and deliberate.

In general, the attitude toward "subjective" randomization is menu-dependent (see Claim 3). In easy choices, decision makers may exhibit the aversion to randomization; that is, they may choose the *superior* alternative with certainty, i.e., $\rho^*(x|\{x,y\}) = 1$.

No Preference Reversal: the Attraction and Compromise effects

We mention that, even in the case of menu-dependent probability-weighting functions, the RAU is not consistent with preference reversals such as the Attraction effect (Huber et al. (1982)) and the Compromise effect (Simonson (1989)). The optimal random choice depends on the utility of alternatives (lotteries), and the ratio of probabilities of choosing alternatives only changes. The probability-weighting function g does not matter. There is another cognitive mechanism behind the Attraction effect and the Compromise effect, and the mechanism is different from preferences for randomization.

6.2 Agranov and Ortoleva (2017): Lottery Choices

We consider two "similar" tasks of lottery choices in Agranov and Ortoleva (2017). We provide an easy task and a hard task. Suppose that X is a finite set of all *prizes*. First, in the easy task, a lottery is a degenerate lottery which gives \$10 with certainty. Another lottery gives \$2 with probability $\frac{4}{5}$, and \$12 with probability $\frac{1}{5}$. Let us denote the former lottery by p, the latter by q, respectively. Next, in the hard task, a lottery gives \$1 with probability $\frac{1}{2}$, and \$9 with probability $\frac{1}{2}$. Another lottery gives \$6 with probability $\frac{1}{2}$, and gives \$3 with probability $\frac{1}{2}$. For each prize $x \in X$, suppose $g(\rho(x)) = \sqrt{\rho(x)}$. We consider the two vNM functions. One is risk neutral, i.e., $v(x) = x$ for each $x \in X$. The other is risk averse, i.e., $v(x) = \log x$ for each $x \in X$. Let us denote the former lottery by p, the latter by q, respectively.

Table 3: Lottery Choices

Tasks	$\rho^*(p)$	$\rho^*(q)$	$V(\{p,q\})$
Easy (Risk Neutral)	0.8621	0.1379	10.769
Easy (Risk Averse)	0.7353	0.2647	1.166
Hard (Risk Neutral)	0.5525	0.4475	6.726
Hard (Risk Averse)	0.5370	0.4630	0.955

7 Literature Review: Comparison with Other Models

We provide an overview of related literature. The study of *preferences for randomization* is mainly categorized into three topics: (i) learning (preferences for flexibility), (ii) attention (bounded rationality), and (iii) deliberate randomization. Since this paper is related to *deliberate randomization*, we focus on to refer to the literature in this topic.[17]

Fudenberg et al. (2015): Implementation Costs

Fudenberg et al. (2015) present a theory of *deliberate randomization* in which the decision maker maximize the expected utility on the alternatives on a menu with a cost function. The stochastic choice function corresponds to the maximization of the sum of expected utility and a perturbation function

$$P(A) := \arg \max_{\rho \in \Delta(A)} \sum_{x \in A} u(x)\rho(x) - c(\rho(x)),$$

where $P(A)$ is the probability distribution of choices from a given menu A, $u : X \to \mathbb{R}$ is the utility function of the decision maker, and $c : [0,1] \to \mathbb{R}$ is a convex perturbation function that can reward the decision maker for deliberate randomization. This utility representation is called a weak Additive Perturbed Utility (APU) representation.

The primitive of the two papers are different. The primitive of Fudenberg et al. (2015) is a stochastic choice function, but the primitive of this paper is a binary relation $\succeq$ on $\mathcal{A}$, i.e., the set of all menus. The key procedural aspect of Fudenberg et al. (2015) is the cost function for subjective randomization. Since the cost function c is strictly convex, their model is presumed that the decision maker has a preference for the desire for randomization; that is, the model corresponds to the case that g is concave in this paper.

We consider a special case. Suppose that $\succeq$ is represented by a pair $\langle u, (g_A)_{A \in \mathcal{A}} \rangle$. Assume that, for each $A \in \mathcal{A}$, g_A is *strictly concave*. Given a menu A, for any $q \in A$, define

$$c_q(\rho(q)) := u(q)[g_A(\rho(q)) - \rho(x)]$$

Then, c is strictly convex, and we can show that c^1 over $(0,1)$. Hence, we have

$$\rho^* \in \arg \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) - c_q(\rho(q)).$$

This form is related to a weaker version of APU in Fudenberg et al. (2014) (Section 5.2).[18] In the menu-invariant additive perturbed utility (Menu-Invariant APU) representation, the

[17] See Kreps (1979) and Dekel et al. (2001) for the framework of *subjective state spaces*. For example, there are Gul and Pesendorfer (2006), and Ahn and Sarver (2013) for the extension for random utility. See Cerreia-Vioglio et al. (2017) about the literature of a relationship between bounded rationality and random choices in detail.

[18] I would like to thank Ryota Iijima for his comment.

cost function depends on the alternatives in menus, but it is invariant with respect to menus. The Menu-Invariant APU is characterized by the following axiom.

Axiom (Menu Acyclicity): If $\rho(p_1|A_1) > \rho(p_1|A_2), \rho(p_k|A_k) \geq^* \rho(p_k|A_{k+1})$ for $a < k < n$, then $\rho(p_n|A_n) \not\geq^* \rho(p_n|A_1)$.

The axiom of *Weak Acyclicity* imposes on a ranking on menus. The axiom of *Menu Acyclicity* can violate the axiom of *Strong Stochastic Transitivity*, but it satisfies the axiom of *Weak Stochastic Transitivity*.

Axiom (Weak Stochastic Transitivity): For any $p, q, r \in \Delta(X)$, $\rho(p|\{p, q\}) \geq 0.5$ and $\rho(q|\{q, r\}) \geq 0.5$ implies $\rho(p|\{p, r\}) \geq 0.5$.

In Proposition 9, if the probability-weighting function g is menu-independent, then the ex-post stochastic choice ρ^* satisfies the axiom of *IIA*. Hence, ρ^* satisfies the axiom of *Strong Stochastic Transitivity*. To compare RAU with (weak) APU, we relax the probability-weighting function g (Claim 3). With the menu-dependent probability functions $(g_A)_{A \in \mathcal{A}}$, the *random anticipated utility representation* can violate the axiom of *IIA*.

Cerreia-Vioglio et al. (2017): Allais-Type Preferences

Cerreia-Vioglio et al. (2017) develops a theory of *deliberate randomization* that is an extension of *Cautious Expected Utility* (Cautious EU) in Cerreia-Vioglio et al. (2015). Let $[w, b] \subset \mathbb{R}$ be an interval of monetary prizes. Let Δ be the set of lotteries (Borel probability measures) over $[w, b]$, endowed with the topology of weak convergence.[19] Denote the set of continuous functions from $[w, b]$ into $\mathbb{R}$ by $C([w, b])$ and metrize it by the supnorm. A stochastic choice function ρ admits a *Cautious Stochastic Choice Representation* if there exists a compact set $\mathcal{W} \subseteq C([w, b])$ such that every function $v \in \mathcal{W}$ is strictly increasing and concave, and

$$\rho(A) \in \arg \max_{p \in \mathrm{co}(A)} \min_{v \in \mathcal{W}} v^{-1}(\mathbb{E}_p(v)),$$

for every $A \in \mathcal{A}$.

The key procedural aspect of Cerreia-Vioglio et al. (2017) is that *deliberate randomization* depends on menu. In a menu, the decision maker exhibits the *desire for randomization*. However, in another menu, the same decision maker exhibits an *aversion to randomization*. On the other hand, in this paper, we identify the attitude toward randomization by g. To allow for menu-dependent probability-weighting functions, we need to relax the axiom of *Randomization*. Or, in Section 4, we provide additional axioms such as the axiom of *Monotonicity*

[19]This case corresponds to $X \subseteq \mathbb{R}^n$ where $n = 1$.

(*preferences for flexibility*). Then, we allow for multiple probability-weighting functions dependent on subjective states.

To compare this paper with Cerreia-Vioglio et al. (2017), we consider menu-dependent probability-weighting functions $(g_A)_{A\in\mathcal{A}}$ (Claim 3). And, we study the following axiom.

Axiom (Negative Certainty Independence (NCI)): For any $p, q \in \Delta(X)$, $x \in X$, and $\lambda \in [0,1]$,

$$\{p\} \succeq \{\delta_x\} \Rightarrow \lambda\{p\} + (1-\lambda)\{q\} \succeq \lambda\{\delta_x\} + (1-\lambda)\{q\}.$$

The axiom of *Negative Certainty Independence* says that if a lottery (alternative) is preferred to a degenerate lottery, then the axiom of *Strong Singleton Independence* holds. This axiom is a weaker version of *Strong Singleton Independence*.

The model of *Cautious Stochastic Choice* is stated as follows. There exists a pair $\langle \mathcal{W}, (g_A)_{A\in\mathcal{A}} \rangle$ where $\mathcal{W}$ is the set of continuous, strictly increasing, and concave functions, and $(g_A)_{A\in\mathcal{A}}$ is a *menu-dependent* probability-weighting function denoted by $g_A : [0,1] \to [0,1]$ satisfying $g(0) = 0$ and $g(1) = 1$ such that the ex-post stochastic choice is defined by

$$\rho(A) \in \arg\max_{\rho\in\Delta(A)} \Big[\sum_{q\in A} \Big(\min_{v\in\mathcal{W}} v^{-1}(\mathbb{E}_q(v))\Big) g_A(q)\Big].$$

Deliberate Randomization

In sum, we provide a summary in Figure 7.

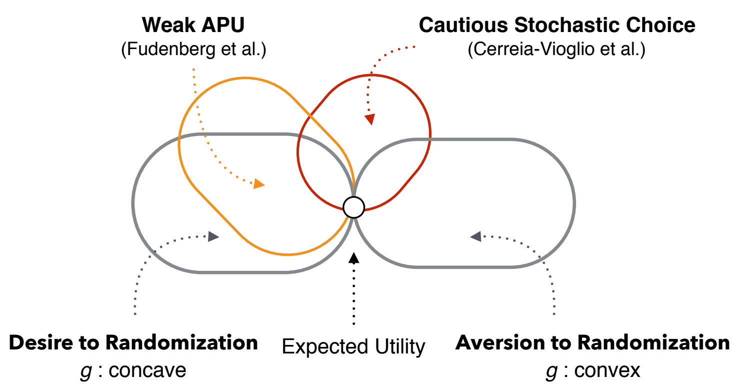

Figure 6: A Summary

Other Literature

In Gul and Pesendorfer (2006), *random expected utility* (Random EU) is studied.[20] They present a theory of random choices of lotteries, which is restricted to vNM expected utility preferences. The model states that the decision makers have a probability distribution on expected utility functions with finite support. When decision makers make choices, they maximize a well-defined preference, but the taste changes randomly over time due to changes in unobservable conditions such as endogenous information, mood, feeling, and so on. To compare this paper with Gul and Pesendorfer (2006), we need to relax the axiom of *Randomization* (see Section 4.1). The key axiom of Gul and Pesendorfer (2006) is called *Linearity*:

Axiom (Linearity): For any $A \in \mathcal{A}$ and $p \in A$,

$$\rho(p|A) = \rho(\lambda p + (1-\lambda)q|\lambda A + (1-\lambda)\{q\}),$$

for any $q \in \Delta(X)$ and $\lambda \in (0,1)$.

