WINPEC Working Paper Series No.E1902
April 2019

Expected Utility Theory with Probability Grids and Preference Formation

Mamoru Kaneko

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan

Expected Utility Theory with Probability Grids and Preference Formation*

Mamoru Kaneko†

April 7, 2019

Abstract

We reformulate expected utility theory, from the viewpoint of bounded rationality, by introducing probability grids and a cognitive bound; we restrict permissible probabilities only to decimal (ℓ-ary in general) fractions of finite depths up to a given cognitive bound. We distinguish between measurements of utilities from pure alternatives and their extensions to lotteries involving more risks. Our theory is constructive, from the viewpoint of the decision maker. When a cognitive bound is small, the preference relation involves many incomparabilities, but these diminish as the cognitive bound is is relaxed. Similarly, the EU hypothesis would hold more for a weaker cognitive bound. The main part of the paper is a study of preferences including incomparabilities in cases with finite cogntive bounds; we give representation theorems in terms of a 2-dimensional vector-valued utility functions. We exemplify the theory with one experimental result reported by Kahneman-Tversky.

JEL Classification Numbers: C72, C79, C91

Key Words: Expected Utility, Measurement of Utility, Bounded Rationality, Probability Grids, Cognitive Bound, Incomparabilities

1 Introduction

We reconsider EU theory from the viewpoint of preference formation and of bounded rationality. We restrict permissible probabilities to decimal (ℓ-ary, in general) fractions up to a given cognitive bound ρ; if ρ is a natural number k, the set of permissible probabilities is given as $\Pi_\rho = \Pi_k = \{\frac{0}{10^k}, \frac{1}{10^k}, ..., \frac{10^k}{10^k}\}$. The decision maker makes preference comparisons step by step using probabilities with small k to those with larger k' to obtain accurate comparisons. The derived preference relation is incomplete in general, but the EU hypothesis holds for some lotteries and would hold more when there is no cognitive bound, i.e., $\rho = \infty$. Our main concern is the case $\rho < \infty$. Since the theory involves various entangled aspects, we first disentangle them.

The concepts of *probability grids* and *cognitive bounds* are introduced based on the idea of "bounded rationality". This idea can be interpreted in many ways such as bounded logical

*The author thanks J. J. Kline, P. Wakker, M. Lewandowski, S. Shiba, M. Cohen, O. Shulte, and Y. Rebille for helpful comments on earlier versions of this paper. In particular, comments given by the two referees improved the paper significantly and are greatly appreciated. The author is supported by Grant-in-Aids for Scientific Research No.26245026 and No.17H02258, Ministry of Education, Science and Culture.

†Waseda University, Tokyo, Japan (mkanekoepi@waseda.jp)

inference, bounded perception ability, though Simon's [30] original concept meant a relaxation of utility maximization. The mathematical components involved in EU theory are classified to two types; object-components used by the decision maker and meta-components used by the outside analyst and possibly by the decision maker himself. The former is primary targets in EU theory, and the latter such as highly complex rational as well as irrational probabilities is added for analytic convenience. A free use of the latter leads to a critique that the theory presumes "super rationality" (Simon [31]).

As a significance level for statistical hypothesis testing is typically 5% or 1%, probability values $\frac{t}{10^2}$ $(t = 0, ..., 10^2)$ are already quite accurate for ordinary people. However, the classical EU theory starts with the full real number theory and makes no separation between the viewpoints of the decision maker and the outside analyst for available probabilities. This is still a problem of degree, but it would be meaningful if they are separated in some manner. The concepts of probability grids and a cognitive bound ρ make this separation.

Turing [32] in his attempt to define computable numbers faced a similar situation: "... *The differences from our point of view between the single and compound symbols is that the compound symbols, if they are too lengthy, cannot be observed at one glace. This is in accordance with experience. We cannot tell at a glance whether* $\underline{0.}$*9999999999999999 and* $\underline{0.}$*999999999999999* (the underlined parts by the author) *are the same* ([32], p.250)."[1] In contrast, it is easy for us to distinguish between 0.999 and 0.99. Turing's theory tried to abstract calculation in the human mind, while a built machine has no such problem since it reads each primitive symbol but a compound symbol not as one symbol, like a human. This is nature of human cognition. A cognitive bound ρ is a bound on such distinguishability and on how deep the decision maker cares about those probabilities.

The set of probability grids up to depth k is given as $\Pi_k = \{\frac{0}{10^k}, \frac{1}{10^k}, ..., \frac{10^k}{10^k}\}$. The decision maker thinks about his preferences with Π_k from a small k to a larger k up to bound ρ; for example, when $\rho = 2$, Π_0, Π_1, and Π_2 are only allowed. This is a constructive approach from the viewpoint of the decision maker in the sense that he finds/forms his own preferences.[2,3]

We turn our attention to the development of our constructive EU theory. Constructiveness needs a start; we take a hint from von Neumann-Morgenstern [33]. They divided the motivating argument into the following two, though this separation was not reflected in their development:

Step B: measurements of utilities from pure alternatives in terms of probabilities;

Step E: extensions of these measurements to lotteries involving more risks.

These steps differ in their natures: Step B is to measure a "satisfaction", "desire", etc. from a pure alternative, while Step E is to extend the measured satisfactions given by Step B to lotteries including more risks. An important difference is that Step B is to find the subjective preferences hidden in the mind of the decision maker, while Step E is to extend logically the

[1] I thank Oliver Schulte for mentioning this quotation

[2] This sounds similar to "constructive decision theory" in Shafer [28], [29] and in Blume *et al.* [4]. These authors study Savage's [27] subjective utility/probability theory so as to introduce certain constructive features for decision making. Our theory is constructive more explicitly with the introduction of probability grids and a cognitive bound. The chief difference is that we formulate how a decision maker finds/forms his own preferences, while they add new constructs like "goals" or "frames" that shape the choices of the decision maker.

[3] Our concept of probability grids may be interpreted as "imprecise probabilities/similarity" (cf. Augustin *et al.* [2], Rubinstein [26]). Imprecision/similarity is defined as an attribute of a probability/a set of probabilities, allowing all real number probabilities. In our approach, however, probability grids in Π_k are exact; the restriction of probabilities to Π_k expresses imprecision in cognitive acts taken by the decision maker.

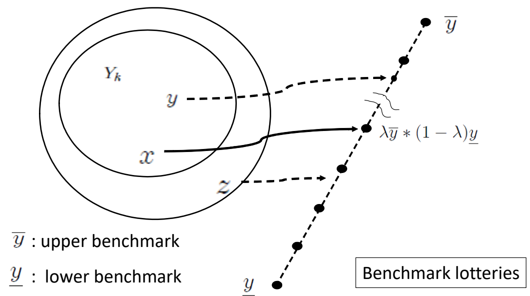

Figure 1: Step B with the benchmark scale

preferences found in Step B to lotteries with more risks.

We develop our theory based on the above two steps and also take two approaches in terms of preferences and utilities; each approach consists of Steps B and E. In this introduction, we focus mainly on the former theory, and we give a brief explanation of the latter.[4]

We assume two pure alternatives $\overline{y}$ and $\underline{y}$, called the *upper and lower benchmarks*; these together with the probability grids Π_k form the *benchmark scale* $B_k(\overline{y}; \underline{y})$ in layer k. In Step B, pure alternatives are measured by this scale. Preferences are constructed in shallow to deeper layers, where preferences are incomplete in the beginning, except for benchmark lotteries as measurement units, and in deeper layers, more precise preferences may be found. In Fig.1, the benchmark scale for layer k is depicted as the right broken line with dots; x is measured exactly by the scale, y need a more precise scale within ρ. However, z is not done within ρ.

Two different roles of probability grids appear in Step E for evaluation of a lottery:

(i) probability grids used for measurement of a pure alternative in Step B;

(ii) probability coefficients to pure alternatives.

By these, relevant cognitive depths of lotteries become more complex especially with a finite cognitive bound; this leads to incomparabilities in preferences and some violation of the EU hypothesis. This is central in our development and is closely related to the issue of "bounded rationality". Let us illustrate (i) and (ii) via an example.

Consider one example with the upper and lower benchmarks $\overline{y}$, $\underline{y}$, and the third pure alternative y with strict preferences $\overline{y} \succ y \succ \underline{y}$. In Step B, the decision maker looks for a probability λ so that y is indifferent to a lottery $[\overline{y}, \lambda; \underline{y}] = \lambda\overline{y} * (1-\lambda)\underline{y}$ with probability λ for $\overline{y}$ and $1-\lambda$ for $\underline{y}$; this indifference is denoted by

$$y \sim [\overline{y}, \lambda; \underline{y}]. \tag{1}$$

Suppose that this λ is uniquely determined as $\lambda = \lambda_y = \frac{83}{10^2} \in \Pi_2$. Here, exact measurement of y is successful in layer 2, where Step B is enough here.

We have the other source of cognitive depths. Consider lottery $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$, which includes the third pure alternative y. The independence condition of the classical EU theory dictates

[4] Our theory is dual to that in terms of certainty equivalent of a lottery (cf. Kontek-Lewandowski [21] and its references). In our method, the set of benchmark lotteries forms a base scale, while the set of monetary amounts is the base scale in the latter (see Section 4.2 in [21]).

that because of (1), $[\overline{y}, \frac{83}{10^2}; \underline{y}]$ is substituted for y in d, and d is reduced to:

$$d = \tfrac{25}{10^2} y * \tfrac{75}{10^2} \underline{y} \sim \tfrac{25}{10^2} [\overline{y}, \tfrac{83}{10^2}; \underline{y}] * \tfrac{75}{10^2} \underline{y} = \tfrac{2075}{10^4} \overline{y} * \tfrac{7925}{10^4} \underline{y}. \tag{2}$$

Thus, y is evaluated as being indifferent to $[\overline{y}, \frac{83}{10^2}; \underline{y}]$ in Step B, but y also has a probability coefficient $\frac{25}{10^2}$ in d, which is taken into account in Step E. These steps leads to probability $\frac{2075}{10^4}$, which is much more precise than either of $\frac{83}{10^2}$ and $\frac{25}{10^2}$.

As indicated in (i) and (ii), lottery $d = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$ has two types of cognitive depths; one is simply a probability coefficient $\frac{25}{10^2}$ and the other is $\lambda_y = \frac{83}{10^2}$ from (1). Although d itself is expressed as a lottery of depth 2, the total depths including these two types is 4, which is beyond the cognitive bound $\rho = 2$. One point is that the resulting probability may be very precise with a relatively small cognitive bound, and the other is that this is intimately related to the EU hypothesis. When ρ is small, the EU hypothesis does not typically hold, while it would hold more as ρ is getting larger.

The preference formation by Steps B and E is formulated as a form of mathematical induction; Step B is the inductive base and Step E is the inductive step. Step B is spread out to layers of various depths, i.e., the induction base is spread too. These steps are described in Table 1.1: the relation $\trianglerighteq_k$ for layer k of row B expresses preferences measured in Steps B. In layer k, $\succsim_k$ is derived from $\trianglerighteq_k$ and $\succsim_{k-1}$; the former is a part of the inductive base and the latter is the inductive step. This is a weak form of "independence condition".

Table 1.1

Layers	0		1	...	$k-1$		k	...	ρ
B: base relations	$\trianglerighteq_0$	$\subseteq$	$\trianglerighteq_1$	$\subseteq \ldots \subseteq$	$\trianglerighteq_{k-1}$	$\subseteq$	$\trianglerighteq_k$	$\subseteq \ldots \subseteq$	$\trianglerighteq_\rho$
	$\downarrow$		$\downarrow$		$\downarrow$		$\downarrow$		$\downarrow$
E: constructed relations	$\succsim_0$	$\rightarrow$	$\succsim_1$	$\rightarrow \ldots \rightarrow$	$\succsim_{k-1}$	$\rightarrow$	$\succsim_k$	$\rightarrow \ldots \rightarrow$	$\succsim_\rho$

We also provide another approach in terms of a 2-dimensional vector-valued utility functions $\langle \boldsymbol{v}_k \rangle_{k<\rho+1} = \langle [\overline{v}_k, \underline{v}_k] \rangle_{k<\rho+1}$ and $\langle \boldsymbol{u}_k \rangle_{k<\rho+1} = \langle [\overline{u}_k, \underline{u}_k] \rangle_{k<\rho+1}$ with Fishburn's [8] interval order $\geq_I$. In each of Steps B and E, this approach is entirely equivalent to the preference approach, depicted in Table 1.2. This may be interpreted as what von Neumann- Morgenstern [33], p.29 indicated. The approaches in terms of preferences and utilities enable us to view Steps B and E in different ways as well as serve different analytic tools for studies of incomparabilities/comparabilities involved.

Table 1.2

Preference theory		Utility theory
Step B (B0 to B3)	$\Longleftrightarrow$ (Sec.3)	Step B (b0 to b3)
$\downarrow\downarrow$ Extension (Sec.4)		$\downarrow\downarrow$ Extension (Sec.5)
Step E (E0 to E3)	$\Longleftrightarrow$ (Sec.5)	Step E (e0 to e3)

Our theory enjoys a weak form of the expected utility hypothesis. This will be discussed in Section 6. In the case of $\rho = \infty$, restricting our attention to the set of measurable pure alternatives, in Section 7, we show that our theory exhibits a form of the classical EU theory. We provide a further extension of $\succsim_\infty$ to have the full form of classical EU theory; this extension involves some unavoidable non-constructive step, which may be interpreted as the criticism of "super rationality" by Simon [31].