The axiom of *Linearity* is analogous to the axiom of *Independence* in the expected utility theorem. Consider a pair $(u, \mathcal{G})$ where $u : \Delta(X) \to \mathbb{R}$ is a non-constant function, and $\mathcal{G}$ is a set of state-dependent probability-weighting functions, i.e., $g : [0,1] \times S \to \mathbb{R}$ is a continuous and strictly increasing function where $g(0) = 0$ and $g(1) = 1$ where S is a subjective state space. Then, we impose on the axiom of *Independence*. Hence, each g_s is linear. This case corresponds to Random EU.

Saito (2015) presents a seminal study of *preferences for randomization*. Saito (2015) axiomatizes preferences for randomization from "deterministic" preferences over sets of Anscombe and Aumann acts. The motivation of his paper is different from that of this paper.[21] In his model, the *uncertainty-averse* decision maker may randomize alternatives (*acts*) to eliminate the effect of uncertainty. Saito (2015) shows that, in his framework, a stronger preference for flexibility is equivalent to a stronger preference for randomization, behaviorally. To apply the idea of Saito (2015) into this paper, consider a different subjective belief of *deliberate randomization*. The decision maker may be uncertain about the effect of *deliberate randomization*. To capture this procedural aspect, consider the following axiom, instead of the axiom of *Randomization*.

Axiom (Dominance): For any $A, B \in \mathcal{A}$, suppose that the following statements hold:

[20] See Ahn and Sarver (2013), who incorporate the decision maker's random choice from menus by using the framework of Dekel et al. (2001) and the random expected utility (Gul and Pesendorfer (2006)). It is shown that both preferences for flexibility in the first period and random choices from the menus chosen in the first period can be, respectively, rationalized by *subjective state spaces*.

[21] Classically, Raiffa (1961) suggests that ambiguity aversion may exhibit a preference for randomization.

(i) for any $\mu \in \Delta(B)$, there exists $\rho \in \Delta(A)$ such that $\rho \trianglerighteq_{\text{dom.}} \mu$; and

(ii) for any $q \in B$, there exists $p \in A$ such that $\{p\} \succeq \{q\}$.

Then, $A \succeq B$.

We can obtain the following corollary.

Corollary 6. $\succeq$ *satisfies Standard Preferences, Dominance, and Strong Singleton Independence, if and only if there exists a triple* $\langle u, g, \alpha \rangle$ *where* $u : \Delta(X) \to \mathbb{R}$ *is a non-constant function,* $g : [0,1] \to [0,1]$ *is a probability-weighting function, and* $\alpha \in [0,1]$ *is a parameter, such that* $\succeq$ *is represented by* $V : \mathcal{A} \to \mathbb{R}$ *defined by*

$$V(A) = \max_{\rho \in \Delta(A)} \Big[\alpha \Big(\sum_{q \in A} u(q) g(\rho(q)) \Big) + (1-\alpha) \Big(\sum_{q \in A} u(q) \rho(q) \Big) \Big],$$

where $\Delta(A)$ *is a probability distribution over* A.

Depending on $\alpha \in [0,1]$, the ex-post stochastic choice can be changeable. When $\alpha = 1$, the utility representation corresponds to a *random anticipated utility representation.* On the other hand, when $\alpha = 0$, the utility representation corresponds to the result in Kreps (1979) that $\succeq$ satisfies the axiom of *Strategic Rationality.*

Notice that preferences for flexibility itself does not have any properties of a *desire* or an *aversion* to randomization. Behaviorally, resulting choice behaviors are stochastic. On the other hand, to capture preferences for randomization, this paper pays much attention to weaker versions of the axioms of *Strategic Rationality* and *Independence* rather than *Monotonicity* (*preferences for flexibility*).

8 Concluding Remarks

In this paper, we provide an axiomatic foundation for a class of *preferences for randomization* (Section 2) by classifying the *desire to randomization*, the *indifference to randomization*, and the *aversion to randomization*, respectively (Section 3). To study preferences for randomization, we have used the framework of *subjective state spaces* (Kreps (1979)/Dekel et al. (2001)), and have identified risk attitudes and randomization attitudes separately. Especially, by relaxing the axioms of *Strategic Rationality* and *Independence*, we provide new axioms called *Randomization* and *Strong Singleton Independence.* Moreover, we provide an axiomatic characterization of ex-post stochastic choices, which makes it clear to the link between this paper and random utility theory (Section 5).

To study general cases, we have relaxed the axioms of *Randomization* and *Strong Singleton Independence.* We have imposed on the axiom of *Monotonicity* to consider multiple attitudes toward the effect of randomization (Section 4).[22] In addition, this paper is easily applicable for a theory of subjective learning (Dillenberger et al. (2014)) and a theory of thinking aversion (Ortoleva (2013)). By extending the domain and adding the axiom in Dillenberger et al. (2014), we have studied the relationship between *preferences for randomization* and *preferences for delay*. The theory of *thinking aversion* is also applied, and the theory is helpful for an identification of the attitude toward *deliberate randomization*.

There are some further tasks. First, the study on deliberate randomization is lack of both experimental and empirical evidence on preferences for randomization. Moreover, risk attitudes and randomization attitudes may be correlated. The relationship should be empirically studied. Second, randomization attitudes may change dynamically. For example, status-quo bias may lead to the aversion to randomization. On the other hand, falling into a rut in intertemporal choices may lead to the desire for randomization. Such aspects should be studied in the theory of intertemporal choices.

[22]Note that multiple attitudes toward randomization are not uniquely identified generally, even though multiple attitudes are intuitive.

A Proof of Theorem

A.1 Sufficiency Part

We show the sufficiency part. Suppose that $\succeq$ satisfies the axioms in Theorem 1: *Standard Preferences*, *Randomization*, and *Singleton Independence*.

STEP 1

We construct a function $V : \mathcal{A} \to \mathbb{R}$. We show that, for any $A, B \in \mathcal{A}$, $A \succeq B \Leftrightarrow V(A) \geq V(B)$.

Lemma 1. *There exists a function* $V : \mathcal{A} \to \mathbb{R}$ *such that*

$$V(A) \geq V(B) \Leftrightarrow A \succeq B$$

for any $A, B \in \mathcal{A}$.

Proof. Let $u : \Delta(X) \to \mathbb{R}$ is a non-constant function. Define $V(\{p\}) = u(p)$ for any $p \in \Delta(X)$. Since $\mathcal{A}$ is compact, by the Berge's Maximum Theorem, u is continuous (see Aliprantis and Border (2006) (p.570)). Fix a menu $A \in \mathcal{A}$. And, notice that A is compact. By the Weierstrass's Theorem, there exists $\overline{p}, \underline{p} \in \mathrm{co}(A)$ such that $\overline{p} \in \arg\max_{p \in \mathrm{co}(A)} u(p)$ and $\underline{p} \in \arg\min_{p \in A} u(p)$ (see Aliprantis and Border (2006) (p.40)).

By the axiom of *Randomization*, we have $\{\overline{p}\} \succeq A \succeq \{\underline{p}\}$. If $\{\overline{p}\} \sim A$, then we define $V(A) = u(\overline{p})$. If $A \sim \{\underline{p}\}$, then we define $V(A) = u(\underline{p})$. If $\{\overline{p}\} \succ A \succ \{\underline{p}\}$, then, by the axiom of *Continuity*, there exists a unique $\lambda \in (0,1)$ such that $A \sim \lambda\{\overline{p}\} + (1-\lambda)\{\underline{p}\}$. Define $V(A) = u(\lambda\overline{p} + (1-\lambda)\underline{p})$. Hence, for any $A, B \in \mathcal{A}$, $A \succeq B \Leftrightarrow V(A) \geq V(B)$. $\square$

STEP 2

The main objective of this step is to elicit a pair $\langle u, \theta \rangle$ where a *cardinal* utility function $u : \Delta(X) \to \mathbb{R}$ (see STEP 1) and a capacity θ^* on $\mathcal{A}^* \subset \mathcal{A}$ separably and uniquely to evaluate binary menus (*doubletons*).[23] In STEP 2, we apply bi-separable preferences (*preference average*) into our framework (Ghirardato and Marinacci (2001)).

First, we show the following lemma. To do so, we present the following axiom.

Axiom ($\frac{1}{2}$-Average): For any $p, q \in \Delta(X)$ such that $\{\delta_{c_p}\} \succeq \{\delta_{c_q}\}$, there exists $r \in \Delta(X)$ where $\{\delta_{c_p}\} \succeq \{\delta_{c_r}\} \succeq \{\delta_{c_q}\}$ such that

$$\frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \frac{1}{2}\left\{\frac{1}{2}\delta_{c_p} + \frac{1}{2}\delta_{c_r}\right\} + \frac{1}{2}\left\{\frac{1}{2}\delta_{c_q} + \frac{1}{2}\delta_{c_r}\right\}.$$

[23] $\mathcal{A}^*$ is the set of *doubleton* menus.

Lemma 2. *Suppose that $\succeq$ satisfies the axioms in Theorem 1. Then, $\succeq$ satisfies $\frac{1}{2}$-Average.*

Proof. Take $p, q \in \Delta(X)$ with $\{p\} \succeq \{q\}$. By the axiom of *Singleton Independence*, for any $r \in \Delta(X)$, we have $\frac{1}{2}\{p\} + \frac{1}{2}\{r\} \succeq \frac{1}{2}\{q\} + \frac{1}{2}\{r\} \Leftrightarrow \{\frac{1}{2}p + \frac{1}{2}r\} \succeq \{\frac{1}{2}q + \frac{1}{2}r\}$. By the axiom of *Continuity*, there exists $r \in \Delta(X)$ such that $\frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \{r\} \Leftrightarrow \{\frac{1}{2}p + \frac{1}{2}q\} \sim \{r\}$. Then, we have $\{r\} \sim \frac{1}{2}\{\frac{1}{2}p + \frac{1}{2}q\} + \frac{1}{2}\{r\} \Leftrightarrow \{\frac{1}{2}p + \frac{1}{2}q\} \sim \{\frac{1}{4}p + \frac{1}{4}q + \frac{1}{2}r\} \Leftrightarrow \frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \frac{1}{2}\Big\{\frac{1}{2}p + \frac{1}{2}r\Big\} + \frac{1}{2}\Big\{\frac{1}{2}q + \frac{1}{2}r\Big\} \Leftrightarrow \frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \frac{1}{2}\Big\{\frac{1}{2}\delta_{c_p} + \frac{1}{2}\delta_{c_r}\Big\} + \frac{1}{2}\Big\{\frac{1}{2}\delta_{c_q} + \frac{1}{2}\delta_{c_r}\Big\}$. $\square$

Lemma 3. *$\succeq$ satisfies Standard Preferences and $\frac{1}{2}$-Average if and only if for any $p, q \in \Delta(X)$, there exists $r \in \Delta(X)$ where $\{p\} \succeq \{r\} \succeq \{q\}$ such that*

$$V(\{r\}) = \frac{1}{2}V(\{p\}) + \frac{1}{2}V(\{q\}).$$

Proof. Take arbitrary $p, q \in \Delta(X)$ such that $\{p\} \succeq \{q\}$. By the axiom of *Continuity*, there exists $r \in \Delta(X)$ such that $\{p\} \succeq \{r\} \succeq \{q\}$ and $\frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \frac{1}{2}\{\frac{1}{2}p + \frac{1}{2}r\} + \frac{1}{2}\{\frac{1}{2}q + \frac{1}{2}r\}$. By Step 1, we have $V(\{p\}) \geq V(\{r\}) \geq V(\{q\})$. And, we have $V(\frac{1}{2}\{p\} + \frac{1}{2}\{q\}) = V(\frac{1}{2}\{\frac{1}{2}p + \frac{1}{2}r\} + \frac{1}{2}\{\frac{1}{2}q + \frac{1}{2}r\})$.