We apply our theory to the Allais paradox, specifically, to an experimental result from Kahneman-Tversky [15]. We show that the paradoxical results remains when the cognitive bound $\rho \geq 3$. However, when $\rho = 2$, the resultant preference relation $\succsim_\rho$ is compatible with their experimental result, where incomparabilities play crucial roles in explaining them.

A remark is on the relationship between k and ρ exhibiting a layer and a cognitive bound. The former is a variable in our theory and the latter is a parameter of the theory. We talk about the sequences $\langle \succsim_k \rangle_{k<\rho+1}$ and $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ describing the process of preference formation layer to layer up to ρ. Nevertheless, the final target preferences and utilities are $\succsim_\rho$ and u_ρ. In the context of the quotation from Turing [32], within the layers up to ρ, the decision maker can distinguish each probability as a single symbol but beyond ρ, he would have a difficulty; it is assumed here that he does not think about his decision problem beyond ρ. When $\rho = \infty$, he can treat any grid probability as a single entity. This remark leads to the view that our theory is a generalization of the classical EU theory, which is discussed in Section 7.

The paper is organized as follows: Section 2 explains the concept of probability grids and other basic concepts. Section 3 formulates Step B in terms preferences and utilities, and states their equivalence. Section 4 discusses Step E in terms of preferences and Section 5 does it in terms of utilities. Section 6 discusses the measurable/non-measurable lotteries, and shows that the expected utility hypothesis holds for the measurable lotteries. Section 7 discusses the connection from our theory to the classical EU theory. In Section 8, we exemplify our theory with an experimental result in Kahneman-Tversky [15]. Section 9 concludes this paper with comments on further possible studies. Proofs of all the results in each section are given in a separate subsection; only proof of Lemma 2.1 is given in Section 10.

2 Preliminaries

Our theory is about preference formation in the context of EU theory. The classical EU theory is the reference point, but our theory deviates from it in various manners. To have clear relations between the classical EU theory and our development, we first mention the classical theory (cf. Herstein-Milnor [13], Fishburn [11]), and then, we start our development. In Section 2.2, we give various basic concepts for our theory and one basic lemma. In Section 2.3, we give definitions of preferences, indifferences, incomparabilities, and their counterparts in terms of vector-valued utility functions.

2.1 Classical EU theory

Let X be a given set of *pure alternatives* with cardinality $|X| \geq 2$. A *lottery* f is a function over X taking real values in $[0,1]$ with $\sum_{x \in S} f(x) = 1$ for some finite subset S of X. This subset S is called a *support* of f. We define $L_{[0,1]}(X) = \{f : f : X \to [0,1] \text{ is a lottery}\}$. The set $L_{[0,1]}(X)$ is uncountable. We define *compound lotteries*: for any $f, g \in L_{[0,1]}(X)$ and $\lambda \in [0,1]$, $\lambda f * (1-\lambda) g$ is a lottery in $L_{[0,1]}(X)$ defined by $(\lambda f * (1-\lambda)g)(x) = \lambda f(x) + (1-\lambda) g(x)$ for all $x \in X$.

Let $\succsim_E$ be a binary relation over $L_{[0,1]}(X)$; and we assume NM0 to NM2 on $\succsim_E$. This system is one among various equivalent systems.

Axiom NM0 (Complete preordering): $\succsim_E$ is a complete and transitive relation on $L_{[0,1]}(X)$.

Axiom NM1 (Intermediate value): For any $f, g, h \in L_{[0,1]}(X)$, if $f \succsim_E g \succsim_E h$, then $\lambda f * (1-\lambda)h \sim_E g$ for some $\lambda \in [0,1]$.

Axiom NM2 (Independence): For any $f, g, h \in L_{[0,1]}(X)$ and $\lambda \in (0,1]$,

ID1: $f \succ_E g$ implies $\lambda f * (1-\lambda)h \succ_E \lambda g * (1-\lambda)h$;

ID2: $f \sim_E g$ implies $\lambda f * (1-\lambda)h \sim_E \lambda g * (1-\lambda)h$,

where the *indifference* part and *strict preference* part of $\succsim_E$ are denoted by $\sim_E$ and $\succ_E$; that is, $f \sim_E g$ means $f \succsim_E g$ & $g \succsim_E f$; and $f \succ_E g$ does $f \succsim_E g$ & not $(g \succsim_E f)$.

The following two are the key theorems in the classical EU theory. For a fruitful development of our theory, we should be conscious of how they remain in our theory.

Theorem 2.1 (Classical EU theorem). A preference relation $\succsim_E$ satisfies Axioms NM0 to NM2 if and only if there is a function $u : X \to \mathbb{R}$ so that for any $f, g \in L_{[0,1]}(X)$,

$$f \succsim_E g \text{ if and only if } E_f(u) \geq E_f(u), \tag{3}$$

where the *expected utility functional* $E_f(u)$ is defined as:

$$E_f(u) = \textstyle\sum_{x \in S} \; f(x)u(x) \text{ for each } f \in L_{[0,1]}(X) \text{ with its support } S. \tag{4}$$

Theorem 2.2 (Uniqueness up to Affine transformations). Suppose that $\succsim_E$ satisfies Axioms NM0 to NM2. If two functions $u, v : X \to \mathbb{R}$ satisfy (3), then there are two real numbers $\alpha > 0$ and β such that $u(x) = \alpha v(x) + \beta$ for all $x \in X$.

In these theorems, preference relation $\succsim_E$ is given with Axioms NM0 to NM2. The theory is silent about how a decision maker finds/forms his preferences. As mentioned in Section 1, we consider this question from simple cases to more complex cases, while distinguishing between Steps B and E. In the above axiomatization, these are mixed in NM1 and NM2. We will make a clear-cut distinction between Steps B and E. In these steps, we avoid the existence of a complete preference relation dictated by Axiom NM0. Another salient restriction in our theory is on the available probabilities and is formulated by the concept of probability grids. This allows us to think about his preferences from simpler lotteries to complex ones step by step. The step-by-step consideration collapses in Axioms NM2 and NM3.

The system $(L_{[0,1]}, \succsim_E)$ with Axioms NM0 to NM2 itself is not in the central part of our theory, but our theory is closely related to the EU hypothesis that preferences are represented by the expected utility functional $E_f(u)$. We will discuss the *EU* hypothesis time to time, and touch the system $(L_{[0,1]}, \succsim_E)$ only in Section 7.

Aumann [3] and Fishburn [9] considered one-way representation theorem (i.e., the *only-if* of (3)), dropping completeness. See Fishburn [10] for further studies. Dubra-Ok [6] and Dubra, *et al.* [7] developed representation theorems in terms of utility comparisons based on all possible expected utility functions for the relation without completeness. In this literature, incomparabilities are given in the preference relation. In contrast, in our approach, incomparabilities are changing with a cognitive bound and may disappear when there are no cognitive bounds.

2.2 Probability grids, lotteries, and decompositions

Let ℓ be an integer with $\ell \geq 2$. This ℓ is the base for describing probability grids; we take $\ell = 10$ in the examples in the paper. The *set of probability grids* Π_k is defined as

$$\Pi_k = \{\tfrac{\nu}{\ell^k} : \nu = 0, 1, ..., \ell^k\} \text{ for any finite } k \geq 0. \tag{5}$$

Here, $\Pi_1 = \{\frac{\nu}{\ell} : \nu = 0, ..., \ell\}$ is the base set of probability grids for measurement, whereas $\Pi_0 = \{0, 1\}$ is needed for technical completeness. Each Π_k is a finite set, and $\Pi_\infty := \cup_{t<\infty}\Pi_t$ is countably infinite. We use the standard arithmetic rules over Π_∞; sum and multiplication are needed.[5] We allow reduction by eliminating common factors; for example, $\frac{20}{10^2}$ is the same as $\frac{2}{10}$. Hence, $\Pi_k \subseteq \Pi_{k+1}$ for $k = 0, 1, ...$ The parameter k is the precision of probabilities that the decision maker uses. We define the *depth* of each $\lambda \in \Pi_\infty$ by: $\delta(\lambda) = k$ iff $\lambda \in \Pi_k - \Pi_{k-1}$. For example, $\delta(\frac{25}{10^2}) = 2$ but $\delta(\frac{20}{10^2}) = \delta(\frac{2}{10}) = 1$. The concept of a *layer* of probability grids up to a given depth k is well defined. The decision maker thinks about his preferences along probability grids from a shallow layer to a deeper one.

We use the standard equality $=$ and strict inequality $>$ over Π_k. Then, trichotomy holds: for any $\lambda, \lambda' \in \Pi_k$,

$$\text{either } \lambda > \lambda',\ \lambda = \lambda',\ \text{or } \lambda < \lambda'. \tag{6}$$

Each element in Π_k is obtained by taking the weighted sums of elements in Π_{k-1} with the equal weights:

$$\Pi_k = \{\sum_{t=1}^{\ell} \tfrac{1}{\ell}\lambda_t : \lambda_1, ..., \lambda_\ell \in \Pi_{k-1}\} \text{ for any } k\ (1 \le k < \infty). \tag{7}$$

This is basic for the connection between layer $k-1$ to the next. A proof of (7) is not given here, but an extension will be given in Lemma 2.1 with a proof given in the Appendix.

The union $\Pi_\infty = \cup_{k<\infty}\Pi_k$ is a proper subset of $[0, 1] \cap \mathbb{Q}$, where $\mathbb{Q}$ is the set of rational numbers. For example, when $\ell = 10$, Π_∞ has no recurring decimals, but they are rationals. We also note that Π_∞ depends upon the base ℓ; for example, Π_1 with $\ell = 3$ has $\frac{1}{3}$, but Π_∞ with $\ell = 10$ has no element corresponding to $\frac{1}{3}$.

For any $k < \infty$, we define $L_k(X)$ by

$$L_k(X) = \{f : f \text{ is a function from } X \text{ to } \Pi_k \text{ with } \textstyle\sum_{x\in X} f(x) = 1\}. \tag{8}$$

We identify each pure alternative x with the lottery having x as its support; so X is regarded as a subset of $L_k(X)$. Specifically, $L_0(X) = X$. Since Π_k is a finite set, every $f \in L_k(X)$ has a finite support. Since $\Pi_k \subseteq \Pi_{k+1}$, it holds that $L_k(X) \subseteq L_{k+1}(X)$. We denote $L_\infty(X) = \cup_{k<\infty}L_k(X)$. As long as X is finite, $L_k(X)$ is also a finite set, but $L_\infty(X)$ is a countable set and is dense in $L_{[0,1]}(X)$.

We define the *depth of a lottery* f in $L_\infty(X)$ by $\delta(f) = k$ iff $f \in L_k(X) - L_{k-1}(X)$. We use the same symbol δ for the depth of a lottery and the depth of a probability. It holds that $\delta(f) = k$ if and only if $\max_{x\in X} \delta(f(x)) = k$. This is relevant in Section 6. Lottery $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$ is in $L_2(X) - L_1(X)$ and its depth $\delta(d) = 2$; but since $d' = \frac{20}{10^2}y * \frac{80}{10^2}\underline{y} = \frac{2}{10}y * \frac{8}{10}\underline{y} \in L_1(X)$, we have $\delta(d') = 1$.

The decision maker thinks about preferences from shallow layers to deeper ones. This stops at a *cognitive bound* ρ, which is a natural number or infinity ∞. If $\rho = k < \infty$, he eventually reaches the set of lotteries $L_\rho(X) = L_k(X)$, and if $\rho = \infty$, he has no cognitive limit; we define $L_\rho(X) = L_\infty(X) = \cup_{k<\infty}L_k(X)$.

We formulate a connection from $L_{k-1}(X)$ to $L_k(X)$. We say that $\widehat{f} = (f_1, ..., f_\ell)$ in $L_{k-1}(X)^\ell = L_{k-1}(X) \times \cdots \times L_{k-1}(X)$ is a *decomposition* of $f \in L_k(X)$ iff for all $x \in X$,

$$f(x) = \textstyle\sum_{t=1}^{\ell} \frac{1}{\ell} \times f_t(x) \text{ and } \delta(f_t(x)) \le \delta(f(x)) \text{ for all } t \le \ell. \tag{9}$$

[5] See Mendelson [23] for related basic mathematics.

We denote this by $\sum_{t=1}^{\ell} \frac{1}{\ell} * f_t$, and letting $\widehat{e} = (\frac{1}{\ell}, ..., \frac{1}{\ell})$, it is written as $\widehat{e}{*}\widehat{f}$. We can regard $\widehat{e}{*}\widehat{f}$ as a compound lottery connecting $L_{k-1}(X)$ to $L_k(X)$ by reducing $\widehat{e}{*}\widehat{f}$ to f in (9). Our theory allows only this form of a compound lotteries and reduction with the depth constraint. The next lemma states that $L_k(X)$ is generated from $L_{k-1}(X)$ by taking all compound lotteries of this kind. It facilitates our induction method described in Table 1.1 reducing an assertion in layer k to layer $k-1$. A proof of Lemma 2.1 is given in Section 10.[6]

Lemma 2.1 (Decomposition of lotteries). Let $1 \leq k < \infty$. Then,

$$L_k(X) = \{f \in L_k(X) : f \text{ has a decomposition } \widehat{f}\}. \tag{10}$$

Furthermore, for any $f \in L_k(X)$ with $\delta(f) > 0$, there is a decomposition of $\widehat{f}$ of f so that

$$\delta(f_t(x)) < \delta(f(x)) \text{ for any } x \in X \text{ with } f(x) > 0. \tag{11}$$

The right-hand side of (10) is the set of composed lotteries from $L_{k-1}(X)$ with the equal weights. The inclusion $\supseteq$ states that the composed lotteries from $L_{k-1}(X)$ belong to $L_k(X)$. The converse inclusion $\subseteq$ is essential and means that each lottery in $L_k(X)$ is decomposed to an equally weighted sum of some $(f_1, ..., f_\ell)$ in $L_{k-1}(X)^\ell$ with the depth constraint in (9). In the trivial case that $f = x \in L_0(X)$ is decomposed to $\widehat{f} = (x, ..., x)$. This will be used in Proposition 4.1.(2). The latter asserts the choice of a strictly shallower decomposition for f with $\delta(f) > 0$.