Let $\gamma \in [0, 1]$. By the axiom of $\frac{1}{2}$-*Average*, we have $V(\{p, q\}) = u(p)\gamma + u(q)(1 - \gamma)$, $V(\{p, r\}) = u(p)\gamma + u(r)(1 - \gamma)$, and $V(\{q, r\}) = u(r)\gamma + u(q)(1 - \gamma)$.

Since $\{p\} \succeq \{r\} \succeq \{q\}$, we have $\{\frac{1}{2}p + \frac{1}{2}r\} \succeq \{\frac{1}{2}q + \frac{1}{2}r\}$. Thus, we have

$$\begin{aligned} V\Big(\frac{1}{2}\Big\{\frac{1}{2}p + \frac{1}{2}r\Big\} + \frac{1}{2}\Big\{\frac{1}{2}q + \frac{1}{2}r\Big\}\Big) &= u\Big(\frac{1}{2}p + \frac{1}{2}r\Big)\gamma + u\Big(\frac{1}{2}q + \frac{1}{2}r\Big)(1 - \gamma) \\ &= V\Big(\Big\{\frac{1}{2}p + \frac{1}{2}r\Big\}\Big)\gamma + V\Big(\Big\{\frac{1}{2}q + \frac{1}{2}r\Big\}\Big)(1 - \gamma) \\ &= [u(p)\gamma + u(r)(1 - \gamma)]\gamma + [u(r)\gamma + u(q)(1 - \gamma)](1 - \gamma) \\ &= u(p)\gamma^2 + u(q)\gamma^2 + 2u(r)\gamma(1 - \gamma). \end{aligned}$$

Remember that $r \in \Delta(X)$ satisfies $\{p\} \succeq \{r\} \succeq \{q\}$ and $\frac{1}{2}\{p\} + \frac{1}{2}\{q\} \sim \frac{1}{2}\{\frac{1}{2}p + \frac{1}{2}r\} + \frac{1}{2}\{\frac{1}{2}q + \frac{1}{2}r\}$. That is, we obtain

$$\begin{aligned} &V\Big(\frac{1}{2}\{p\} + \frac{1}{2}\{q\}\Big) = V\Big(\frac{1}{2}\Big\{\frac{1}{2}p + \frac{1}{2}r\Big\} + \frac{1}{2}\Big\{\frac{1}{2}q + \frac{1}{2}r\Big\}\Big) \\ &\Leftrightarrow u(p)\gamma + u(q)(1 - \gamma) = u(p)\gamma^2 + u(q)\gamma^2 + 2u(r)\gamma(1 - \gamma) \\ &\Leftrightarrow u(p)\gamma(1 - \gamma) + u(q)\gamma(1 - \gamma) = 2u(r)\gamma(1 - \gamma) \end{aligned}$$

Hence, we have $u(r) = \frac{1}{2}u(p) + \frac{1}{2}u(q)$, i.e., $V(\{r\}) = \frac{1}{2}V(\{p\}) + \frac{1}{2}V(\{q\})$. $\square$

Next, we consider *dyadic rational*, a number $\gamma \in (0, 1)$ such that for some finite integer $M \in \mathbb{N}$,

$$\gamma = \sum_{i=1}^{M} \frac{a_i}{2^i},$$

where $a_i \in \{0,1\}$ for every $i \in \{1, \cdots, M\}$ and $a_M = 1$. By the axiom of $\frac{1}{2}$-*Average*, we can consider the following preference average:

$$\frac{1}{2}\{r_1\} \oplus \frac{1}{2}\Big(\cdots\Big(\frac{1}{2}\{r_{M-1}\} \oplus \frac{1}{2}\Big(\frac{1}{2}\{r_M\} \oplus \frac{1}{2}\{q\}\Big)\Big)\cdots\Big),$$

where for any $i \in \{1, 2, \cdots, M\}$, $r_i = p$ if $a_i = 1$, $r_i = q$ otherwise.

We show the following lemma of dyadic rational.

Lemma 4. *Let $p, q \in \Delta(X)$ and γ be a dyadic rational. Then, we have*

$$V(\{\gamma p + (1-\gamma)q\}) = \gamma V(\{p\}) + (1-\gamma)V(\{q\}).$$

Proof. Let $\gamma = \sum_{i=1}^{\infty} \frac{a_i}{2^i}$. We define $l(\gamma) = \max\{i \geq 1 \mid a_i = 1\}$. We prove this lemma by induction on $l(\gamma)$. If $l(\delta) = 1$, then $\gamma = \frac{1}{2}$ by Lemma 3.

Suppose that the result holds for all γ such that $l(\gamma) \leq m$. Let $\delta = \sum_{i=1}^{m+1} \frac{a_i}{2^i}$. Then, we have $\gamma = \frac{1}{2}a_1 + \frac{1}{2}\delta$. Thus, we obtain

$$\begin{aligned} V\Big(\{\gamma p + (1-\gamma)q\}\Big) &= V\Big(\frac{1}{2}r_1 + \frac{1}{2}(\delta x + (1-\delta)q)\Big) \\ &= \frac{1}{2}V(\{r_1\}) + \frac{1}{2}\Big(\delta V(\{p\}) + (1-\delta)V(\{q\})\Big). \end{aligned}$$

If $a_1 = 0$, then $r_1 = 1$, or if $a_1 = 1$, then $r_1 = x$. Therefore, we obtain the following: if $a_1 = 0$, then $V\Big(\{\gamma p + (1-\gamma)q\}\Big) = \frac{1}{2}\delta V(\{p\}) + (1 - \frac{1}{2}\delta)V(\{q\})$. On the other hand, if $a_1 = 0$, then $(\frac{1}{2} + \frac{1}{2}\delta)V(\{p\}) + \frac{1}{2}(1-\delta)V(\{q\})$. Thus, it is shown that $V\Big(\{\gamma p + (1-\gamma)q\}\Big) = \gamma V(\{p\}) + (1-\gamma)V(\{q\})$. □

Finally, in Step 2, we show the following lemma. Remeber that $\mathcal{A}^*$ be the set of all *binary menus*. Any binary menus are represented by a pair $\langle u, \theta^* \rangle$ where u is a *cardinal* utility function, and θ^* is a *capacity* on binary menus (doubleton menus).

Lemma 5. *$\succeq$ satisfies Standard Preferences and Randomization if and only if there exists a non-constant utility function $u : \Delta(X) \to \mathbb{R}$ and a capacity $\theta^* : \mathcal{A}^* \to [0,1]$ such that $\succeq$ is represented by $V : \mathcal{A}^* \to \mathbb{R}$, for any $p, q \in \Delta(X)$ where $\{p\} \succeq \{q\}$ such that*

$$V(\{p,q\}) = u(p)\theta^*(\{p\}) + u(q)\theta^*(\{q\}).$$

Proof. By Lemma 1, for any $p \in \Delta(X)$, $V(\{p\}) = u(p)$. By the compactness of $\Delta(X)$, there exists $\overline{p}, \underline{p} \in \Delta(X)$ such that $\{\overline{p}\} \succeq A \succeq \{\underline{p}\}$ for any $A \in \mathcal{A}$. Without loss of generality, we normalize $V(\{\overline{p}\}) = 1$ and $V(\{\underline{p}\}) = 0$. Suppose that $\{\overline{p}\} \succeq \{p\} \succeq \{q\} \succeq \{\underline{p}\}$. Then, we show the following:

$$V(\{p,q\}) = V(\{q\}) + [V(\{p\}) - V(\{q\})]V(\{\lambda\overline{p} + (1-\lambda)\underline{p}\}),$$

for some $\lambda \in [0,1]$.

Take $\alpha \in (0,1]$ and $\beta \in [0,1)$ such that $\{p\} \sim \{\alpha\overline{p} + (1-\alpha)q\}$ and $\{q\} \sim \{\beta\overline{p} + (1-\beta)\underline{p}\}$. Then, by the normalization, we have $V(\{q\}) = \beta$, and $V(\{p\}) = \alpha + (1-\alpha)V(\{q\})$. Thus, we have $V(\{p\}) - V(\{q\}) = \alpha(1 - V(\{q\})) = \alpha(1-\beta)$.

We have $V(\{p,q\}) = V(\{\alpha\overline{p} + (1-\alpha)q\}) + (1-\alpha)V(\{q\}) = \alpha V(\{\alpha\overline{p} + (1-\alpha)q\}) + (1-\alpha)V(\{q\})$ by Lemma 4. Hence,

$$\begin{aligned}
V(\{p,q\}) &= \alpha V(\{\alpha\overline{p} + (1-\alpha)(\beta\overline{q} + (1-\beta)\underline{p})) + (1-\alpha)V(\{q\}) \\
&= \alpha V(\{\beta\overline{p} + (1-\beta)(\alpha\overline{p} + (1-\alpha)\underline{p})\}) + (1-\alpha)V(\{q\}) \\
&= \alpha\beta V(\{\overline{p}\}) + \alpha(1-\beta)V(\{\alpha\overline{p} + (1-\alpha)\underline{p}\}) + (1-\alpha)V(\{q\}) \\
&= \alpha V(\{q\}) + \alpha(1-\beta)V(\{\alpha\overline{p} + (1-\alpha)\underline{p}\}) + (1-\alpha)V(\{q\}) \\
&= V(\{q\}) + \alpha(1-\beta)V(\{\alpha\overline{p} + (1-\alpha)\underline{p}\}) \\
&= V(\{q\}) + [V(\{p\}) - V(\{q\})]V(\{\alpha\overline{p} + (1-\alpha)\underline{p}\}).
\end{aligned}$$

Consider a binary menu $\{p,q\}$ such that $\{p\} \succeq \{q\}$. We define θ^* as follows. If $\delta_{c_p} > \delta_{c_q}$, then $\theta^*(\{p\}) = 1$, and $\theta(\{q\}) = 0$. If $\delta_{c_p} < \delta_{c_q}$, then $\theta^*(\{p\}) = 0$, and $\theta^*(\{q\}) = 1$. Otherwise, $\theta^*(\{p\}) = V(\{\alpha\overline{p} + (1-\alpha)\underline{p}\})$. Hence, we obtain $V(\{p,q\}) = u(p)\theta^*(\{p\}) + u(q)\theta^*(\{q\})$ where $\theta^*(\{q\}) = 1 - \theta^*(\{p\})$. □

It is shown that a utility of doubletons (binary menus) is a weighted average of the two alternatives.

STEP 3

We induce a binary relation $\trianglerighteq^*$ on $\mathcal{A}$, a ranking of deliberate randomization. We show that for all $A, B \in \mathcal{A}$, $A \trianglerighteq^* B \Leftrightarrow \theta(A) \geq \theta(B)$. Moreover, we show that θ is *weakly additive*. Finally, we show that for any $A \in \mathcal{A}$, $\theta(A) = g \circ \rho(A)$ for some continuous and strictly increasing function $g : [0,1] \to [0,1]$ satisfying $g(0) = 0$ and $g(1) = 1$.

We write $\rho \trianglerighteq_{\text{dom}} \mu$ if ρ *dominates* μ. Define, for any $A, B \in \mathcal{A}$,

$$A \trianglerighteq^* B$$

if for any $\mu \in \Delta(B)$ there exists $\rho \in \Delta(A)$ $\rho \trianglerighteq_{\text{dom}} \mu$. A set function $\theta : \mathcal{A} \to [0,1]$ is said to be a *measure* if (i) $B \subseteq A \Rightarrow \theta(B) \leq \theta(A)$, and (ii) $\theta(\varnothing) = 0$ and $\theta(X) = 1$.

Definition 10. We say that a measure θ agrees with a binary relation $\trianglerighteq^*$ if and only if $\theta(A) \geq \theta(B) \Leftrightarrow A \trianglerighteq^* B$.

Moreover, let us define the following property of capacities. We say that a capacity θ is *locally convex-valued* if the following holds:

$$B \subseteq A \text{ and } r \in [\theta(B), \theta(A)] \Rightarrow \exists C \subset X \text{ with } B \subset C \subset A \text{ such that } \theta(C) = r.$$

In addition, we introduce the following property of θ.

Definition 11. We say that a capacity θ is *weakly additive* if, for any $A, B, C, C' \subseteq X$, $C \subseteq A \cap B$, $C' \subseteq (A \cup B)^c$,

$$\theta((A \setminus C) \cup C') > \theta((B \setminus C) \cup C') \Rightarrow \theta(A) > \theta(B).$$

First, we show that θ agrees with a binary relation $\trianglerighteq^*$. Take $A, B \in \mathcal{A}$ with $A \trianglerighteq^* B$. Then, by the axiom of *Randomization*,

$$\begin{aligned} A \trianglerighteq^* B &\Leftrightarrow \rho \trianglerighteq_{\text{dom}} \mu \\ &\Leftrightarrow A \succeq B \\ &\Leftrightarrow V(A) \geq V(B) \\ &\overset{\text{def.}}{\Leftrightarrow} \theta(A) \geq \theta(B) \end{aligned}$$

by $\theta : \mathcal{A} \to [0, 1]$ satisfying (i) $B \subset A \Rightarrow \theta(B) \leq \theta(A)$, and (ii) $\theta(\varnothing) = 0$ and $\theta(X) = 1$.

Next, we show that the capacity θ that agrees with $\trianglerighteq^*$ satisfies *weak additivity*. Take $A, B, C \in \mathcal{A}$ with $A \cap B \cap C = \varnothing$, and $\rho \in \Delta(A), \mu \in \Delta(B)$, and $\gamma \in \Delta(C)$. Then, $A \cap B = \varnothing$, and $C \in (A \cup B)^c$ hold. Take $\lambda \in (0, 1]$. Fix A, B, C, and λ. Suppose $\lambda\rho + (1-\lambda)\gamma \trianglerighteq_{\text{dom.}} \lambda\mu + (1-\lambda)\gamma$. Then, $\lambda A + (1-\lambda)C \succeq \lambda B + (1-\lambda)C \Leftrightarrow V(\lambda A + (1-\lambda)C) \geq V(\lambda B + (1-\lambda)C) \Leftrightarrow \theta(\lambda A + (1-\lambda)C) \geq \theta(\lambda B + (1-\lambda)C) \Leftrightarrow \lambda A + (1-\lambda)C \trianglerighteq^* \lambda B + (1-\lambda)C$. This holds for any arbitrary $\lambda \in (0, 1]$. Hence, $A \trianglerighteq^* B$. We obtain $\theta(A) \geq \theta(B)$.

Consider $A, B \in \mathcal{A}$ with $A \cap B = \varnothing$. Then, take $r \in \Delta(X) \setminus (A \cup B)$. Suppose $\theta(A \cup \{r\}) > \theta(B \cup \{r\})$. By definition, for any $\widehat{\mu} \in \Delta(B \cup \{r\})$, there exists $\widehat{\rho} \in \Delta(A \cup \{r\})$ such that $\widehat{\rho} \trianglerighteq_{\text{dom.}} \widehat{\mu}$. We can find that there exist $\rho \in \Delta(A)$ and $\mu \in \Delta(B)$ such that $\rho \trianglerighteq_{\text{dom.}} \mu$. Hence, $A \succeq B$. Let $A' = A \cup \{r\}$, and $B' = B \cup \{r\}$. Take $r' \in \Delta(X)$. Suppose $\theta(A' \cup \{r'\}) > \theta(B' \cup \{r'\})$. Then, by definition, for $\widehat{\mu}' \in \Delta(B \cup \{r\})$, there exists $\widehat{\rho}' \in \Delta(A \cup \{r\})$ such that $\widehat{\rho}' \trianglerighteq_{\text{dom.}} \widehat{\mu}'$. In the same way, consider A' and B'. Then, we have $\widehat{\rho} \trianglerighteq_{\text{dom.}} \widehat{\mu}$. Hence, $A' \succeq B'$. In this way, it is shown that θ is *weakly additive*.

Lemma 6. *Let θ be a locally convex-valued measure on X. Then, θ is a weakly additive measure if and only if there exists a pair $\langle g, \rho \rangle$ where $g : [0, 1] \to [0, 1]$ is a strictly increasing function, and ρ is a locally convex-valued additive measure on X such that $\theta = g \circ \rho$. Moreover, $\langle g, \rho \rangle$ is unique.*

Proof. Step 1: If θ is a weakly additive measure, then $\trianglerighteq^*$ is a qualitative probability relation.

A binary relation $\trianglerighteq$ is a qualitative probability relation if and only if (i) $A \trianglerighteq B, B \trianglerighteq C \Rightarrow A \trianglerighteq C$, (ii) $A \trianglerighteq B$ or $B \trianglerighteq A$, (iii) $A \trianglerighteq \varnothing$, (iv) $\neg(\varnothing \triangleright X)$, and (v) $A \cap C = B \cap C = \varnothing \Rightarrow [A \trianglerighteq B \Leftrightarrow A \cup C \trianglerighteq B \cup C]$. We can show that $\trianglerighteq^*$ satisfies (i) - (v). (i) is easily shown. Since θ agrees with $\trianglerighteq^*$, $\trianglerighteq^*$ is *complete*. Then, (ii) is shown. (iii) and (iv) are obvious for any measure on X. (v) is easily shown by the *weak additivity* of θ.

Step 2: If θ is a locally convex-valued weakly additive measure, then $\trianglerighteq^*$ is *tight*.

Let us introduce the following. Denote a binary relation on $\mathcal{A}$ as follows: $B \trianglerighteq' C$ if and only if $B \cup E \trianglerighteq^* C$ for all $E \subset X$ with $E \triangleright^* \varnothing$ for which $B \cap E = \varnothing$. If $B \trianglerighteq' C$ and $C \trianglerighteq' B$, then denote $B \simeq' C$.

Definition 12. $\trianglerighteq^*$ is *tight* if and only if $B \simeq' C \Rightarrow B \simeq^* C$.

Suppose $B \triangleright^* C$. We show this step by the way of contradiction. Assume $C \trianglerighteq' B$. If $\theta(B) > \theta(C)$, then there exists $E \subset X$ with $E \cap C = \varnothing$ such that $\theta(B) > \theta(C \cup E) > \theta(C)$. Hence, we have $E \triangleright^* \varnothing$. This is a contradiction.

Step 3: If θ is a locally convex-valued weakly additive measure, then $\trianglerighteq^*$ is *fine*.

We introduce the following. We say that A and B are *nC-equivalent* for $n \in \mathbb{N}$ and $C \subset \Delta(X)$ if and only if there exists $\{A_i\}_{i=1}^n$ and $\{B_i\}_{i=1}^n$ such that (i) $C \trianglerighteq^* A_i, B_i$, (ii) $B \trianglerighteq^* A \setminus \cup_i A_i$, and (iii) $A \trianglerighteq^* B \setminus \cup_i B_i$.

Definition 13. $\trianglerighteq^*$ is *fine* if and only if for any $E \subset \Delta(X)$ with $E \triangleright^* \varnothing$, there exists $n \in \mathbb{N}$ suhc that $\varnothing$ and X are *nE-equivalent*.

Take $E \subset \Delta(X)$ with $E \triangleright^* \varnothing$. Define $\approx$ on $\mathcal{A}$ by $A \approx B$ if and only if there exists $n \in \mathbb{N}$ for which A and B are *nE-equivalent*. Note that $\approx$ is an *equivalence relation*. For any equivalence class $\mathcal{B} \subset \mathcal{A}$, denote $J_{\mathcal{B}} = \{\theta(B)\}_{B \in \mathcal{B}}$. Note that

(i) $\mathcal{B}_1 \neq \mathcal{B}_2$, $J_{\mathcal{B}_1} \cap J_{\mathcal{B}_2} = \varnothing$ ($\because \theta(B_1) = \theta(B_2) \Rightarrow B_1 \approx B_2$).

(ii) For all $\mathcal{B}$, $J_{\mathcal{B}}$ is *convex*.

(iii) For all $\mathcal{B}$, $J_{\mathcal{B}}$ is (relatively) open in $[0,1]$.

Hence, $\{J_{\mathcal{B}}\}_{\mathcal{B} \subset \mathcal{A} \setminus \approx}$ is a partition of $[0,1]$ into disjoint open intervals. Then, it must be $\varnothing \approx X$. Therefore, $\trianglerighteq^*$ is *fine*.

The necessity part is easily shown. The uniqueness result implies that ρ is bound to be unique under the three steps. Then, g is also unique. □

It is shown that, for any $A \in \mathcal{A}$, $\theta(A) = g \circ \rho(A)$ for some $g : [0,1] \to [0,1]$.