One remark is that when f is a benchmark lottery in $B_k(\overline{y}; \underline{y})$, for its decomposition $\widehat{f} = (f_1, ..., f_\ell)$, each f_t is a benchmark lottery in $B_{k-1}(\overline{y}; \underline{y})$. This fact will be used without referring.

For the set of lotteries over subset X' of X, i.e., $L_k(X')$, we introduce the following convention. We define $L_k(X') = \{f \in L_k(X) : f(x) > 0 \text{ implies } x \in X'\}$. Hence, $L_k(X')$ is a subset of $L_k(X)$. Lemma 2.1 hods for $L_k(X')$ and $L_{k-1}(X')$.

The lottery $d = [y, \frac{25}{10^2}; \underline{y}]$ has three types of decompositions:

$$d = \tfrac{t}{10} * y + \tfrac{5-2t}{10} * [y, \tfrac{5}{10}; \underline{y}] + \tfrac{5+t}{10} * \underline{y} \ \text{ for } t = 0, 1, 2. \tag{12}$$

Here, a decomposition $\widehat{f} = (f_1, ..., f_{10})$ is given as $f_1 = ... = f_t = y$, $f_{t+1} = ... = f_{5-t} = [y, \frac{5}{10}; \underline{y}]$ and $f_{5-t+1} = ... = f_{10} = \underline{y}$. We use this short-hand expressions rather than a full specification of $\widehat{f} = (f_1, ..., f_{10})$. We should be careful about this multiplicity.

The reason for explicit considerations of layers for $L_k(X)$ and also preference relation $\succsim_k$ is to avoid collapse from a layer to a shallower one. Without them, we may have a difficulty in identifying the sources for preferences. For example, the weighted sum $\frac{5}{10}[\frac{25}{10^2}y * \frac{75}{10^2}\underline{y}] * \frac{5}{10}[\frac{75}{10^2}y * \frac{25}{10^2}\underline{y}]$ is reduced to $\frac{5}{10}y * \frac{5}{10}\underline{y}$; preferences about $\frac{5}{10}y * \frac{5}{10}\underline{y}$ may possibly come from layer 2 or from layer 0. To prohibit such collapse, we take explicitly depths of layers into account in (9).

2.3 Incomplete preference relations and vector-valued utility functions

We consider two methods to represent the decision maker's desires: a preference relation and a utility function. We starts with incomplete preferences and, correspondingly, representing

[6] When $\ell > 2$, binary decompositions are not enough for Lemma 2.1, For example, consider lottery $f = \frac{3}{10}\overline{y} * \frac{3}{10}y * \frac{4}{10}\underline{y}$. This is not expressed by a binary combination of elements in $L_0(X) = X$ with weights in Π_1.

utility functions become vector-valued with the interval order. These are first departures from the classical EU theory.

Let $\succsim$ be a preference relation over a given set, say A. For $f, g \in A$, the expression $f \succsim g$ means that f is strictly preferred to g or is indifferent to g. We define the *strict (preference) relation* $\succ$, *indifference relation* $\sim$, and *incomparability relation* $\bowtie$ by

$$\begin{aligned} f \ &\succ\ g \text{ if and only if } f \succsim g \text{ and not } g \succsim f; \\ f \ &\sim\ g \text{ if and only if } f \succsim g \text{ and } g \succsim f; \\ f \ &\bowtie\ g \text{ if and only if neither } f \succsim g \text{ nor } g \succsim f. \end{aligned} \tag{13}$$

All the axioms are given on the relations $\succsim$, $\succ$, $\sim$, and the relation $\bowtie$ is defined as the residual part of $\succsim$. Although $\sim$ and $\bowtie$ are sometimes regarded as closely related (cf. Shafer [28], p.469), they are well separated in Theorem 6.2 in our theory.

In the classical theory in Section 2.1, the preference relation $\succsim_E$ is assumed to be complete. Since, however, we consider a formation of preferences, our theory should avoid this completeness assumption. Nevertheless, it appears as a result when a domain of lotteries is restricted.

Another method of measurement of desires is by a vector-valued function $\boldsymbol{u}$ with the interval order introduced by Fishburn [8]. Let $\boldsymbol{u}(f) = [\overline{u}(f), \underline{u}(f)]$ be a 2-dimensional vector-valued function from its domain A to the set $\mathbb{Q}^2 = \mathbb{Q} \times \mathbb{Q}$ with $\overline{u}(f) \geq \underline{u}(f)$ for each $f \in A$. The components $\overline{u}(f)$ and $\underline{u}(f)$ are interpreted as the least upper and greatest lower bounds of possible utilities from f. We say that $\boldsymbol{u}(f)$ is *effectively single-valued* iff $\overline{u}(f) = \underline{u}(f)$; in this case, we write $\overline{u}(f) = \underline{u}(f) = u(f)$, dropping the upper and lower bars. We use the interval order $\geq_I$ over the values of $\boldsymbol{u}$; for $f, g \in A$,

$$\boldsymbol{u}(f) \geq_I \boldsymbol{u}(g) \ \text{ if and only if } \ \underline{u}(f) \geq \overline{u}(g). \tag{14}$$

That is, f and g are ordered if and only if the greatest lower bound $\underline{u}(f)$ from f is larger than or equal to the least upper bound $\overline{u}(g)$ from g. This $\geq_I$ allows incomparabilities, for example, if $\boldsymbol{u}(f) = [\frac{9}{10}, \frac{7}{10}]$ and $\boldsymbol{u}(g) = [\frac{83}{10^2}, \frac{83}{10^2}]$, then f and g are incomparable by $\geq_I$. The relation $\geq_I$ is transitive, but $(\boldsymbol{u}(f) \geq_I \boldsymbol{u}(g)$ & $\boldsymbol{u}(g) \geq_I \boldsymbol{u}(f))$ is equivalent to $u(f) = \underline{u}(f) = \overline{u}(g) = \underline{u}(g)$, i.e., this is the case only when the values $\boldsymbol{u}(f)$ and $\boldsymbol{u}(g)$ are effectively single-valued and identical.

3 Measurement Step

We formulate Step B of measurement of pure alternatives up to cognitive bound ρ. This has two sides: in terms of preference relations $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ and in terms of vector-valued utility $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$. We show the representation theorem on $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ by $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$, and the uniqueness theorem on $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ up to positive linear transformations. Finally, we mention that these are well interpreted in terms of Simon's [30] satisficing/aspiration argument.

3.1 Base preference streams

The set of pure alternatives X is assumed to contain two distinguished elements $\overline{y}$ and $\underline{y}$, which we call the *upper* and *lower benchmarks*. Let $k < \infty$. We call an $f \in L_k(X)$ a *benchmark lottery* of depth (at most) k iff $f(\overline{y}) = \lambda$ and $f(\underline{y}) = 1 - \lambda$ for some $\lambda \in \Pi_k$, which we denote by $[\overline{y}, \lambda; \underline{y}]$. The *benchmark scale of* depth k is the set $B_k(\overline{y}; \underline{y}) = \{[\overline{y}, \lambda; \underline{y}] : \lambda \in \Pi_k\}$. In particular, $B_0(\overline{y}; \underline{y}) = \{\overline{y}, \underline{y}\}$. The dots in Fig.1 express the benchmark lotteries. We define $B_\infty(\overline{y}; \underline{y}) = \cup_{k<\infty} B_k(\overline{y}; \underline{y})$. The depth of a benchmark lottery $[\overline{y}, \lambda; \underline{y}]$ is determined to be the depth of λ, i.e., $\delta([\overline{y}, \lambda; \underline{y}]) = \delta(\lambda)$.

We denote a *cognitive bound* by ρ, which is a natural number or $\rho = \infty$. We use k as a *variable* expressing a natural number of a layer within the theory, but ρ as a constant parameter of it. Stipulating $\infty + 1 = \infty$, "$k < \rho + 1$" expresses the two statements "$k \leq \rho$ if $\rho < \infty$" and "$k < \rho$ if $\rho = \infty$". This constant ρ plays an active role as a small constraint such as $\rho = 2$ or 3 in Example 5.1 and Section 8, and as $\rho = \infty$ in Section 7 for consideration of the expected utility hypothesis.

Let $\trianglerighteq_k$ be a subset of

$$D_k = B_k(\overline{y}; \underline{y})^2 \cup \{(x, g), (g, x) : x \in X \text{ and } g \in B_k(\overline{y}; \underline{y})\}. \tag{15}$$

Thus, $\trianglerighteq_k$ consists of the scale part of the benchmarks and the measurement part of pure alternatives. The scale part allows the decision maker to make comparisons between any grids of depth k. For a pure alternative $x \in X$, he thinks about where x is located in the benchmark scale $B_k(\overline{y}; \underline{y})$; it may or may not correspond to a grid, which is seen in Fig.1. For example, if $(x, g) \in \trianglerighteq_k$ but $(g, x) \notin \trianglerighteq_k$, then x is strictly better than the grid g; and if $(x, g) \notin \trianglerighteq_k$ and $(g, x) \notin \trianglerighteq_k$, then x and g are incomparable for him.

We make four axioms on $\langle \trianglerighteq_k \rangle_{k<\rho+1}$. Axiom B0 requires pure alternatives be between the upper and lower benchmarks $\overline{y}$, $\underline{y}$.

Axiom B0 (Benchmarks): $\overline{y} \trianglerighteq_0 x$ and $x \trianglerighteq_0 \underline{y}$ for all $x \in X$.

The next states that preferences over $B_k(\overline{y}; \underline{y})$ are the same as the natural order on Π_k.

Axiom B1 (Benchmark scale): For $\lambda, \lambda' \in \Pi_k$, $[\overline{y}, \lambda; \underline{y}] \trianglerighteq_k [\overline{y}, \lambda'; \underline{y}]$ if and only if $\lambda \geq \lambda'$

It follows from Axiom B1 that for $\lambda, \lambda' \in \Pi_k$,

$$[\overline{y}, \lambda; \underline{y}] \triangleright_k [\overline{y}, \lambda'; \underline{y}] \text{ if and only if } \lambda > \lambda'. \tag{16}$$

Also, $\lambda = \lambda'$ if and only if $[\overline{y}, \lambda; \underline{y}]$ and $[\overline{y}, \lambda'; \underline{y}]$ are indifferent. Thus, $\trianglerighteq_k$ is a complete relation over $B_k(\overline{y}; \underline{y})$ by (6). This is the scale part of $\trianglerighteq_k$, and is precise up to Π_k. Since $\overline{y} = [\overline{y}, 1; \underline{y}]$ and $\underline{y} = [\overline{y}, 0; \underline{y}]$, it follows from (16) that $\overline{y} \succ_{B,0} \underline{y}$.

Measurement is required to be coherent with the scale part given by Axiom B1.

Axiom B2 (Monotonicity): For all $x \in X$ and $\lambda, \lambda' \in \Pi_k$, if $[\overline{y}, \lambda; \underline{y}] \trianglerighteq_k x$ and $\lambda' > \lambda$, then $[\overline{y}, \lambda'; \underline{y}] \triangleright_k x$; and if $x \trianglerighteq_k [\overline{y}, \lambda; \underline{y}]$ and $\lambda > \lambda'$, then $x \triangleright_k [\overline{y}, \lambda'; \underline{y}]$.

This implies no reversals with Axiom B1; if $[\overline{y}, \lambda; \underline{y}] \trianglerighteq_k x$ and $x \trianglerighteq_k [\overline{y}, \lambda'; \underline{y}]$, then $\lambda \geq \lambda'$. Indeed, if $\lambda < \lambda'$, then $[\overline{y}, \lambda'; \underline{y}] \triangleright_k x$ by B2, which implies not $x \trianglerighteq_k [\overline{y}, \lambda'; \underline{y}]$. If we assume transitivity for $\trianglerighteq_k$ over D_k, B2 could be derived from B1, but we adopt B2 instead of transitivity, since B2 gives a more specific property to the measurement step.

The last requires the preferences in layer k be preserved in the next layer $k + 1$. This is expressed by the set-theoretical inclusion $\subseteq$ in Table 1.1.

Axiom B3 (Preservation): For all $f, g \in D_k$, $f \trianglerighteq_k g$ implies $f \trianglerighteq_{k+1} g$.

The above axioms still allow great freedom for base preference relations $\langle \trianglerighteq_k \rangle_{k<\rho+1}$. To see this fact as well as how the measurement step B of utilities from pure alternatives goes on, we consider vector-valued utility functions with the interval order $\geq_I$ in Section 3.2.