STEP 4

First, we show that $\succeq$ on $\mathcal{A}^*$ is represented by $V^* : \mathcal{A}^* \to \mathbb{R}$ defined by a pair $\langle u, g \rangle$. Next, we show that the representation of $\succeq$ is extended into the whole domain $\mathcal{A}$.

In Kreps (1979), it is shown that $\succeq$ satisfies *Standard Preferences* and *Strategic Rationality* if and only if there exists a non-constant and continuous function $u : X \to \mathbb{R}$ such that $\succeq$ is represented by $V : \mathcal{A} \to \mathbb{R}$ defined by

$$V(A) = \max_{x \in A} u(x).$$

This is equivalent to the following: $V(A) = \max_{x \in A} u(x) = \max_{\rho \in \Delta(A)} \sum_{x \in A} u(x)\rho(x)$ where $\Delta(A)$ is the set of all probability distributions on A. This representation says that the *best* alternative in a menu A is chosen with probability 1. We need to relax the axiom of *Strategic Rationality* to obtain the representation in Theorem 1.

We rule out the axiom of *Strategic Rationality*. Instead, we impose on the axiom of *Randomization*. First, consider binary menus. Then, we consider the whole domain $\mathcal{A}$.

Lemma 7. *$\succeq$ satisfies Standard Preferences, Randomization, and Singleton Independence if and only if there exists a non-constant utility function $u : \Delta(X) \to \mathbb{R}$ and a continuous and strictly increasing function $g : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$ such that $\succsim$ is represented by $V^* : \mathcal{A}^* \to \mathbb{R}$, for any $p, q \in \Delta(X)$ such that*

$$V^*(\{p,q\}) = \max_{\rho \in \Delta(\{p,q\})} u(p)g(\rho(p)) + u(q)g(\rho(q)),$$

where $\Delta(\{p,q\})$ is the set of all probability distributions on $\{p,q\}$.

Proof. We show the sufficiency part. By Lemma 5, there exists a pair $\langle u, \theta \rangle$ where $u : \Delta(X) \to \mathbb{R}$ and θ is a capacity on $\Delta(X)$. By Step 3, θ is *decomposable*. Then, there exists $g : [0,1] \to [0,1]$ such that $\theta = g \circ \rho$ where ρ is an additive measure on $\Delta(X)$.

For any $p \in \Delta(X)$, define $V(\{p\}) = u(p)$. We show that, for any $p, q \in \Delta(X)$,

$$\{p\} \succeq \{q\} \Leftrightarrow g(\rho(p)) \geq g(\rho(q)),$$

for some continuous and strictly increasing function $g : [0,1] \to [0,1]$. By the axiom of *Standard Preferences*, $\succeq$ is a weak order. By the axiom of *Randomization*, for any $p, q \in \Delta(X)$, if $\delta_{c_p} \geq \delta_{c_q}$, then $\{p\} \succeq \{q\}$ where δ_{c_p} is a lottery that gives the certainty equivalent of p with certainty, and δ_{c_q} is a lottery that gives the certainty equivalent of q with certainty. By the axiom of *Singleton Independence*, for any $p, q, r \in \Delta(X)$ and $\lambda \in [0,1]$, $\{p\} \succeq \{q\} \Leftrightarrow \{\lambda p + (1-\lambda)r\} \succeq \{\lambda q + (1-\lambda)r\}$.

Take $p, q \in \Delta(X)$ with $\{p\} \succeq \{q\}$. Then, by definition, we have $V^*(\{p\}) \geq V^*(\{q\}) \Leftrightarrow u(p) \geq u(q)$. First, suppose $\{p,q\} \sim \{p\}$. Then, $\rho(p) = 1$, so $g(\rho(p)) = 1$ and $g(\rho(q)) = 0$.

Second, suppose $\{p,q\} \succ \{p\}$. Without loss of generality, suppose $\{p\} \succeq \{q\}$. Then, $u(p) \geq u(q)$. Since g is strictly increasing, if $u(p) \geq u(q)$, then it must be $\rho(p) \geq \rho(q)$. Thus, we have $u(p) \geq u(q) \Leftrightarrow g(\rho(p)) \geq g(\rho(q))$. Since both u and g is continuous, V^* is continuous. Hence, we obtain, for any $p,q \in \Delta(X)$, $V^*(\{p,q\}) = \max_{\rho \in \Delta(\{p,q\})} u(p)g(\rho(p)) + u(q)g(\rho(q))$.

The necessity part is trivial. Take $p,q \in \Delta(X)$ with $g(\rho(p)) \geq g(\rho(q))$. Since g is strictly increasing, $\rho(p) \geq \rho(q) \Leftrightarrow u(p) \geq u(q)$. Hence, V^* represents $\succeq$. $\square$

Finally, we show that the representation holds in the whole domain. We provide a weaker version of *Strong Singleton Independence*:

Axiom (Singleton Independence): For any $\lambda \in [0,1]$ and $p \in \Delta(X)$,

$$A \succeq B \Rightarrow \lambda A + (1-\lambda)\{p\} \succsim \lambda B + (1-\lambda)\{p\}.$$

It is easily shown that $\succeq$ satisfies *Singleton Independence* if $\succeq$ satisfies *Randomization* and *Strong Singleton Independence*.

Take $A, B \in \mathcal{A}$ with $A \succeq B$. Fix such two menus A, B. It is applicable for the whole domain by induction. First, consider the case $|A| = 1$ with $A = \{p\}$. By definition, $V(\{p\}) = u(p)$. Next, consider the case $|A| = 2$. This is the result of Lemma 7.

Suppose that the case $|A| = n$ holds. The utility of a menu A is represented by $V : \mathcal{A} \to \mathbb{R}$. Consider a menu $A \cup \{p_{n+1}\}$. Let $\overline{p} \in A$ be the *best* alternative in the menu A. Let $\underline{p} \in A$ be the *worst* alternative in the menu A. Let $A' \equiv A \cup \{p_{n+1}\}$. We show that the utility of the menu A' is represented by $V : \mathcal{A} \to \mathbb{R}$.

First, consider the case $\{p_{n+1}\} \succeq \{p\}$ for any $p \in A$. Then, $u(p_{n+1}) \geq u(p)$ for any $x \in A$. Since g is strictly increasing, $\rho(p_{n+1}|A) \geq \rho(p|A)$ for any $x \in A$. We obtain $\rho(p_{n+1}|A) \in (0,1]$. Thus, $V(A') \geq V(B)$, which represents $\succeq$. Next, consider the case $\{\overline{p}\} \succeq \{p_{n+1}\} \succeq \{\underline{p}\}$. In this case, we have $\rho(p_{n+1}|A) \in [0,1)$. Then, $V(A') \geq V(B)$, which represents $\succeq$. Finally, consider the case $\{p\} \succeq \{p_{n+1}\}$ for any $x \in A$. In this case, $\rho_A(p_{n+1}) \in [0,1)$. Then, $g(\rho(p_{n+1}|A)) \in [0,1)$. $V(A') \geq V(B)$, which represents $\succeq$.

By the axiom of *Weak Singleton Independence*, we have $V(\lambda A + (1-\lambda)\{p_{n+1}\}) = \lambda V(A) + (1-\lambda)V(\{p_{n+1}\})$ for any $\lambda \in [0,1]$. Fix $\lambda \in (0,1)$. Let $A_\lambda = \lambda A + (1-\lambda)\{p_{n+1}\}$. Notice that $|A_\lambda| = n$. Then, $V(A_\lambda) = \lambda V(A) + (1-\lambda)u(p_{n+1})$. Define $V^* = \max_{\rho \in \Delta(A_\lambda)} \sum_{p \in A_\lambda} u(p)g(\rho(p))$.

We show that $V(A_\lambda) = V^*$, i.e., (i) $V(A_\lambda) \geq V^*$ and (ii) $V(A_\lambda) \leq V^*$. First, we show $V(A_\lambda) \geq V^*$. By the definition of V, $V(A_\lambda) \in \mathbb{R}$. We can find an alternative $r \in \Delta(X)$ such that $A \sim \{r\}$. Then, $V(A) = u(r)$. By the axiom of *Singleton Independence*, $V(A_\lambda) = \lambda u(r) + (1-\lambda)u(p_{n+1})$. On the other hand, we can find an alternative $r' \in X$ such that $u(r') = U^*$. By the axiom of *Randomization*, there exists $\alpha \in [0,1]$ such that $\{\alpha r + (1-\alpha)p_{n+1}\} \sim \{r'\}$. By the axiom of *Continuity*, we can find an alternative $r'' \in A_\lambda$

such that $\{r''\} \sim \{\alpha r + (1-\alpha)p_{n+1}\}$. Then, $V(A_\lambda) = u(r'')$. By the axiom of *Randomization*, we have $A_\lambda \succeq \{r''\}$. Thus, we can obtain $V(A_\lambda) \geq u(r'') = u(r') = U^*$. Next, we show $V(A_\lambda) \leq V^*$. Consider an alternative $r' \in \Delta(X)$ such that $u(r') = V^*$. We can find an alternative $r'' \in \mathrm{co}(A \cup \{r_{n+1}\})$ such that $\{r''\} \sim \{r'\}$. Since $\mathrm{co}(A \cup \{r_{n+1}\}) \succeq A_\lambda$ holds, $\{r''\} \succeq \{p\}$ for any $r \in A_\lambda$. Then, we obtain $V(A_\lambda) \leq V^*$. Hence, $V(A_\lambda) = V^*$.

We complete the representation. For any $A \in \mathcal{A}$, $\succsim$ satisfying *Standard Preferences*, *Randomization*, and *Singleton Independence* is represented by $V : \mathcal{A} \to \mathbb{R}$ defined by a non-constant function $u : \Delta(X) \to \mathbb{R}$ and a strictly increasing and continuous function $g : [0,1] \to [0,1]$:

$$V(A) = \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q)),$$

where $\Delta(A)$ is the set of all probability distributions on the menu A. $\square$

A.2 Necessity Part

We verify that V satisfies the axiom of *Randomization*. Other axioms are easily verified. The *random anticipated utility representation* with $\langle u, g \rangle$ represents $\succeq$ on $\mathcal{A}$. Take $A, B \in \mathcal{A}$ such that for any $\mu \in \Delta(B)$ there exists $\rho \in \Delta(A)$ such that ρ *dominates* μ. Without loss of generality, assume that $\rho_p = 1$ and $\mu_q = 1$ for some $p \in A$ and $q \in B$. Then, $\delta_{c_p} \geq \delta_{c_q}$; that is, for any $q \in B$ there exists $\delta_{c_p} \geq \delta_{c_q}$. Since g is strictly increasing, we obtain $V(A) \geq V(B)$. $\square$

B Proof of Propositions

B.1 Proof of Proposition 1

The uniqueness result is easily obtained by the property of bi-separability for $\succeq$. Suppose that both two random anticipated utility representation $\langle u, g \rangle$ and $\langle u', g' \rangle$ represent the same preference $\succeq$. Remember that, by the axiom of *Randomization*, a weaker version of *Strategic Rationality*, the support on the subjective state space is one.