3.2 Base utility streams

We consider another way of Step B in terms of vector-valued utility functions with the interval order $\geq_I$. Let $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1} = \langle [\overline{\upsilon}_k, \underline{\upsilon}_k] \rangle_{k<\rho+1}$ be a sequence of vector-valued functions so that for each $k < \rho + 1$, $\boldsymbol{\upsilon}_k$ is a function from $B_k(\overline{y}; \underline{y}) \cup X$ to $\mathbb{Q}^2$ such that $\overline{\upsilon}_k(f) \geq \underline{\upsilon}_k(f)$ for all $f \in B_k(\overline{y}; \underline{y}) \cup X$. Recall that when $\boldsymbol{\upsilon}_k(f)$ is effectively single-valued, we write $\overline{\upsilon}_k(f) = \underline{\upsilon}_k(f) = \upsilon_k(f)$. The following conditions on $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ are not exactly parallel to the axiomatic system B0 to B3, but these two systems are equivalent, which is stated in Theorem 3.1.

We define a *base (upper-lower) utility stream* $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1} = \langle [\overline{\upsilon}_k, \underline{\upsilon}_k] \rangle_{k<\rho+1}$ by $b0$ to $b3$:

***b*0**: $\upsilon_0(\overline{y}) > \upsilon_0(\underline{y})$;

and for $k < \rho + 1$,

***b*1**: $\upsilon_k([\overline{y}, \lambda; \underline{y}]) = \lambda \upsilon_k(\overline{y}) + (1 - \lambda)\upsilon_k(\underline{y})$ for all $[\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$;

***b*2**: for each $x \in X$, $\overline{\upsilon}_k(x) = \upsilon_k([\overline{y}, \overline{\lambda}_x; \underline{y}])$ and $\underline{\upsilon}_k(x) = \upsilon_k([\overline{y}, \underline{\lambda}_x; \underline{y}])$ for some $\overline{\lambda}_x$ and $\underline{\lambda}_x$ in Π_k;

***b*3**: for each $x \in X$, $\overline{\upsilon}_k(x) \geq \overline{\upsilon}_{k+1}(x) \geq \underline{\upsilon}_{k+1}(x) \geq \underline{\upsilon}_k(x)$.

Condition $b0$ fixes the utility values from the upper and lower benchmarks $\overline{y}$ and $\underline{y}$, which corresponds to the implication of $B0$. Then, $b1$ means that for benchmark lotteries $[\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$, $\boldsymbol{\upsilon}_k([\overline{y}, \lambda; \underline{y}])$ is effectively single-valued and takes the expected utility value of $\overline{y}$ and $\underline{y}$, which corresponds to B1. Here, the EU hypothesis is included. $b2$ states that the least upper and greatest lower utilities of $x \in X$ are measured by the benchmark scale $B_k(\overline{y}; \underline{y})$; this does not exactly correspond to B2, but it does an implication of B2 with the help of transitivity for $\boldsymbol{\upsilon}_k$ implied by the interval order $\geq_I$. Corresponding to B3, $b3$ states that $\overline{\upsilon}_k(x)$ and $\underline{\upsilon}_k(x)$ are getting more accurate as k increases. These imply

$$\upsilon_k(\overline{y}) = \upsilon_0(\overline{y}) \text{ and } \upsilon_k(\underline{y}) = \upsilon_0(\underline{y}) \text{ for any } k < \rho + 1. \tag{17}$$

Now, we have Theorem 3.1. As stated, all proofs are given in separate subsections.

Theorem 3.1 (Representation for Step *B*). A base preference stream $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ satisfies Axioms B0 to B3 if and only if there is a base utility stream $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ satisfying $b0$ to $b3$ such that for any $k < \rho + 1$ and $(f, g) \in D_k$,

$$f \trianglerighteq_k g \text{ if and only if } \boldsymbol{\upsilon}_k(f) \geq_I \boldsymbol{\upsilon}_k(g). \tag{18}$$

Although $\boldsymbol{\upsilon}_k$ is vector-valued, it satisfies the EU hypothesis, since by $b1$, $\upsilon_k([\overline{y}, \lambda; \underline{y}]) = \lambda \upsilon_k(\overline{y}) + (1 - \lambda)\upsilon_k(\underline{y})$ for $[\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$, and by $b2$, $\overline{\upsilon}_k(x) = \overline{\lambda}_x \upsilon_k(\overline{y}) + (1 - \overline{\lambda}_x)\upsilon_k(\underline{y})$ and $\underline{\upsilon}_k(x) = \underline{\lambda}_x \upsilon_k(\overline{y}) + (1 - \underline{\lambda}_x)\upsilon_k(\underline{y})$ for $x \in X$. The EU hypothesis does not hold for the representation theorem for Step E (Theorem 5.1).

We have the uniqueness theorem.

Theorem 3.2 (Uniqueness). Let $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ satisfy Axioms B0 to B3. If $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ and $\langle \boldsymbol{\upsilon}'_k \rangle_{k<\rho+1}$ satisfying $b0$ to $b3$ represent $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ in the sense of (18), there are rational numbers $\alpha > 0$ and β such that $\boldsymbol{\upsilon}'_k(x) = \alpha \boldsymbol{\upsilon}_k(x) + \beta = [\alpha \overline{\upsilon}_k(x) + \beta, \alpha \underline{\upsilon}_k(x) + \beta]$ for all $x \in X$ and $k < \rho + 1$.

Conditions $b0$ to $b3$ require $\boldsymbol{\upsilon}_k(x) = (\overline{\upsilon}_k(x), \underline{\upsilon}_k(x))$ be represented essentially by two values λ and λ' in Π_k with $\upsilon_0(\overline{y})$ and $\upsilon_0(\underline{y})$. However, $\upsilon_0(\overline{y})$ and $\upsilon_0(\underline{y})$ for each $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ are allowed

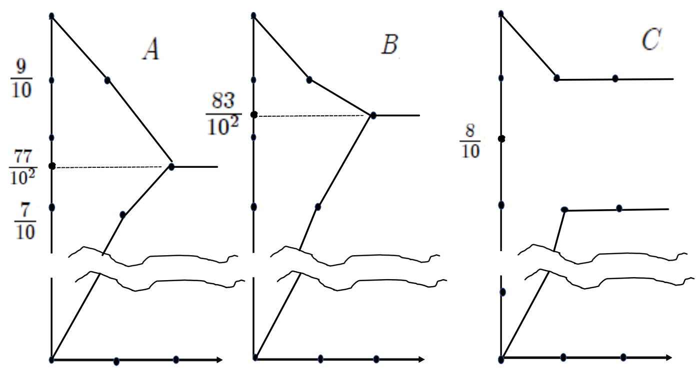

Figure 2: upper and lower utility functions

to take any two rational numbers in $\mathbb{Q}$ only with $v_0(\overline{y}) > v_0(\underline{y})$. This is the reason for the above uniqueness result.

The uniqueness up to a positive linear transformation plays a crucial role in the literature of bargaining theories such as Nash [24] and the Nash welfare function theory (Kaneko- Nakamura [19]). The rational number scalars are enough for the 2-person case and the real-algebraic numbers are enough for the general n-person case (cf. Kaneko [16] for the 2-person case). It is easy to generalize Theorem 3.2 for the real numbers scalars, but the problem is how much we restrict the scalars. Theorem 3.2 is suggestive of how bounded rationality is incorporated to these theories.

The processes described in terms of $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ and/or $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ are thought experiments by the decision maker to search preferences/utilities in his mind. From the viewpoint of "bounded rationality", he may stop his search when he is satisfied and/or is already tired. This is the same as Simon's [30] argument of satisficing/aspiration. First, we consider Example 3.1, and then we exemplify the satisficing/aspiration argument.

Example 3.1. Let $X = \{\overline{y}, y, \underline{y}\}$, $\boldsymbol{v}_0(\overline{y}) = [1,1]$, $\boldsymbol{v}_0(\underline{y}) = [0,0]$, and $\boldsymbol{v}_0(y) = [1,0]$. Also, let $\boldsymbol{v}_1(y) = [\frac{9}{10}, \frac{7}{10}]$. Then, $\boldsymbol{v}_1(f) = [\frac{8}{10}, \frac{8}{10}]$ for $f = [\overline{y}, \frac{8}{10}; \underline{y}]$ by $b1$. Then $\boldsymbol{v}_1(y) \not\geq_I \boldsymbol{v}_1(f)$ and $\boldsymbol{v}_1(f) \not\geq_I \boldsymbol{v}_1(y)$; so y and f are incomparable with respect to $\trianglerighteq_1$ by (18). In Fig.2, $\langle \boldsymbol{v}_k(y) \rangle_{k<\rho+1} = \langle [\overline{v}_k(y), \underline{v}_k(y)] \rangle_{k<\rho+1}$ is described as solid lines in cases A, B, and C. Since $\boldsymbol{v}_0(y) = [1,0]$, we have $\overline{y} \triangleright_0 y \triangleright_0 \underline{y}$ by (18). For $k = 2$, in A, $\boldsymbol{v}_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$ and the decision maker prefers $f = [\overline{y}, \frac{8}{10}; \underline{y}]$ to y; and in B, $\boldsymbol{v}_2(y) = [\frac{83}{10^2}, \frac{83}{10^2}]$; he prefers y to f. In C, $\boldsymbol{v}_k(y) = [\frac{9}{10}, \frac{7}{10}]$ is constant for $k \geq 2$; he gives up comparisons between y and f after $k = 1$.

Introspection Process of Simon's' satisficing/aspiration: The decision maker starts evaluating of a pure alternative $y \in X$ with the benchmark scale $B_0(\overline{y}; \underline{y})$. Suppose that he finds $\boldsymbol{v}_0(y) = [1,0]$, i.e., he attaches the upper value 1 and lower value 0 to y. If $\rho = 0$, his introspection is over. Let $\rho \geq 1$. Then, he goes to layer 1 and uses the more precise scale $B_1(\overline{y}; \underline{y})$ to measure y. In Example 3.1, y is better than $[\overline{y}, \frac{7}{10}; \underline{y}]$ but worse than $[\overline{y}, \frac{9}{10}; \underline{y}]$. Still, he has not reached a very precise measurement. If $\rho = 1$, he stops introspection. Let $\rho \geq 2$. Then, he goes to layer $k = 2$; in A of Fig.2, he reaches precise utility values $\boldsymbol{v}_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$, but in C, he has still imprecise values $\boldsymbol{\upsilon}_2(y) = [\frac{9}{10}, \frac{7}{10}]$ and does not improve them any more even for $k > 2$.

In case A, there are several possible interpretations: one is that $\boldsymbol{\upsilon}_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$ expresses his preferences precisely, and the other is that he is still unsure about the value of y, for example, the lower and upper values may be $\frac{77}{10^2}$ and $\frac{78}{10^2}$, but according his aspiration level, the difference $\frac{78}{10^2} - \frac{77}{10^2} = \frac{1}{10^2}$ is tiny and he does not care about the choice between $\frac{77}{10^2}$ and $\frac{78}{10^2}$. By chance, he chooses $\boldsymbol{\upsilon}_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$. In case C, $\boldsymbol{\upsilon}_2(y) = [\frac{9}{10}, \frac{7}{10}]$ is good enough for him, and he forgets further updating.

Thus, exact values may include some imprecision induced by his aspiration level. This is an attribute of cognitive acts by the decision maker, instead of an attribute of probabilities.

3.3 Proofs

Proof of Theorem 3.1. **(If)**: Suppose that $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ satisfies $b0$ to $b3$ and that (18) holds for $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ and $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$

B0: We have, by $b0$, $b1$ and (17), $\upsilon_0(\overline{y}) \geq \overline{\upsilon}_0(x)$ and $\underline{\upsilon}_0(x) \geq \upsilon_0(\underline{y})$, i.e., $\overline{y} \trianglerighteq_0 x \trianglerighteq_0 \underline{y}$. Thus, B0.

B1: By (18), $b1$, and (17), we have $[\overline{y}, \lambda; \underline{y}] \trianglerighteq_k [\overline{y}, \lambda'; \underline{y}]$ if and only if $\boldsymbol{\upsilon}_k([\overline{y}, \lambda; \underline{y}]) \geq_I \boldsymbol{\upsilon}_k([\overline{y}, \lambda'; \underline{y}])$ if and only if $\lambda \upsilon_0(\overline{y}) + (1-\lambda)\upsilon_0(\underline{y}) \geq \lambda' \upsilon_0(\overline{y}) + (1-\lambda')\upsilon_0(\underline{y})$ if and only if $\lambda \geq \lambda'$. That is, B1.

B2: Let $[\overline{y}, \lambda; \underline{y}] \trianglerighteq_k x$ and $\lambda' > \lambda$. By $b2$ and (18), we have $\lambda' \upsilon_k(\overline{y}) + (1-\lambda')\upsilon_k(\underline{y}) > \lambda \upsilon_k(\overline{y}) + (1-\lambda)\upsilon_k(\underline{y}) \geq \overline{\upsilon}_k(x)$. Thus, by $b1$, $\upsilon_k([\overline{y}, \lambda'; \underline{y}]) = \lambda' \upsilon_k(\overline{y}) + (1-\lambda')\upsilon_k(\underline{y}) > \overline{\upsilon}_k(x)$. By (18), we have $[\overline{y}, \lambda'; \underline{y}] \triangleright_k x$. The other case is symmetric.

B3: Let $f \trianglerighteq_k g$. By (18), we have $\underline{\upsilon}_k(f) \geq \overline{\upsilon}_k(g)$. Let $f = x \in X$ and $g = [\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$. Then, $\upsilon_k(g) = \upsilon_{k+1}(g)$ by $b1$. Then, by $b3$, we have $\underline{\upsilon}_{k+1}(f) \geq \underline{\upsilon}_k(f) \geq \upsilon_k(g) = \upsilon_{k+1}(g)$. By (18), $f \trianglerighteq_{k+1} g$. The case $f \in B_k(\overline{y}; \underline{y})$, $g = x \in X$ is parallel. The case $f = [\overline{y}, \lambda; \underline{y}]$, $g = [\overline{y}, \lambda'; \underline{y}] \in B_k(\overline{y}; \underline{y})$ is similar.

(Only-if): Suppose that $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ satisfying Axioms B0 to B3 is given. We construct a base utility stream $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ satisfying (18). We define $[\overline{\upsilon}_k, \underline{\upsilon}_k]$, $k < \rho + 1$, as follows: for any $f \in B_k(\overline{y}; \underline{y}) \cup X$,

$$
\begin{aligned}
\overline{\upsilon}_k(f) &= \min\{\lambda \in \Pi_k : [\overline{y}, \lambda; \underline{y}] \trianglerighteq_k f\}; \\
\underline{\upsilon}_k(f) &= \max\{\lambda \in \Pi_k : f \trianglerighteq_k [\overline{y}, \lambda; \underline{y}]\}.
\end{aligned}
\tag{19}
$$

It holds that $\overline{\upsilon}_k(f) = \underline{\upsilon}_k(f)$ for $f \in B_k(\overline{y}; \underline{y})$. Consider $f = [\overline{y}, \lambda_f; \underline{y}], g = [\overline{y}, \lambda_g; \underline{y}] \in B_k(\overline{y}; \underline{y})$. Then, $f \trianglerighteq_k g$ if and only if $[\overline{y}, \lambda_f; \underline{y}] \trianglerighteq_k [\overline{y}, \lambda_g; \underline{y}]$ if and only if $\lambda_f \geq \lambda_g$, i.e., $\boldsymbol{\upsilon}_k(f) \geq_I \boldsymbol{\upsilon}_k(g)$ by B1. Let $f \in B_k(\overline{y}; \underline{y})$ and $g = x \in X$. Denote $\upsilon_k(f) = \lambda_f$ and $\overline{\upsilon}_k(x) = \overline{\lambda}_x$. Suppose $f \trianglerighteq_k x$. By (19), $[\overline{y}, \lambda_f; \underline{y}] = f \trianglerighteq_k [\overline{y}, \overline{\lambda}_x; \underline{y}]$. By B1, $\lambda_f \geq \overline{\lambda}_x$, i.e., $\boldsymbol{\upsilon}_k(f) \geq_I \boldsymbol{\upsilon}_k(x)$. The converse is obtained by tracing this back. Thus, $f \trianglerighteq_k x$ if and only if $\boldsymbol{\upsilon}_k(f) \geq_I \boldsymbol{\upsilon}_k(x)$. The case $f = x \in X$, $g \in B_k(\overline{y}; \underline{y})$ is parallel.

By (16) and B0, we have $b0$. By (19), we have $b2$ and $b3$. Consider $b1$. Since $\upsilon_k(\overline{y}) = 1$ and $\upsilon_k(\underline{y}) = 0$, the set $\{\upsilon_k(f) : f \in B_k(\overline{y}; \underline{y})\}$ is the same as Π_k. For any $f = [\overline{y}, \lambda; \underline{y}], g = [\overline{y}, \lambda'; \underline{y}] \in B_k(\overline{y}; \underline{y})$, we have, by (17), $\upsilon_k(f) > \upsilon_k(g)$ if and only if $f \triangleright_k g$ if and only if $\lambda > \lambda'$. Hence, $\upsilon_k(f) = \lambda = \lambda \upsilon_k(\overline{y}) + (1-\lambda)\upsilon_k(\underline{y})$, which is $b1$.$\blacksquare$

Proof of Theorem 3.2. Let $\alpha = (\upsilon_0'(\overline{y}) - \upsilon_0'(\underline{y}))/(\upsilon_0(\overline{y}) - \upsilon_0(\underline{y}))$ and $\beta = (\upsilon_0(\overline{y})\upsilon_0'(\underline{y}) - \upsilon_0'(\overline{y})\upsilon_0(\underline{y}))/(\upsilon_0(\overline{y}) - \upsilon_0(\underline{y}))$. Noting (17), we have $\boldsymbol{\upsilon}_k'(\overline{y}) = \alpha \boldsymbol{\upsilon}_k(\overline{y}) + \beta$ and $\boldsymbol{\upsilon}_k'(\underline{y}) = \alpha \boldsymbol{\upsilon}_k(\underline{y}) + \beta$. For any $[\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$, we have $\boldsymbol{\upsilon}_k'([\overline{y}, \lambda; \underline{y}]) = \lambda \boldsymbol{\upsilon}_k(\overline{y}) + (1-\lambda)\boldsymbol{\upsilon}_k(\underline{y}) = \alpha \boldsymbol{\upsilon}_k([\overline{y}, \lambda; \underline{y}]) + \beta$ by $b1$.

For any $x \in X$, we have $\overline{\lambda}_x$ and $\underline{\lambda}_x$ in Π_k by $b2$ for $\boldsymbol{\upsilon}_k$ such that $\boldsymbol{\upsilon}_k(x) = [\upsilon_k([\overline{y}, \overline{\lambda}_x; \underline{y}]), \upsilon_k([\overline{y}, \underline{\lambda}_x; \underline{y}])]$.

Let $\overline{\lambda}'_x$ and $\underline{\lambda}'_x$ be given by $b2$ for $\boldsymbol{\upsilon}'_k$. Suppose $\overline{\lambda}'_x \neq \overline{\lambda}_x$, say, $\overline{\lambda}_x > \overline{\lambda}'_x$. Then, $\boldsymbol{\upsilon}_k([\overline{y}, \overline{\lambda}_x; \underline{y}]) \geq_I \boldsymbol{\upsilon}_k(x)$, but $\overline{\upsilon}_k(x) = \upsilon_k([\overline{y}, \overline{\lambda}'_x; \underline{y}]) > \upsilon_k([\overline{y}, \overline{\lambda}'_x; \underline{y}])$. Hence, $\boldsymbol{\upsilon}_k([\overline{y}, \overline{\lambda}'_x; \underline{y}]) \not\geq_I \boldsymbol{\upsilon}_k(x)$. However, by definition of $\overline{\lambda}'_x$, we have $\boldsymbol{\upsilon}'_k([\overline{y}, \overline{\lambda}'_x; \underline{y}]) \geq_I \boldsymbol{\upsilon}'_k(x)$. This is impossible since $\boldsymbol{\upsilon}_k$ and $\boldsymbol{\upsilon}'_k$ represent the same $\trianglerighteq_k$. The case $\overline{\lambda}_x < \overline{\lambda}'_x$ is parallel. Thus, $\overline{\lambda}'_x = \overline{\lambda}_x$, and similarly, $\underline{\lambda}'_x = \underline{\lambda}_x$, which imply $\boldsymbol{\upsilon}'_k(x) = [\upsilon'_k([\overline{y}, \overline{\lambda}_x; \underline{y}]), \upsilon'_k([\overline{y}, \underline{\lambda}_x; \underline{y}])] = \alpha[\upsilon_k([\overline{y}, \overline{\lambda}_x; \underline{y}]), \upsilon_k([\overline{y}, \underline{\lambda}_x; \underline{y}])] + \beta = \alpha\boldsymbol{\upsilon}_k(x) + \beta$.■

4 Extension Step: Extended Preference Streams

Step B is an introspective process to find preferences hidden in the mind of the decision maker. On the other hand, Step E is a logical process to extend base preferences found in Step B. This difference may create some logical difficulty in adopting the standard method of representing preferences in terms of a binary relation. We actually show that we can avoid this difficulty. Keeping this remark in mind, we present our axiomatic system for Step E. Throughout this section, let $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ be a given base preference stream satisfying Axioms B0 to B3.

4.1 Extended preference stream

Here, we consider how $\trianglerighteq_k$ is extended to $L_k(X)$ for $k < \rho + 1$. Axiom E0 is to convert base preferences $\trianglerighteq_k$ to $\succsim_k$ for each $k < \rho + 1$, depicted as the vertical arrows in Table 1.1.

Axiom E0 (Extension)(i): For any $(f, g) \in D_0$, $f \trianglerighteq_0 g$ if and only if $f \succsim_0 g$.

(ii): For any k $(1 \leq k < \rho + 1)$ and $(f, g) \in D_k$, if $f \trianglerighteq_k g$, then $f \succsim_k g$.

This is the ultimate source for preferences for Step E. For $k = 0$, the base preferences are only the direct source for $\succsim_0$. For $k \geq 1$, in addition to the base preferences, there is another source from the previous $\succsim_{k-1}$; thus, (ii) has only one direction. However, we will show that as long as the domain D_k is concerned, the converse of (ii) holds for our intended preference stream $\langle \succsim_k \rangle_{k<\rho+1}$.

Consider the connection between layers $k - 1$ and k. For $\widehat{f} = (f_1, ..., f_\ell)$ and $\widehat{g} = (g_1, ..., g_\ell)$, we write $\widehat{f} \succsim_k \widehat{g}$ iff $f_t \succsim_k g_t$ for all $t = 1, ..., \ell$. Recall that a decomposition of $f \in L_k(X)$ is defined by (9). We formulate a derivation of $\succsim_k$ from $\succsim_{k-1}$ as follows: let $1 \leq k < \rho + 1$.

Axiom E1 (Derivation from the previous layer): Let $f \in L_k(X)$, $g \in B_k(\overline{y}; \underline{y})$, and $\widehat{f}$, $\widehat{g}$ their decompositions. If $\widehat{f} \succsim_{k-1} \widehat{g}$ or $\widehat{g} \succsim_{k-1} \widehat{f}$, then $f \succsim_k g$ or $g \succsim_k f$, respectively.

In layer $k-1$, each f_t of $\widehat{f} = (f_1, ..., f_\ell)$ is compared with the corresponding benchmark lottery g_t. These preferences are extended to layer k. In Table 1.1, the horizontal arrows indicate this derivation. When $\rho = 2$, the lottery $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$ in the example of (2) should be evaluated in terms of lotteries in $B_2(\overline{y}; \underline{y})$. Axiom NM2 (Independence) is much stronger in that comparisons jump from one layer to a layer of any depth. In E1, a connection from one layer to the next with equal weights describes the step-by-step extension of preferences by the decision maker.

The preferences derived by the above axioms are extended by transitivity: let $0 \leq k < \rho + 1$.

Axiom E2 (Transitivity): For any $f, g, h \in L_k(X)$, if $f \succsim_k g$ and $g \succsim_k h$, then $f \succsim_k h$.

Here, we regard Axioms E0 to E2 as inference rules, rather than properties to be satisfied by $\succsim_k$. This means that the decision maker constructs $\succsim_0, \succsim_1, ...,$ step by step, using these axioms.

As mentioned above, this view may involve some difficulty; it is logically possible that Axioms E0 to E2 may lead to new unintended preferences. Theorem 4.1 states that this is not the case for the constructed preferences. These also have the following additional conditions:

E0*: for all $k < \rho + 1$ and $(f, g) \in D_k$, $f \trianglerighteq_k g$ if and only if $f \succsim_k g$;

E1*: E1 holds and if the premise of E1 includes strict preferences, so does the conclusion.

Condition E0* states that $\langle \succsim_k \rangle_{k<\rho+1}$ is a faithful extension of $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ in that as long as a pair of lotteries in D_k is concerned, the extended relation $\succsim_k$ has no superfluous preferences. E1* is a strengthening of E1, too. Without these, some preferences would be added in the derivation process of $\succsim_0, \succsim_1, ...,$ and also we would have inconveniences in applications. Note that E2 (transitivity) preserves strict preferences in the same way as E1*.

To prove that our constructive extended stream $\langle \succsim_k \rangle_{k<\rho+1}$ enjoys E0*, E1*, and E2, we first show the following lemma using the EU hypothesis. However, this differs from the intended stream, which enjoys the EU hypothesis only partially. For this lemma, a base utility stream $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$ satisfying (18) given in Theorem 3.1 is used.

Lemma 4.1 (Consistency of E0*, E1*, and E2). There is a stream of binary relations $\langle \succsim_k^* \rangle_{k<\rho+1}$ satisfying Axioms E0*, E1*, and E2. One of such a $\langle \succsim_k^* \rangle_{k<\rho+1}$ is given as follows: for all $k < \rho + 1$, we define $\succsim_k^*$ by

$$f \succsim_k^* g \text{ if and only if } E_f(\underline{\upsilon}_k) \geq E_g(\overline{\upsilon}_k). \tag{20}$$

A point of this lemma is that we have no problem in regarding E0 to E2 as requirements for binary relations $\langle \succsim_k^* \rangle_{k<\rho+1}$, and also, some satisfies E0* and E2*.

Now, we prepare a few concepts for the main theorem, i.e., Theorem 4.1, of this section. Let $\langle \succsim_k \rangle_{k<\rho+1}$ be a stream satisfying E0 to E2. We say that $\langle \succsim_k \rangle_{k<\rho+1}$ is the *smallest* stream iff for any $\langle \succsim_k' \rangle_{k<\rho+1}$ satisfying E0 to E2, and $f, g \in L_k(X)$, $k < \rho + 1$,

$$f \succsim_k g \text{ implies } f \succsim_k' g. \tag{21}$$

Also, the set of preferences over $L_k(X)$ derived from $\succsim_{k-1}$ by E1 is denoted by $(\succsim_{k-1})^{\mathrm{E1}}$, and the set of *transitive closure* of $F \subseteq L_k(X)^2$ is denoted by F^{tr}, i.e., $(f, g) \in F^{tr}$ if and only if there is a finite sequence $f = h_0, h_1, ..., h_m = g$ such that $(h_t, h_{t+1}) \in F$ for $t = 0, ..., m - 1$.