By the bi-separability of $\succeq$, the utility function u is cardinal. That is, u is unique to a positive affine transformation. There exists $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$.

Moreover, by the bi-separability of $\succeq$, the capacity θ is also unique. By Lemma 6, for all $A \in \mathcal{A}$, $\theta(A) = \theta'(A)$. By Theorem 1, for any $p, q \in \Delta(X)$, there exists $g : [0,1] \to [0,1]$ such that $u(p) \geq u(q) \Leftrightarrow g(\rho(p)) \geq g(\rho(q))$. Hence, for each $A \in \mathcal{A}$, $\theta(A)$ is decomposable, i.e., $\theta(A) = g \circ \rho(A)$. We obtain $g = g'$. $\square$

B.2 Proof of Proposition 2

We prove (i). In the similar way, we can prove (ii), so we omit the proof of (ii).

First, we show the sufficiency part. Take $p, q \in \Delta(X)$ such that $\{p\} \sim \{q\}$. Suppose that $\succeq$ satisfies *Desire for Randomization*. Fix $\lambda \in [0,1]$. By Theorem 1, we have $V(\lambda \circ \{p\} \oplus (1-\lambda) \circ \{q\}) \geq V(\{p\}) \Leftrightarrow V(\lambda \circ \{p\} \oplus (1-\lambda) \circ \{q\}) \geq \min\{V(\{p\}), V(\{q\})\}$. Remember that u is mixture-linear by the axiom of *Strong Singleton Independence*. Hence, V is quasi-concave. Suppose $V(\lambda \circ \{p\} \oplus (1-\lambda) \circ \{q\}) > U(\{p\}) = u(p)$. This holds only if g is concave. Suppose that g is not concave. Without loss of generality, assume g is strictly convex. Then, we have $V(\{\lambda p + (1-\lambda) q\}) = V(\{p\}) = u(p)$. Thus, $V(\{\lambda p + (1-\lambda) q\}) \geq V(\{p\})$ only if g is concave.

Next, we show the necessity part. Suppose that g is concave. And, suppose $\{p\} \sim \{q\}$. Then, $V(\{p\}) = V(\{q\})$. Fix $\lambda \in [0,1]$. By the concavity of g, we can obtain $V(\lambda \circ \{p\} \oplus (1-\lambda) \circ \{q\}) \geq V(\{p\}) = u(p)$. V is quasi-concave. Therefore, by Theorem 1, we have $\lambda \circ \{p\} \oplus (1-\lambda) \circ \{q\} \succeq \{p\}$. $\square$

B.3 Proof of Proposition 3

First, we show the sufficiency part. Suppose that $\succeq$ satisfies the axiom of *Independence*. By the axiom of *Completeness*, *Transitivity*, and *Independence*, there exists a *mixture-linear* function $V : \mathcal{A} \to \mathbb{R}$. We show that g is linear, by the way of contradiction. Suppose not. Without loss of generality, suppose that g is concave. By Proposition 2, for any $p, q \in \Delta(X)$ with $\{p\} \sim \{q\}$, and $\lambda \in [0,1]$, $V(\lambda\{p\} + (1-\lambda)\{q\}) \geq \lambda V(\{p\}) + (1-\lambda) V(\{q\})$. This is a contradiction. Hence, g is linear.

Next, we show the necessity part. Suppose that g is linear. Then, for any $A \in \mathcal{A}$, $V(A) = \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q)\rho(q) = \max_{q \in A} u(q)$. Thus, we have, for any $A, B \in \mathcal{A}$ and $\lambda \in [0,1]$, $V(\lambda A + (1-\lambda) B) = \lambda V(A) + (1-\lambda) V(B)$. Hence, $\succeq$ satisfies the axiom of *Independence*. $\square$

B.4 Proof of Proposition 4

First, we show the sufficiency part. Suppose that $\succeq_1$ exhibts a stronger preference for randomization than $\succeq_2$. Take an arbitrary menu A with $|A| \geq 2$. Then, for any $p \in A$, $A \succeq_2 \{p\} \Rightarrow A \succeq_1 \{p\}$. Suppose that $\succeq_1$ is represented by V_1, and that $\succeq_2$ is represented by V_2. We obtain $V_2(A) \geq V_2(\{p\}) \Rightarrow V_1(A) \geq V_1(\{p\})$. Notice that $V_2(\{p\}) = V_1(\{p\}) = u(p)$. Thus, we have $V_1(A) \geq V_2(A)$. Since $u_1 = u_2 = u$, we obtain the desired statement: g_1 is more concave than g_2.

Next, we show the necessity part. Suppose that g_1 is more concave than g_2. Take an arbitrary menu $A \in \mathcal{A}$ with $|A| \geq 2$. For any $p \in \Delta(X)$, $V(\{p\}) = u(p)$. By the concavity of g, for any $p \in A$, $V(A) \geq V(\{p\}) = u(p)$. Since g_1 is more concave than g_2, we have $V_1(A) \geq V_2(A)$. Hence, the desired result is obtained. $\square$

B.5 Proof of Proposition 5

We show the sufficiency part. We modify the proof steps in the proof of Theorem 1. We introduce the axiom of *Monotonic Randomization*, instead of the axiom of *Randomization*. Suppose that $\succeq$ satisfies the axioms in Proposition 5.By applying Dekel et al. (2001)'s Theorem 4 by imposing on the axioms of *Monotonic Randomization* and *Weak Independence*.

By the axiom of *Monotonic Randomization*, take $A, A' \in \mathcal{A}$ with (i) for any $\mu \in \Delta(A')$, there exists $\rho \in \Delta(A)$ such that ρ *dominates* μ, i.e., $\rho \trianglerighteq_{\text{dom.}} \mu$, and (ii) $\text{co}(A') \subseteq \text{co}(A)$. Then, $A \succeq A'$. The condition (ii) is a monotonic condition for *deliberate randomization*.

By the axiom of *Weak Independence*, for any $A, B, C \in \mathcal{A}$ and $\lambda \in (0,1)$, there exist $A', B' \in \mathcal{A}$ such that $A' \sim \lambda A + (1-\lambda)C$ and $B' \sim \lambda B + (1-\lambda)C$. Then, $V(A') = V(\lambda A + (1-\lambda)C)$. And, we have $V(B') = V(\lambda B + (1-\lambda)C)$. Suppose $A \succ B$. Then, it is equivalent to $V(A) > V(B) \Leftrightarrow V(A') > V(B') \Leftrightarrow A' \succ B'$. That is, $A \succ B \Leftrightarrow A' \succ B' \Leftrightarrow \lambda A + (1-\lambda)C \succ \lambda B + (1-\lambda)C$. Hence, $V(A') = V(\lambda A + (1-\lambda)C) = \lambda V(A) + (1-\lambda)V(C)$. And, we have $V(B') = V(\lambda B + (1-\lambda)C) = \lambda V(B) + (1-\lambda)V(C)$. Thus, V is mixture-linear. Therefore, we have the representation: for any $A \in \mathcal{A}$,

$$V(A) = \int_S \max_{q \in A} v(q,s) d\mu(s),$$

where μ is a probability distribution on the subjective state space S and $v : \Delta(X) \times S \to \mathbb{R}$. In other words, the representation is the aggregation of state-dependent utility function v.

By the axioms of *Monotonic Randomization* and *Weak Independence*, V is mixture-linear. Hence, there exists a probability distribution μ on the subjective state space S. By the axiom of *Strong Singleton Independence*, u is state-independent. For any $s \in S$, let $v(x,s) = \sum_{q \in A} u(q) g(\rho(q), s)$ for some $g : [0,1] \times S \to [0,1]$ since $g(0) = 0$ and $g(1) = 1$ by the axioms of *Singleton Independence* and *Randomization*. Hence, for any $q \in \Delta(X)$, we have $V(\{q\}) = \int_S u(q) d\mu(s) = u(q)$. Thus, we obtain the desired representation: for any $A \in \mathcal{A}$,

$$V(A) = \int_S \max_{\rho \in \Delta(A)} \sum_{x \in A} u(x) g(\rho(x), s),$$

for some $g : [0,1] \times S \to [0,1]$, i.e., a state-dependent probability-weighting function. $\square$

B.6 Proof of Proposition 6

We show that the framework in this paper is equivalent to that in Dekel et al. (2001). Then, it is applicapable for subjective partitional learning in Dillenberger et al. (2014).

By the axioms in Theorem 1, u is unique up to positive affine transformations. By the axiom of *Dominance*, for each $\omega \in \Omega$, g_w is uniquely identified. Let $N = |n|$ ($\Omega = \{\omega_1, \cdots, \omega_n\}$). By the uniqueness result, we normalize the utility space by $\mathcal{U} := [0,1]^N$.

We embed the utility space $\mathcal{U}$ into a $N+1$ dimensional space. That is,

$$[0,1]^N \ni (v_1, \cdots, v_N) \mapsto \Big(v_1, \cdots, v_N, N - \sum_n v_n\Big) \in [0,1]^N \times [0, N].$$

Thus, we obtain

$$\mathcal{U}' := \Big\{ v' \in [0,1]^N \times [0,N] \mid \sum_{i=1}^{N+1} v'_i = N \Big\}.$$

Moreover, we obtain

$$\mathcal{U}'' := \Big\{ v' \in \mathbb{R}^{N+1} \mid \sum_{i=1}^{N+1} v'_i = N \Big\},$$

i.e., $\mathcal{U}'$ is a subset of $\mathcal{U}''$ ($\mathcal{U}' \subseteq \mathcal{U}''$).

$\mathcal{U}$ and $\mathcal{U}'$ are isomorphic. Let $\mathcal{A}'$ be the set of all non-empty compact subsets of $\mathcal{U}'$. Consider the binary relation $\succeq'$ on $\mathcal{A}'$. For $v' \in \mathcal{A}'$, denote the vector that agrees with the first n comoponents of v' by v'^n. Define $A' \succeq' B'$ if and only if $A^* \succeq B^*$ where $A^* = \{v \in [0,1]^N \mid v = v'^n \text{ for some } v' \in A' \}$ and $B^* = \{v \in [0,1]^N \mid v = v'^n \text{ for some } v' \in B' \}$.

Remark. $\succeq'$ satisfies *Independence*.