Theorem 4.1 (Smallest extended stream). The sequence $\langle \succsim_k \rangle_{k<\rho+1}$ of the sets generated by the following induction:

$$\succsim_0 \ = \ (\trianglerighteq_0)^{tr}; \text{ and } \succsim_k \ = [(\succsim_{k-1})^{\mathrm{E1}} \cup (\trianglerighteq_k)]^{tr} \text{ for each } k \ (1 \leq k < \rho + 1) \tag{22}$$

is the smallest stream satisfying E0 to E2. Also, $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies E0* and E1*.

The construction starts with $\succsim_0 = (\trianglerighteq_0)^{tr}$, which is well defined since $\trianglerighteq_0$ is a binary relation in D_0. Then, provided that $\succsim_{k-1}$ and $\trianglerighteq_k$ are already given, $\succsim_k$ is defined to be $[(\succsim_{k-1})^{\mathrm{E1}} \cup (\trianglerighteq_k)]^{tr}$. This is a subset of $L_k(X)^2$; thus, it is a binary relation. This preference stream is unique, and is the smallest among the streams satisfying E0 to E2. Furthermore, the constructed stream satisfies E0* and E1*.

We extract the essential addition in (22) to Axioms E0 to E2 and formulate it as Axiom E3. It states that a preference $f \succsim_k g$ is based on comparisons with the benchmark scale $B_k(\overline{y}; \underline{y})$ with either $\trianglerighteq_k$ or $\succsim_{k-1}$. We note that h in E3 may be the same as f or g.

Axiom E3 (i): For any $f, g \in L_0(X)$, if $f \succsim_0 g$, then $f \succsim_0 h \succsim_0 g$ for some $h \in B_0(\overline{y}; \underline{y})$.

(ii): For any $f, g \in L_k(X)$ $(k \geq 1)$, if $f \succsim_k g$, then there is an $h \in B_k(\overline{y}; \underline{y})$ with $f \succsim_k h \succsim_k g$ such that for the first pair (f, h), $f \trianglerighteq_k h$ holds or f, h have decompositions $\widehat{f}, \widehat{h}$ with $\widehat{h} \succsim_{k-1} \widehat{g}$; and the same holds for the second pair (h, g).

The key of Axiom E3 is to include the depth constraint; comparison $f \succsim_k g$ effectively comes from the benchmark scale $B_k(\overline{y}; \underline{y})$ of the same depth k. This constraint with a finite cognitive bound ρ makes the EU hypothesis hold partially. The preference stream given by (20) of Lemma 4.1 enjoys the EU hypothesis, but since it does not take the depths of f, g into account, Axiom E3 is violated. This violation will be seen in Example 5.1.

The stream $\langle \succsim_k \rangle_{k<\rho+1}$ given by (22) is characterized by adding E3 to E0 to E2.

Theorem 4.2 (Uniqueness by E0 to E3). Any extended stream satisfying E0 to E3 is the same as the preference stream $\langle \succsim_k \rangle_{k<\rho+1}$ given by Theorem 4.1.

Throughout the following, the stream given by (22) is denoted by $\langle \succsim_k \rangle_{k<\rho+1}$. Other streams may have some additional superscripts such as $'$, $*$.

Proposition 4.1 will be used in the subsequent analyses: (1) is the horizontal arrows in Table 1.1, and (2) that $\succsim_k$ is bounded in $L_k(X)$ by the upper and lower benchmarks $\overline{y}$ and $\underline{y}$.

Proposition 4.1. Let $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies E0 to E3, and $1 \leq k < \rho + 1$.

(1)(Preservation of preferences): For any $f, g \in L_{k-1}(X)$, $f \succsim_{k-1} g$ implies $f \succsim_k g$.

(2): $\overline{y} \succsim_k f \succsim_k \underline{y}$ for any $f \in L_k(X)$.

We have extended an already built preference relation $\succsim_{k-1}$ and a given base relation $\trianglerighteq_k$ to $\succsim_k$ by Axioms E1 and E2, which is the weakest relation. This extension process is somewhat similar to Dubra-Ok's [6] argument: they extend a preference relation on a finite set of lotteries to the smallest relation satisfying Axiom NM2.ID1, and they show the extended relation is represented by a set of expected utilities. Our extension process is weaker than theirs in that it requires only Axioms E1 and E2 (and E3), and as stated above, it is much weaker than NM2.

Remark 4.1 (Partial EU hypothesis in the two systems). Axiom B1 and $b1$ assume the EU hypothesis along the benchmark scale $B_k(\overline{y}; \underline{y})$, and E1 is a very weak form of Axiom NM2 (independence). The other axioms are related to it only in that preferences are considered through comparisons with $B_k(\overline{y}; \underline{y})$. As mentioned above, Axiom E3 includes the depth constraint on preference comparisons; it follows from Lemma 4.1 and Theorem 4.1 that there are possibly many preference streams satisfying E0 to E2, among which the EU representation in (20) is allowed. A departure from the EU hypothesis is caused by two types of depths included in a lottery and their interactions with a cognitive bound $\rho < \infty$. For example, lottery $d = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$ involves the depths of coefficient $\frac{25}{10^2}$ and of evaluation λ_y. This and E3 make the EU hypothesis hold only for some partial domain, which is explicitly studied in Section 6.

4.2 Proofs

Proof of Lemma 4.1. We show that $\langle \succsim_k^* \rangle_{k<\rho+1}$ given by (20) satisfies E0*, E1*, and E2. By Theorem 3.2, we can assume that $\upsilon_k(\overline{y}) = 1$ and $\upsilon_k(\underline{y}) = 0$.

Since $E_f(\underline{\upsilon}_k) = \lambda$ if $f = [\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$ and $E_x(\overline{\upsilon}_k) = \overline{\upsilon}_k(x)$ if $f = x \in X$. Hence, by (18) and $b2$, $f \trianglerighteq_k x$ if and only if $\lambda \geq \overline{\upsilon}_k(x)$ if and only if $E_f(\underline{\upsilon}_k) \geq E_x(\overline{\upsilon}_k)$. The other cases are symmetric. Thus, E0* holds for any $(f, g) \in D_k$.

It remains to show that $\langle \succsim_k^* \rangle_{k<\rho+1}$ satisfies E1* and E2. Since (20) gives the interval order over the set $\{[E_f(\overline{v}_k), E_f(\underline{v}_k)] : f \in L_k(X)\}$, E2 holds. We show E1*. Let $f \in L_k(X), g \in B_k[\overline{y}; \underline{y}]$ and their decompositions $\widehat{f}$ and $\widehat{g}$ with $\widehat{f} \succsim_{k-1}^* \widehat{g}$. By (20), $E_{f_t}(\underline{v}_k) \geq E_{g_t}(\overline{v}_k)$ for all $t = 1, ..., \ell$. Then, $E_f(\underline{v}_k) = E_{\widehat{e}*\widehat{f}}(\underline{v}_k) = \sum_{t=1}^{\ell} \frac{1}{\ell} E_{f_t}(\underline{v}_k) \geq \sum_{t=1}^{\ell} \frac{1}{\ell} E_{g_t}(\overline{v}_k) = E_{\widehat{e}*\widehat{g}}(\overline{v}_k) = E_g(\overline{v}_k)$. If strict preferences are included in the decompositions, the conclusion is strict; thus we have E1*.■

Proof of Theorem 4.1. This has the three assertions: (a) $\langle \succsim_k \rangle_{k<\rho+1}$ is a sequence of a binary relations satisfying Axioms E0 to E2; (b) it is the smallest in the sense of (21) among streams $\langle \succsim_k' \rangle_{k<\rho+1}$ satisfying E0 to E2; and (c) E0*, E1* hold for $\langle \succsim_k \rangle_{k<\rho+1}$.

(a): E2 follow directly from (22). Consider E0. (ii) follows from (22). We show that $\succsim_0 = (\trianglerighteq_0)^{tr}$ satisfies that for any $(f, g) \in D_0$, $f \succsim_0 g$ implies $f \trianglerighteq_0 g$. Since $\succsim_0 = (\trianglerighteq_0)^{tr}$, there is a sequence $f = h_0 \trianglerighteq_0 ... \trianglerighteq_0 h_m = g$. If $h_t \in X - B_0(\overline{y}; \underline{y})$, then $h_{t-1} \in B_0(\overline{y}; \underline{y})$ and $h_{t+1} \in B_0(\overline{y}; \underline{y})$. By B2, $\lambda_{t-1} \geq \lambda_{t+1}$, where $h_{t-1} = [\overline{y}; \lambda_{t-1}, \underline{y}]$ and $h_{t+1} = [\overline{y}; \lambda_{t+1}, \underline{y}]$. If $h_t, h_{t+1} \in B_0(\overline{y}; \underline{y})$, then $\lambda_t \geq \lambda_{t+1}$. Hence, we can shorten the sequence to $f = h_0 \trianglerighteq_0 h_m = g$. Thus, $f \trianglerighteq_0 g$.

Consider E1. Suppose that $f \in B_k[\overline{y}; \underline{y}]$ and $g \in L_k(X)$ have decompositions $\widehat{f}, \widehat{g} \in L_{k-1}(X)$ with $\widehat{f} \succsim_{k-1} \widehat{g}$. By (22), we have $f = e * \widehat{f} \succsim_k e * \widehat{g} = g$. The symmetric case, $\widehat{g} \succsim_{k-1} \widehat{f}$, is similar.

(b): We prove by induction on k that $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies (21) for any $\langle \succsim_k' \rangle_{k<\rho+1}$ satisfying E0 to E2. When $k = 0$, we have $\succsim_0 = (\trianglerighteq_0)^{tr}$ by (22). Let $f \succsim_0 g$, i.e., $f \ (\trianglerighteq_0)^{tr} g$, which implies that there is a sequence $f = h_0 \trianglerighteq_0 h_1 \trianglerighteq_0 ... \trianglerighteq_0 h_m = g$. By E0.(i), we have $f = h_0 \succsim_0' h_1 \succsim_0' ... \succsim_0' h_m = g$. By E2 for $\succsim_0'$, we have $f \succsim_0' g$.

Now, we assume that (21) holds for $k-1$. Let $f \succsim_k g$. By (22), there is a sequence $f = h_0 \succsim_k ... \succsim_k h_m = g$ such that each $h_t \succsim_k h_{t+1}$ is a consequence of E1 or $h_t \succsim_k h_{t+1}$ is $h_t \trianglerighteq_k h_{t+1}$. In the first case, there are decompositions $\widehat{h}_t, \widehat{h}_{t+1}$ of h_t, h_{t+1} such that $\widehat{h}_t \succsim_{k-1} \widehat{h}_{t+1}$. By the induction hypothesis, we have $\widehat{h}_t \succsim_{k-1}' \widehat{h}_{t+1}$. Thus, $h_t \succsim_k' h_{t+1}$ by E1 for $\succsim_k'$. In the second case, $h_t \trianglerighteq_k h_{t+1}$ implies $h_t \succsim_k' h_{t+1}$ by E0.(ii) for $\succsim_k'$. Hence, $f \succsim_k' g$ by E2 for $\succsim_k'$.

(c): Take $\langle \succsim_k^* \rangle_{k<\rho+1}$ given by Lemma 4.1. Since $\langle \succsim_k^* \rangle_{k<\rho+1}$ satisfies E0 to E2, it holds that for all $k < \rho + 1$ and $f, g \in L_k(X)$,

$$f \succsim_k g \text{ implies } f \succsim_k^* g. \tag{23}$$

E0*: Since E0* holds for $\succsim_k^*$ by Lemma 4.1, we have: for any $(f, g) \in D_k$, $f \succsim_k^* g$ implies $f \trianglerighteq_k g$; thus, $f \succsim_k g$ implies $f \trianglerighteq_k g$. The converse is from (22).

E1*: Let $f, g \in L_k(X)$. Let $\widehat{f}, \widehat{g}$ be decompositions of f, g so that $\widehat{f} \succsim_{k-1} \widehat{g}$ with strict preferences for some components. Hence, by (23), the same holds for $\succsim_{k-1}^*$. Hence, by E2 for $\succsim_{k-1}^*$, we have $f \succ_k^* g$. In this case, $g \succsim_k f$ is impossible; if it was the case, we would have, by (23), $f \sim_k^* g$, a contradiction. Hence, $f \succ_k g$.■

Proof of Theorem 4.2. Let $\langle \succsim_k^* \rangle_{k<\rho+1}$ be any extended stream satisfying E0 to E3. We prove by induction on $k < \rho + 1$ that for any $f, g \in L_k(X)$,

$$f \succsim_k^* g \text{ if and only if } f \succsim_k g. \tag{24}$$

Since $\langle \succsim_k \rangle_{k<\rho+1}$ is the smallest stream satisfying E0 to E2 by Theorem 4.1, the *if* part holds for any $k < \rho + 1$. Consider the *only-if* part. Let $k = 0$. Let $f, g \in L_0(X)$ with $f \succsim_0^* g$. Then, by E3.(i), we have an $h \in B_0(\overline{y}; \underline{y})$ with $f \succsim_0^* h \succsim_0^* g$. But this h is either $\overline{y}$ or $\underline{y}$. If $h = \overline{y}$, then $h \trianglerighteq_0 f$ and $f \trianglerighteq_0 h$ by B0. Since $h \trianglerighteq_0 g$ by B0, we have $f \ (\trianglerighteq_0)^{tr} g$, i.e., $f \succsim_0 g$ by (22). The case $h = \underline{y}$ is similar.

We make the induction hypothesis that the *only-if* part holds for $k-1$. Let $f \succsim_k^* g$. Then, by E3, we have $h_0 := f \succsim_k^* h_1 \succsim_k^* h_2 := g$ for some $h_1 \in B_k(\overline{y}; \underline{y})$. If $h_0 = x \in X$, then, $h_0 = x$ has no decomposition; thus, by E3.(ii), $h_0 \trianglerighteq_k h_1$, which implies $h_0 \succsim_k h_1$ by E0* for $\succsim_k$. Let $h_0 \notin X$. By E3.(ii), $\widehat{h}_0 \succsim_{k-1}^* \widehat{h}_1$ for some decompositions $\widehat{h}_0$, $\widehat{h}_1$ of h_0, h_1. By the induction hypothesis, we have $\widehat{h}_0 \succsim_{k-1} \widehat{h}_1$. Thus, by E1 for $\succsim_k$, we have $h_0 \succsim_k h_1$. By the same argument, we have $h_1 \succsim_k h_2$. Thus, by (22), $h_0 \succsim_k h_2$, i.e., $f \succsim_k g$.■

Proof of Proposition 4.1. **(1)**: Let $f \in L_{k-1}(X)$ and $g \in B_{k-1}(\overline{y}; \underline{y})$. Suppose $f \succsim_{k-1} g$. Then, $f, g \in L_{k-1}(X) \subseteq L_k(X)$. Let $f_1 = ... = f_\ell = f$ and $g_1 = ... = g_\ell = g$. Then, $f = \sum_{t=1}^{\ell} \frac{1}{\ell} * f_t$ and $f = \sum_{t=1}^{\ell} \frac{1}{\ell} * g_t$. By E1, we have $f \succsim_k g$. The case $g \succsim_{k-1} f$ is similar.

Let $f, g \in L_{k-1}(X)$ with $f \succsim_{k-1} g$. Then, by E3.(ii) for k, $f \succsim_{k-1} h \succsim_{k-1} g$ for some $h \in B_{k-1}(\overline{y}; \underline{y})$. It follows from the conclusion of the above paragraph that $f \succsim_k h \succsim_k g$. By E2, we have $f \succsim_k g$.

(2): Let $f \in L_0(X) = X$. By B0 and $\succsim_0 = \trianglerighteq_0$, we have the assertion for $k = 0$. Suppose the induction hypothesis that $\overline{y} \succsim_{k-1} f \succsim_{k-1} \underline{y}$ for any $f \in L_{k-1}(X)$. Consider $f \in L_k(X)$. Then, by Lemma 2.1, there is a vector $\widehat{f} \in L_{k-1}(X)^\ell$ such that $f = \widehat{e} * \widehat{f}$. By the induction hypothesis, $\overline{y} \succsim_{k-1} f_t \succsim_{k-1} \underline{y}$ for any $t \leq \ell$. By E1, $\overline{y} = \widehat{e} * \overline{y} \succsim_k f = \widehat{e} * \widehat{f} \succsim_k \widehat{e} * \underline{y} = \underline{y}$.■

5 Extension Step: Vector-valued Utility Stream

We extend a base utility stream $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ to $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ so that each $\boldsymbol{u}_k$ is a function over $L_k(X)$. We show that this approach is equivalent to that given in Section 4. It provides clear-cut interpretations and mathematical tractability of the entire theory.

5.1 Extended utility stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$

Let a base utility stream $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ satisfying conditions $b0$ to $b3$ be given. We consider a stream of functions $\langle \boldsymbol{u}_k \rangle_{k<\rho+1} = \langle [\overline{u}_k, \underline{u}_k] \rangle_{k<\rho+1}$ so that each $\boldsymbol{u}_k = [\overline{u}_k, \underline{u}_k]$ is a function from $L_k(X)$ to $\mathbb{Q}^2$ with $\overline{u}_k(f) \geq \underline{u}_k(f)$ for all $f \in L_k(X)$. As for a base utility stream, the values $\overline{u}_k(f)$ and $\underline{u}_k(f)$ are interpreted as the least upper and greatest lower bounds of possible utilities from f. For $\widehat{f} = (f_1, ..., f_\ell)$, we write $\overline{u}_k(\widehat{f}) = (\overline{u}_k(f_1), ..., \overline{u}_k(f_\ell))$ and $\underline{u}_k(\widehat{f}) = (\underline{u}_k(f_1), ..., \underline{u}_k(f_\ell))$. Also, recall that when $\boldsymbol{u}_k = [\overline{u}_k, \underline{u}_k]$ is effectively single-valued for f, we drop the upper and lower bars from $\overline{u}_k(f), \underline{u}_k(f)$, and write it as $u_k(f)$.

We assume the following four conditions on $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$: for each $k < \rho + 1$,

***e*0**: The restriction of $\boldsymbol{u}_k$ to $B_k(\overline{y}; \underline{y}) \cup X$ coincides with $\boldsymbol{v}_k$.

***e*1**: Let $f \in L_k(X)$, $g \in B_k(\overline{y}; \underline{y})$, and $\widehat{f}$, $\widehat{g}$ be their decompositions. If $u_{k-1}(\widehat{g}) \geq \overline{u}_{k-1}(\widehat{f})$ or $\underline{u}_{k-1}(\widehat{f}) \geq u_{k-1}(\widehat{g})$, then $u_k(g) \geq \overline{u}_k(f)$ or $\underline{u}_k(f) \geq u_k(g)$, respectively.

***e*2**: For any $f \in L_k(X)$ with $\delta(f) \geq 1$, there are decompositions $\widehat{f}, \widehat{f'} \in L_{k-1}(X)^\ell$ such that $\overline{u}_k(f) = \widehat{e} * \overline{u}_{k-1}(\widehat{f})$ and $\underline{u}_k(f) = \widehat{e} * \underline{u}_{k-1}(\widehat{f'})$.

***e*3**: For any $f \in L_k(X)$, there are g and h in $B_k(\overline{y}; \underline{y})$ such that $\overline{u}_k(f) = u_k(g)$ and $\underline{u}_k(f) = u_k(h)$.

Conditions $e0$ and $e1$ correspond to E0 and E1, while $e2$ does not to E2, since transitivity is already included in the interval order $\geq_I$. Condition $e3$ requires that the least upper and greatest lower utilities $\overline{u}_k(f)$ and $\underline{u}_k(f)$ come from those of some decompositions. Condition $e3$

is a depth constraint corresponding to Axiom $E3$: it requires $\overline{u}_k(f)$ and $\underline{u}_k(f)$ be in $B_k(\overline{y};\underline{y})$ of the same same layer.

First, we present the uniqueness of a possible utility stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ extended by $e0$ to $e3$.

Proposition 5.1 (Unique extension). If $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ and $\langle \boldsymbol{u}'_k \rangle_{k<\rho+1}$ are extended streams from a base stream $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$, then $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ and $\langle \boldsymbol{u}'_k \rangle_{k<\rho+1}$ are the same.

Here, the extended utility stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ is uniquely determined relative to a given base preference stream $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ representing $\langle \trianglerighteq_k \rangle_{k<\rho+1}$. Hence, it follows from Theorem 3.2 that the stream of pair $\langle \boldsymbol{v}_k, \boldsymbol{u}_k \rangle_{k<\rho+1}$ is unique up to positive linear transformation.

The existence of $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ is guaranteed by the next theorem and Theorems 4.2. Recall that $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ with B0 to B3 are assumed behind E0 and that $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ with $b0$ to $b3$ are assumed behind $e0$. They are are connected by (18).

Theorem 5.1 (Representation of $\langle \succsim_k \rangle_{k<\rho+1}$ by $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$). A preference stream $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies E0 to E3 if and only if there is a utility stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ satisfying $e0$ to $e3$ such that for any $k < \rho + 1$ and $f, g \in L_k(X)$,

$$f \succsim_k g \text{ if and only if } \boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g). \tag{25}$$

Theorem 5.1 can regarded as a substantiation of the indication, by von Neumann-Morgenstern [33], p.29, of a possibility of a representation of a preference relation involving incomparabilities in terms of a higher-dimensional vector-valued function.

The structure of the arguments given in Sections 3 to 5 is summarized in Table 1.2. We started with the theory of a base preference stream $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ and of a base utility stream $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$. In Section 4, $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ is extended to $\langle \succsim_k \rangle_{k<\rho+1}$; Theorems 4.1 and 4.2 show the unique existence of $\langle \succsim_k \rangle_{k<\rho+1}$ satisfying E0 to E3, relative to $\langle \trianglerighteq_k \rangle_{k<\rho+1}$. Proposition 5.1 implies that this $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ is uniquely determined relative to $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$. Theorem 5.1 implies the existence of an extended utility stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$. Instead of this way of proving the existence of $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$, we can prove it directly from a given $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$. However, the present way through Theorem 4.1 to the introduction E3 explains better our motivation of the construction of extended preference stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$.

The EU representation given in Lemma 4.1 differs from that in Theorem 5.1 since it violates condition $e3$. In this sense, Theorem 5.1 is different from Theorem 3.1. This creates some difficulty in practical calculation of $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ for the case $\rho < \infty$; for practical purpose, it may be useful to observe the following: for for any $f \in L_k(X)$ with $\delta(f) > 0$,

$$\begin{aligned} \overline{u}_k(f) &= \min\{\widehat{e} * \overline{u}_{k-1}(\widehat{f}) : \widehat{f} \text{ is a decomposition of } f\}; \\ \underline{u}_k(f) &= \max\{\widehat{e} * \underline{u}_{k-1}(\widehat{f}) : \widehat{f} \text{ is a decomposition of } f\}. \end{aligned} \tag{26}$$

This can be proved by using conditions $e2$ and $e3$. Using this, we calculate $\overline{u}_k$ and $\underline{u}_k$ in an example, which will be used in Section 8. In fact, the general existence result of $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ mentioned above is based on this observation.

Example 5.1. Consider A and B of Example 3.1 with $X = \{\overline{y}, y, \underline{y}\}$. Recall $\boldsymbol{v}_0(\overline{y}) = \boldsymbol{v}_1(\overline{y}) = [1, 1]$, $\boldsymbol{v}_0(\underline{y}) = \boldsymbol{v}_1(\underline{y}) = [0, 0]$, $\boldsymbol{v}_0(y) = [1, 0]$, $\boldsymbol{v}_1(y) = [\frac{9}{10}, \frac{7}{10}]$. These values are the same as $\boldsymbol{u}_0, \boldsymbol{u}_1$. Keeping $\boldsymbol{v}_0$ and $\boldsymbol{v}_1$ in mind, consider the two cases:

$$A : \boldsymbol{v}_2(y) = \boldsymbol{u}_2(y) = [\tfrac{77}{10^2}, \tfrac{77}{10^2}] \text{ and } B : \ \boldsymbol{v}'_2(y) = \boldsymbol{u}'_2(y) = [\tfrac{83}{10^2}, \tfrac{83}{10^2}]. \tag{27}$$

In these cases, we consider how to calculate $\boldsymbol{u}_2(d)$ and $\boldsymbol{u}_3(d)$ for $d = \frac{25}{10^2}y * \frac{75}{10^2}\underline{y}$. We mainly consider case A, and will mention how to adjust the calculation in case B.

The lottery d has the three types of decompositions $\frac{t}{10} * y + \frac{5-2t}{10} * [y, \frac{5}{10}; \underline{y}] + \frac{5+t}{10} * \underline{y}$ for $t = 0, 1, 2$ indicated in (12). Among these, the one with $t = 2$ gives $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$, and the others give wider intervals; since y is evaluated in a shallower layer through $\frac{5-2t}{10} * [y, \frac{5}{10}; \underline{y}]$ than $\frac{t}{10} * y$, it would be more accurate to use y than $[y, \frac{5}{10}; \underline{y}]$. Thus, we take the largest weight, $t = 2$, to y. By this remark, the min and max operators in (26) become unnecessary.

First, since $\frac{5}{10}y * \frac{5}{10}\underline{y}$ itself is regarded as a unique decomposition, we have, by $e1$,

$$\begin{aligned} \boldsymbol{u}_1(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) &= \tfrac{5}{10}\boldsymbol{v}_0(y) + \tfrac{5}{10}\boldsymbol{v}_0(\underline{y}) = [\tfrac{5}{10}, 0]; \\ \boldsymbol{u}_2(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) &= \tfrac{5}{10}\boldsymbol{v}_1(y) + \tfrac{5}{10}\boldsymbol{v}_1(\underline{y}) = [\tfrac{45}{10^2}, \tfrac{35}{10^2}]. \end{aligned} \tag{28}$$

Plugging these to the decomposition $\frac{2}{10}y * \frac{1}{10}(\frac{5}{10}y * \frac{5}{10}\underline{y}) * \frac{7}{10}\underline{y}$, we have, by $e1$,

$$\begin{aligned} \boldsymbol{u}_2(d) &= \boldsymbol{u}_2(\tfrac{2}{10}y * \tfrac{1}{10}(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) * \tfrac{7}{10}\underline{y}) = \tfrac{2}{10}\boldsymbol{u}_1(y) + \tfrac{1}{10}\boldsymbol{u}_1(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) + \tfrac{7}{10}\boldsymbol{u}_1(\underline{y}) \\ &= \tfrac{2}{10}[\tfrac{9}{10}, \tfrac{7}{10}] + \tfrac{1}{10}[\tfrac{5}{10}, 0] + \tfrac{1}{10}[0, 0] = [\tfrac{23}{10^2}, \tfrac{14}{10^2}]. \end{aligned} \tag{29}$$

This is compared with $\boldsymbol{u}_2(c) = \boldsymbol{u}_2(\frac{2}{10}\overline{y} * \frac{8}{10}\underline{y}) = [\frac{2}{10}, \frac{2}{10}]$, and these imply that c and d are incomparable with respect to $\precsim_2$. This incomparability will be interpreted in the experimental environment in Section 8.