Proof. For all $A', B', C' \in \mathcal{A}'$ and $\lambda \in [0,1]$, $A' \succeq' B' \Leftrightarrow A \succeq B$. There exists $C \in \mathcal{A}$ such that $\lambda\rho + (1-\lambda)\gamma \trianglerighteq_{\text{dom.}} \lambda\mu + (1-\lambda)\gamma$ where $\rho \in \Delta(A), \mu \in \Delta(B)$, and $\gamma \in \Delta(C)$. By definition, we have $\lambda A + (1-\lambda)C \succeq \lambda B + (1-\lambda)C$. Then, by the definiton of $\succeq'$, $(\lambda A + (1-\lambda)C)' \succeq' (\lambda B + (1-\lambda)C)'$. Hence, we obtain $\lambda A' + (1-\lambda)C' \succeq' \lambda B' + (1-\lambda)C'$. It is shown that $\succeq'$ satisfies *Independence*. $\square$

Next, let $\mathcal{A}''$ be the set of all non-empty compact subsets of $\mathcal{U}''$. consider the binary relation $\succeq''$ on $\mathcal{A}''$ in the following way: $A \succeq'' B$ if and only if for any $\varepsilon < \frac{1}{N^2}$, $\varepsilon A'' + (1-\varepsilon)A^{N+1} \succeq' \varepsilon B'' + (1-\varepsilon)A^{N+1}$ where $A^{N+1} := \{(\frac{N}{N+1}, \cdots, \frac{N}{N+1})\} \in \mathcal{A}'$. Observe that $\varepsilon A'' + (1-\varepsilon)A^{N+1} \in \mathcal{A}'$.

$\succeq''$ is the unique extension of $\succeq'$ to $\mathcal{A}''$ (see Claim 4 in Dillenberger et al. (2014)). $\succeq''$ satisifies the axioms of *Completeness*, *Transitivity*, *Continuity*, *Independence*, and *Monotonicity*. We show that $\succeq''$ satisfies *Independence*. By definition, for any $A'', B'' \in \mathcal{A}''$, $A'' \succeq'' B'' \Leftrightarrow \varepsilon < \frac{1}{N^2}$, $\varepsilon A'' + (1-\varepsilon)A^{N+1} \succeq' \varepsilon B'' + (1-\varepsilon)A^{N+1} \Leftrightarrow \lambda(\varepsilon A'' + (1-\varepsilon)A^{N+1}) + (1-\lambda)(\varepsilon C'' + (1-\varepsilon)A^{N+1}) \succeq'' \lambda(\varepsilon B'' + (1-\varepsilon)A^{N+1}) + (1-\lambda)(\varepsilon C'' + (1-\varepsilon)A^{N+1})$. Since A^{N+1} is a singleton menu, we have $\lambda(\varepsilon A'' + (1-\varepsilon)A^{N+1}) + (1-\lambda)(\varepsilon C'' + (1-\varepsilon)A^{N+1}) = \varepsilon(\lambda A'' + (1-\lambda)C'') + (1-\varepsilon)A^{N+1}$ and $\lambda(\varepsilon B'' + (1-\varepsilon)A^{N+1}) + (1-\lambda)(\varepsilon C'' + (1-\varepsilon)A^{N+1}) = \varepsilon(\lambda B'' + (1-\lambda)C'') + (1-\varepsilon)A^{N+1}$. By the axiom of *Weak Independence* of $\succeq$, we can find $\widehat{A}'', \widehat{B}'' \in \mathcal{A}''$ such that $\widehat{A}'' \sim \lambda A'' + (1-\lambda)C''$ and $\widehat{B}'' \sim \lambda B'' + (1-\lambda)C''$. Then, $A'' \succ B'' \Leftrightarrow \widehat{A}'' \succ'' \widehat{B}''$. Hence, we obtain $\lambda A'' + (1-\lambda)C'' \succeq'' \lambda B'' + (1-\lambda)C''$. The axiom of

Monotonicity is satisfied in the sense that $\succeq''$ satisfies the axiom of *Monotonic Randomization* (Proposition 4.1).

The domain $\mathcal{A}''$ with the Hausdorff metric is equivalent to that of Dekel et al. (2001) with $N+1$ outcomes/alternatives. Since $\succeq''$ satisfies the axioms in Dillenberger et al. (2014), it is applicapable for the representation of subjective partitional learning. $\square$

B.7 Proof of Proposition 7

We show the sufficiency part, and follow the proof steps in Ortoleva (2013). The domain of this paper is different from that of Ortoleva (2013). We show that a representation of thinking aversion as costly randomization is represented in our domain. Define $\succeq^*$ on $\mathcal{A}$ in the following way. We say that $A \succeq^* B$ if for any $\mu_B \in \Delta(B)$ there exists $\rho_A \in \Delta(A)$ such that $\rho_A \trianglerighteq \mu_B$.

First, we show that $\succeq^*$ is *well-defined.*

Lemma 8. *Suppose that $\succeq$ satisfies Thinking Aversion, i.e., for any $A, B \in \mathcal{A}$ there exists $\lambda, \lambda' \in (0,1)$ with $\lambda > \lambda'$, $\lambda A + (1-\lambda)B \succeq \lambda' A + (1-\lambda')B$. Then, for any $\mu, \mu' \in (0,1)$ with $\mu > \mu'$, $\mu A + (1-\mu)B \succeq \mu' A + (1-\mu')B$.*

Proof. Suppose that for any $A, B \in \mathcal{A}$, there exists $\lambda, \lambda' \in (0,1)$ such that $\lambda A + (1-\lambda)B \succeq \lambda' A + (1-\lambda')B$. Suppose that there exist $\overline{A}, \widehat{A} \in \mathcal{A}$ such that $\overline{A} \sim \lambda A + (1-\lambda)B$ and $\widehat{A} \sim \lambda' A + (1-\lambda')B$. Then, we obtain $\overline{A} \succeq \widehat{A}$. Take $\lambda_k, \lambda'_k \in (0,1)$ with $\lambda_k > \lambda'_k$ and $\lambda_k \to \lambda$, $\lambda'_k \to \lambda'$ as $k \to \infty$. Suppose, by the way of contradiction, $\lambda_k A + (1-\lambda_k)B \prec \lambda'_k A + (1-\lambda'_k)B$. In the same way, we can find $\overline{A}_k, \widehat{A}_k \in \mathcal{A}$ such that $\overline{A}_k \sim \lambda_k A + (1-\lambda_k)B$ and $\widehat{A}_k \sim \lambda'_k A + (1-\lambda'_k)B$. Hence, $\overline{A}_k \prec \widehat{A}_k$. However, as $k \to \infty$, $\overline{A} \prec \widehat{A}$. This is a contradiction. Thus, for any $\mu, \mu' \in (0,1)$ with $\mu > \mu'$, $\mu A + (1-\mu)B \succeq \mu' A + (1-\mu')B$. $\square$

Next, we show the following.

Lemma 9. *$\succeq^*$ is a weak order.*

Proof. First, we show that $\succeq^*$ is *complete*. This follows from Lemma 8. Second, we show that $\succeq^*$ is *transitive*. Take $A, B, C \in \mathcal{A}$ with $A \succeq^* B \succeq^* C$. We show $A \succeq^* C$. Take $\lambda, \lambda' \in (0,1)$ with $\lambda > \lambda'$. Since $A \succeq^* B$, $\lambda A + (1-\lambda)B \succeq \lambda' A + (1-\lambda')B$. In the same way, since $B \succeq^* C$, $\lambda B + (1-\lambda)C \succeq \lambda' B + (1-\lambda')C$. By the axiom of *Singleton Independence*, there exist $\widehat{A}, \widehat{B} \in \mathcal{A}$ such that $\widehat{A} \sim \lambda A + (1-\lambda)B$ and $\widehat{B} \sim \lambda B + (1-\lambda)C$. Without loss of generality, assme $B \sim^* C$. Then, by the axiom of *Singleton Independence*, $A \succ B \Leftrightarrow A \succ^* B \Leftrightarrow \widehat{A} \succ^* \widehat{B}$. Take $\mu, \mu' \in (0,1)$ with $\mu > \mu'$. By $\widehat{A} \succ^* \widehat{B}$, we have the following: $\mu\widehat{A} + (1-\mu)\widehat{B} \succsim \mu'\widehat{A} + (1-\mu')\widehat{B}$. Since $\widehat{A} \sim \lambda A + (1-\lambda)B$ and $\widehat{B} \sim \lambda B + (1-\lambda)C$, $\mu[\lambda A + (1-\lambda)B] + (1-\mu)[\lambda B + (1-\lambda)C] \succeq \mu'[\lambda A + (1-\lambda)B] + (1-\mu')[\lambda B + (1-\lambda)C]$. Thus, we have $\mu\lambda > (1-\mu)(1-\lambda)$, $A \succeq^* C$. $\square$

By the convexity and compactness of X, define

$$\mathcal{H} := \{A \in \mathcal{A} \mid \{\overline{x}\} \succeq A \succeq \{\underline{x}\}\}$$

for some $\overline{x}, \underline{x} \in X$ with $\{\delta_{\overline{x}}\} \succ \{\delta_{\underline{x}}\}$.[24]

Lemma 10. *The following statements hold.*

(i) For any $A \in \mathcal{H}$, there exists $p \in \Delta(X)$ such that $\{p\} \sim^* A$.

(ii) For any $A \in \mathcal{H}$, there exists $p \in \Delta(X)$ such that $\{p\} \sim A$.

The proof is in Ortoleva (2013). We omit it.

$\succeq^*$ satisfies *Completeness, Transitivity, Weak Randomization*, and *Strong Singleton Independence*. Then, there exists $V^* : \mathcal{A} \to \mathbb{R}$ such that V^* represents $\succeq^*$; that is, for any $A \in \mathcal{A}$,

$$V^*(A) = \max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g_A(\rho(q)),$$

for some $u : \Delta(X) \to \mathbb{R}$ and $g_A : [0,1] \to [0,1]$ with $g(0) = 0$ and $g(1) = 1$ (for each $A \in \mathcal{A}$).[25]

Define $V' : \mathcal{H} \to \mathbb{R}$ by for all $A \in \mathcal{A}$ with $\{r\} \sim A$,

$$V'(A) = V(\{r\}) = u(\delta_{c_r}),$$

for some $\delta_{c_r} \in \Delta(X)$. In singleton menus, for any $p, q \in \Delta(X)$, we have $\{p\} \succeq \{q\} \Leftrightarrow \{p\} \succeq^* \{q\}$. Take $p, q \in \Delta(X)$ with $\{p\} \succeq \{q\}$. Then, for any $\lambda, \lambda' \in (0,1)$ with $\lambda > \lambda'$, $\lambda\{p\} + (1-\lambda)\{q\} \succeq \lambda'\{p\} + (1-\lambda')\{q\}$. Then, we obtain $\{p\} \succeq^* \{q\}$. By definition, $\{p\} \succeq \{q\} \Leftrightarrow \{p\} \succeq^* \{q\}$.

We now consider costs of thinking. Define $c : \mathcal{H} \to \mathbb{R}$ by,

$$\widehat{c}(A) := V^*(A) - V'(A).$$

In the following, we show that $\widehat{c}$ has certain properties.

Lemma 11. *For any $A \in \mathcal{H}$, $p \in \Delta(X)$, and $\lambda \in (0,1)$, $\widehat{c}(A) = \widehat{c}(\lambda A + (1-\lambda)\{x\})$.*

Proof. Take $A \in \mathcal{H}$, $p \in \Delta(X)$, and $\lambda \in (0,1)$. We can find out $q \in \Delta(X)$ such that $\{q\} \sim A$. $\succeq$ is represented by $V'(\cdot) = V^*(\cdot) - \widehat{c}(\cdot)$. Since V^* is mixture-linear with respect to singleton menus, this implies that $\widehat{c}(A) = \widehat{c}(\lambda A + (1-\lambda)\{p\})$. □

Lemma 12. *For any $A, B \in \mathcal{A}$, if $B \subseteq A$, then $\widehat{c}(A) \geq \widehat{c}(B)$.*

[24] Note that $\mathcal{H} \subset \mathcal{A}$.

[25] In the following, we want to show that for each $A \in \mathcal{A}$,$V(A) = V^*(A) - c(\text{supp}(\rho))$.

Proof. Take an arbitrary menu $A \in \mathcal{H}$ with $|A| \geq 2$. Then, by Lemma 10, there exists $q \in \Delta(X)$ such that $\{q\} \sim^* A$. By definition, $V^*(\{q\}) = V'(\{q\})$. Hence, we have $\widehat{c}(\{q\}) = 0$. By Lemma 11, we obtain $\widehat{c}(A) \geq \widehat{c}(\{q\})$. Since for any $p \in \Delta(X)$, $\widehat{c}(\{p\}) = 0$, by taking $p \in A$, we have $\widehat{c}(A) \geq \widehat{c}(\{p\})$. By the axiom of *Thinking Aversion* and Lemma 11, we obtain the desired result. if $B \subseteq A$, then $\widehat{c}(A) \geq \widehat{c}(B)$. $\square$

Define $c : \mathcal{A} \to \mathbb{R}$ by, for any $\rho \in \Delta(A)$,

$$c(\text{supp}(\rho)) := \widehat{c}(A')$$

for some $A' \in \mathcal{H}$ with $A \sim A'$. We can show that c is well-defined by Lemma 11. For any $p \in \Delta(X)$, we have $V^*(\{p\}) = V'(\{p\})$. Hence, $\widehat{c}(\{p\}) = 0$. In addition, $c(\varnothing) = 0$. We have $c(\text{supp}(\rho)) \geq 0 = c(\varnothing)$. Hence, $V'(A) = V^*(A) - c(\text{supp}(\rho))$, for all $A \in \mathcal{A}$.

Finally, we extend the whole domain $\mathcal{A}$. Define $V : \mathcal{A} \to \mathbb{R}$ by $V(A) = V^*(A) - c(\text{supp}(\rho))$, for all $A \in \mathcal{A}$. Notice that, $V(A) = V'(A)$ if $A \in \mathcal{H}$. By the axioms of *Strong Singleton Independence* and *Weak Randomization*, for any $A \in \mathcal{A}$, there exists $A' \in \mathcal{H}$ such that $A \sim A'$. Then, V is well-defined, and V represents $\succsim$. $\square$

B.8 Proof of Proposition 8

Suppose that both two *random anticipated utility representation with thinking aversion* $\langle u, (g_A)_{A \in \mathcal{A}}, c \rangle$ and $\langle u', (g'_A)_{A \in \mathcal{A}}, c' \rangle$ represent the same preference $\succsim$.

First, we show that u is unique. By Proposition 1, By the bi-separability of $\succsim$, the utility function u is cardinal. That is, u is unique to a positive affine transformation. There exists $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$. u is unique up to positive affine transformation.

Second, we show that g is unique. By Lemma 6, the capacity θ is also unique. That is, $\theta = \theta'$. Since $\theta = g \circ \rho$, we obtain $g = g'$. We can easily apply the case of menu-dependence into this result.

Third, we show that c is unique. By the uniqueness of u, we obtain $V = aV' + b$. We can find $A \in \mathcal{A}$ such that $A \sim \{p\}$ for some $p \in \Delta(X)$. Then, $V(\{p\}) = V(A) - c(\text{supp}(\rho))$ and $V'(\{p\}) = V'(A) - c'(\text{supp}(\rho))$. Since we have $V = aV' + b$, we obtain $c = ac'$. $\square$

B.9 Proof of Proposition 9

Let $\rho^* \in \arg\max_{\rho \in \Delta(A)} \sum_{q \in A} u(q) g(\rho(q))$. We show that ρ^* satisfies the axiom of *IIA*. Take $p, q, r, r' \in \Delta(X)$. Suppose $\{p\} \succsim \{q\}$. Then, we have $u(p) \geq u(q)$. If $\{p\} \sim \{p, q\}$, then $V(\{p, q\}) = V(\{p\}) = u(p)$. That is, $\rho^*(p|\{p, q\}) = 1$. If $\{p, q\} \succ \{p\}$, then $\rho^*(p) > 0$ and $\rho^*(q) > 0$. Since g is strictly increasing and $u(p) \geq u(q)$, $g(\rho^*(p)) > g(\rho^*(q))$. This implies that $\rho^*(p|\{p, q\}) \geq \rho^*(q|\{p, q\}) \Leftrightarrow u(p) \geq u(q)$. Hence, $\rho^*(p|\{p, r\}) \geq \rho^*(q|\{q, r\})$ holds if $u(p) \geq u(q)$. This holds for any $r \in \Delta(X)$ because g is strictly increasing. $\rho^*(p|\{p, r\}) \geq \rho^*(q|\{q, r\})$ implies $\rho^*(p|\{p, r'\}) \geq \rho^*(q|\{q, r'\})$. $\square$

References

Agranov, M., and Ortoleva, P. (2017): "Stochastic Choice and Preferences for Randomization," *Journal of Political Economy*, 125(1), 40-68.

Ahn, D. S., and Sarver, T (2013): "Preference for Flexibility and Random Choice," *Econometrica*, 81(1), 341-361.

Aliprantis, C. D., and Border, K. (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer Science and Business Media.

Anscombe, F. J., and Aumann, R. J. (1963): "A Definition of Subjective Probability," *Annals of Mathematical Statistics*, 34(1), 199-205.

Cerreia-Vioglio, S., Dillenberger, D., and Ortoleva, P. (2015): "Cautious Expected Utility and the Certainty Effect," *Econometrica*: 83(2), 693-728.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., and Riella, G. (2017): "Deliberately Stochastic," Working Paper.

Dean, M., and McNeill, J. (2015): "Preference for Flexibility and Random Choice: An Experimental Analysis," Working Paper.

Dekel, E., Lipman, B. L., and Rustichini, A. (2001): "Representing Preferences with a Unique Subjective State Space," *Econometrica*: 69(4), 891-934.

Dillenberger, D., Lleras, J. S., Sadowski, P., and Takeoka, N. (2014): "A Theory of Subjective Learning," *Journal of Economic Theory*, 153, 287-312.

Dwenger, N., Kübler, D., and Weizsäcker, G. (2014): "Flipping a Coin: Theory and Evidence," Working Paper.

Eliaz, K., and Ok, E. A. (2006): "Indifference or Indecisiveness? Choice-Theoretic Foundations of Incomplete Preferences," *Games and Economic Behavior*: 56(1), 61-86.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage axioms," *Quarterly Journal of Economics*: 643-669.

Ergin, H., and Sarver, T. (2010): "A Unique Costly Contemplation Representation," *Econometrica*: 78(4), 1285-1339.

Fudenberg, D., Iijima, R., and Strzalecki, T. (2014): "Stochastic Choice and Revealed Perturbed Utility," Working Paper.

Fudenberg, D., Iijima, R., and Strzalecki, T. (2015): "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83(6), 2371-2409.

Ghirardato, P., Maccheroni, F., Marinacci, M., and Siniscalchi, M. (2003): "A Subjective Spin on Roulette Wheels," *Econometrica*: 71(6), 1897-1908.

Ghirardato, P., and Marinacci, M. (2001): "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Mathematics of operations research*: 26(4), 864-890.

Gilboa, I. (1985): "Subjective Distortions of Probabilities and Non-Additive Probabilities," Working Paper, Tel Aviv University.

Gilboa, I., and Schmeidler, D. (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*: 18(2), 141-153.

Gul, F., Natenzon, P., and Pesendorfer, W. (2014): "Random Choice as Behavioral Optimization," *Econometrica*: 82(5), 1873-1912.

Gul, F., and Pesendorfer, W. (2001): "Temptation and Self-Control," *Econometrica*: 69(6), 1403-1435.

Gul, F., and Pesendorfer, W. (2006): "Random Expected Utility," *Econometrica*: 74(1), 121-146.

Hashidate, Y. (2017): "Preferences for Randomization and Anticipation," *Journal of Behavioral Economics and Finance*, 9, 99-101.

Huber, J., Payne, J. W., and Puto, C. (1982): "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*: 9(1), 90-98.

Kreps, D. M. (1988): *Notes on the Theory of Choice*, Westview Press.

Kreps, D. M. (1979): "A Representation Theorem for "Preference for Flexibility"," *Econometrica: Journal of the Econometric Society*: 565-577.

Kreps, D. M., and Porteus, E. L. (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica: Journal of the Econometric Society*: 185-200.

Luce, R. D. (1959): *Individual Choice Behavior: A Theoretical Analysis*, Wiley: New York.

Machina, M. J. (1984): "Temporal Risk and the Nature of Induced Preferences," *Journal of Economic Theory*: 33(2), 199-231.

Machina, M. J. (1985): "Stochastic Choice Functions generated from Deterministic Preferences over Lotteries," *Economic Journal*: 95(379), 575-594.

Noor, J., and Takeoka, N. (2015): "Menu-Dependent Self-Control," *Journal of Mathematical Economics*: 61, 1-20.

Ortoleva, P. (2013): "The Price of Flexibility: Towards a Theory of Thinking Aversion," *Journal of Economic Theory*: 148(3), 903-934.

Piermont, E., Takeoka, N., and Teper, R. (2016): "Learning the Krepsian State: Exploration through Consumption," *Games and Economic Behavior*: 100, 69-94.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*: 3(4), 323-343.

Raiffa, H. (1961): "Risk, Ambiguity, and the Savage Axioms: Comment," *Quarterly Journal of Economics*: 690-694.

Saito, K. (2015): "Preferences for Flexibility and Randomization under Uncertainty," *American Economic Review*: 105(3), 1246-1271.

Savage, L. (1954): *The Foundation of Statistics*, Wiley, New York (2nd Edition: 1972, Dover Publications).

Sen, A. (1993): "Internal Consistency of Choice," *Econometrica: Journal of the Econometric Society*:495-521.

Simonson, I. (1989): "Choice Based on Reasons: The Case of Attraction and Compromise effects," *Journal of Consumer Research*: 158-174.

Tversky, A., and Russo, J. E. (1969): "Substitutability and Similarity in Binary Choices," *Journal of Mathematical Psychology*:6(1), 1-12.

Wakker, P. P. (2010): *Prospect Theory: For Risk and Ambiguity*, Cambridge University press.