Incidentally, for the EU relation $\succsim_2^*$ defined by (20) of Lemma 4.1, it holds that $c \succ_2^* d$ for case A, since $E_d(\boldsymbol{v}_2) = [\frac{1925}{10^4}, \frac{1925}{10^4}]$ and $E_c(\upsilon_2) = [\frac{2}{10}, \frac{2}{10}]$. This comparability is possible since (20) ignores the depth constraint; Axiom E3 and $b3$ are violated.

Based on the above results, we can calculate $\boldsymbol{u}_3(d)$ for case $\rho = 3$:

$$\begin{aligned} \boldsymbol{u}_3(d) &= \boldsymbol{u}_3(\tfrac{2}{10}y * \tfrac{1}{10}(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) * \tfrac{7}{10}\underline{y}) = \tfrac{2}{10}\boldsymbol{u}_2(y) + \tfrac{1}{10}\boldsymbol{u}_2(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) + \tfrac{7}{10}\boldsymbol{u}_2(\underline{y}) \\ &= \tfrac{2}{10}[\tfrac{77}{10^2}, \tfrac{77}{10^2}] + \tfrac{1}{10}[\tfrac{45}{10^2}, \tfrac{35}{10^2}] + \tfrac{7}{10}[0, 0] = [\tfrac{199}{10^3}, \tfrac{189}{10^3}]. \end{aligned} \tag{30}$$

In this case, c is strictly prefers to d, since $\boldsymbol{u}_3(c) = \boldsymbol{u}_3(\frac{2}{10}\overline{y} * \frac{8}{10}\underline{y}) = [\frac{2}{10}, \frac{2}{10}]$, though d is not yet measurable. Incidentally, d becomes measurable for $k \geq 4$, since $\boldsymbol{u}_k(d) = [\frac{1925}{10^4}, \frac{1925}{10^4}]$, and c is strictly preferred to d.

Consider case B : $\boldsymbol{u}_2'(y) = [\frac{83}{10^2}, \frac{83}{10^2}]$ for $\rho = 2$. Then, the above calculation (29) for $\boldsymbol{u}_2(d)$ remains the same for $\boldsymbol{u}_2'(d)$ with $\boldsymbol{u}_2'(d) = [\frac{23}{10^2}, \frac{14}{10^2}]$, but for $\rho = 3$, $\boldsymbol{u}_3'(d)$ is calculated as follows:

$$\begin{aligned} \boldsymbol{u}_3'(d) &= \boldsymbol{u}_3'(\tfrac{2}{10}y * \tfrac{1}{10}(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) * \tfrac{7}{10}\underline{y}) = \tfrac{2}{10}\boldsymbol{u}_2(y) + \tfrac{1}{10}\boldsymbol{u}_2(\tfrac{5}{10}y * \tfrac{5}{10}\underline{y}) + \tfrac{7}{10}\boldsymbol{u}_2(\underline{y}) \\ &= \tfrac{2}{10}[\tfrac{83}{10^2}, \tfrac{83}{10^2}] + \tfrac{1}{10}[\tfrac{45}{10^2}, \tfrac{35}{10^2}] + \tfrac{7}{10}[0, 0] = [\tfrac{211}{10^3}, \tfrac{201}{10^3}] \end{aligned} \tag{31}$$

Here, d is strictly preferred to c. For $k \geq 4$, $\boldsymbol{u}_k'(d) = [\frac{2075}{10^4}, \frac{2075}{10^4}]$; d is also strictly preferred to c.

5.2 Proofs

First, we show that the condition corresponding to Axiom E3 holds for $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$.

Lemma 5.1. Let $k < \rho+1$. For any $f, g \in L_k(X)$ with $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$, there is an $h \in B_k(\overline{y}; \underline{y})$ such that $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(h) \geq_I \boldsymbol{u}_k(g)$, and $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(h)$ or f, h have decompositions $\widehat{f}, \widehat{h}$ with $\boldsymbol{u}_{k-1}(\widehat{f}) \geq_I \boldsymbol{u}_{k-1}(\widehat{h})$, and the same holds for h, g.[7]

[7]When f or g belongs to $B_k(\overline{y}; \underline{y})$, h can be f or g, respectively.

Proof. Let $f, g \in L_k(X)$ with $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$, i.e., $\underline{u}_k(f) \geq \overline{u}_k(g)$. If $(f,g) \in D_k$, then, $\underline{u}_k(f) \geq \overline{u}_k(g)$ is equivalent to $\underline{v}_k(f) \geq \overline{v}_k(g)$ by $e0$. We consider the case $(f,g) \in L_k(X) - D_k$.

By $e3$, there are $\underline{\lambda}_f$ and $\overline{\lambda}_g$ in such Π_k that $\underline{u}_k(f) = \underline{\lambda}_f$ and $\overline{u}_k(g) = \overline{\lambda}_g$. Then, $\underline{\lambda}_f \geq \overline{\lambda}_g$. Now, let $h = [\overline{y}, \underline{\lambda}_f; \underline{y}]$. Then

$$\underline{u}_k(f) = u_k(h) = \underline{\lambda}_f \geq \overline{\lambda}_g = \overline{u}_k(g). \tag{32}$$

Thus, it remains to show that (f,h) and (h,g) have decompositions $(\widehat{f}, \widehat{h})$ and $(\widehat{h}', \widehat{g})$ such that $\underline{u}_{k-1}(\widehat{f}) = u_{k-1}(\widehat{h})$ and $u_{k-1}(\widehat{h}^*) \geq u_{k-1}(\widehat{g})$. Note that the second holds in inequality.

By $e2$, there is a decomposition of $\widehat{f}$ of f such that $\widehat{e} * \underline{u}_{k-1}(\widehat{f}) = \underline{u}_k(f)$. By $e3$, there are $\underline{\lambda}_{f_1}, ..., \underline{\lambda}_{f_\ell}$ in Π_k such that $\underline{u}_{k-1}(f_t) = \underline{\lambda}_{f_t}$ for $t \leq \ell$. Let $h_t = [\overline{y}, \underline{\lambda}_{f_t}; \underline{y}]$ for $t \leq \ell$ and let $\widehat{h} = (h_1, ..., h_\ell)$. Since $\widehat{f} \in L_{k-1}(X)^{\ell-1}$, $\widehat{h} = (h_1, ..., h_\ell)$ belongs to $L_{k-1}(X)^{\ell-1}$. Since $\underline{\lambda}_f = \underline{u}_k(f) = \widehat{e} * \underline{u}_{k-1}(\widehat{f}) = \sum \frac{1}{\ell} \cdot \underline{\lambda}_{f_t}$, we have $h = [\overline{y}, \underline{\lambda}_f; \underline{y}] = \sum_t \frac{1}{\ell} * h_t$. Hence, $\widehat{h}$ is a decomposition of h with $\underline{u}_{k-1}(\widehat{f}) = u_{k-1}(\widehat{h})$. This implies $\boldsymbol{u}_{k-1}(\widehat{f}) \geq_I \boldsymbol{u}_{k-1}(\widehat{h})$.

In the same manner, we can show that (h,g) has decompositions $(\widehat{h}', \widehat{g})$ such that $u_{k-1}(\widehat{h}') = \overline{u}_{k-1}(\widehat{g})$. Since $\sum_t \frac{1}{\ell} \cdot \underline{\lambda}_{f_t} = \underline{\lambda}_f \geq \overline{\lambda}_g = \sum_t \frac{1}{\ell} \cdot \overline{\lambda}_{g_t}$, we find $(\overline{\lambda}^*_{g_1}, ..., \overline{\lambda}^*_{g_\ell})$ so that $\sum_t \overline{\lambda}^*_{g_t} = \sum_t \underline{\lambda}_{f_t}$ and $\overline{\lambda}^*_{g_t} \geq \overline{\lambda}_{g_t}$ for all $t \leq \ell$. Then, $\widehat{h}^* = ([y, \overline{\lambda}^*_{g_t}; y] : t \leq \ell)$ is a decomposition of h and $u_{k-1}(\widehat{h}^*) \geq \overline{u}_{k-1}(\widehat{g})$. Thus, $\boldsymbol{u}_{k-1}(\widehat{h}) \geq_I \boldsymbol{u}_{k-1}(\widehat{g})$.■

Proof of Proposition 5.1. We prove by induction that $\boldsymbol{u}_k = [\overline{u}_k, \underline{u}_k] = [\overline{v}_k, \underline{v}_k] = \boldsymbol{v}_k$ for all $k < \rho + 1$. For all $k < \rho + 1$, by $e0$, $\boldsymbol{u}_k(f) = \boldsymbol{v}_k(f) = \boldsymbol{v}_k(f)$ for all $f \in B_k(\overline{y}; \underline{y}) \cup X$. The case $k = 0$ is the induction base.

Suppose that $\boldsymbol{u}_{k-1} = \boldsymbol{v}_{k-1}$. First, we show that for any $f, g \in L_k(X)$, $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g) \Longleftrightarrow \boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(g)$. Let $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$. By Lemma 5.1, there is an $h \in B_k(\overline{y}; \underline{y})$ with $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(h) \geq_I \boldsymbol{u}_k(g)$ such that $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(h)$ or f, h have decompositions $\widehat{f}, \widehat{h}$ with $\boldsymbol{u}_{k-1}(\widehat{f}) \geq_I \boldsymbol{u}_{k-1}(\widehat{h})$, and the same holds for h, g. By the above remark, $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(h)$ implies $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(h)$. By the induction hypothesis, we have $\boldsymbol{v}_{k-1}(\widehat{f}) \geq_I \boldsymbol{v}_{k-1}(\widehat{h})$. By $e1$, we have $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(h)$. Similarly, we have $\boldsymbol{v}_k(h) \geq_I \boldsymbol{v}_k(g)$. By transitivity of $\geq_I$, we have $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(g)$. The converse can be proved in the symmetric manner. Thus, $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$ if and only if $\boldsymbol{v}_k(f) \geq_I \boldsymbol{v}_k(g)$.

It remains to show that $\overline{u}_k(f) = \overline{v}_k(f)$ and $\underline{u}_k(f) = \underline{v}_k(f)$ for all $f \in L_k(X)$. By $e3$, there are $\lambda, \lambda' \in \Pi_k$ such that $\overline{u}_k(f) = v_k([y, \lambda; y])$ and $\overline{v}_k(f) = v_k([y, \lambda'; y])$. If $\lambda \neq \lambda'$, say $\lambda > \lambda'$, then, by $b1$ and $e0$, $\overline{u}_k(f) = v_k([y, \lambda; y]) > v_k([y, \lambda'; y]) = u_k([y, \lambda'; y])$. By the result of the above paragraph, this implies $\overline{v}_k(f) > v_k([y, \lambda'; y])$, which contradicts $\overline{v}_k(f) = v_k([y, \lambda'; y]) = \overline{v}_k([y, \lambda'; y])$. Hence, $\lambda = \lambda'$, i.e., $\overline{u}_k(f) = \overline{v}_k(f)$. The proof of $\underline{u}_k(f) = \underline{v}_k(f)$ is symmetric.■

Proof of Theorem 5.1. (***If***): Let $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ be the extended utility stream satisfying $e0$ to $e3$. Let $\succsim_k$ be the binary relation over $L_k(X)$ defined by (25) and its restriction to D_k by $\trianglerighteq_k$. Since $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ is the restriction of $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ to $\langle B_k(\overline{y}; \underline{y}) \cup X \rangle_{k<\rho+1}$, $\langle \boldsymbol{v}_k \rangle_{k<\rho+1}$ satisfies $b0$ to $b3$ by Theorem 3.1. Thus, we have E0. It remains to show that $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies E1 to E3. Since the relation $\geq_I$ over $\{\boldsymbol{u}_k(f) : f \in L_k(X)\}$ is transitive, E2 is satisfied.

Consider E1. Let $\widehat{f} \in L_{k-1}(X)^\ell$ and $\widehat{g} \in B_{k-1}(\overline{y}; \underline{y})^\ell$ be decompositions of $f \in L_k(X)$ and $g \in B_k(\overline{y}; \underline{y})$. Suppose that $\widehat{f} \succsim_{k-1} \widehat{g}$, which implies $\boldsymbol{u}_{k-1}(\widehat{f}) \geq_I \boldsymbol{u}_{k-1}(\widehat{g})$ by (25). Thus, $\widehat{e} * \underline{u}_{k-1}(\widehat{f}) \geq \widehat{e} * u_{k-1}(\widehat{g}) = u_{k-1}(\widehat{e} * \widehat{g}) = u_k(g)$. By $e1$, we have $\underline{u}_k(f) \geq \widehat{e} * u_{k-1}(\widehat{f})$. Hence, $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$: thus, $f \succsim_k g$ by (25). The other case of $\widehat{g} \succsim_{k-1} \widehat{f}$ is symmetric.

Axiom E3 follows from Lemma 5.1.

(***Only-if***): Suppose that $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies Axioms E0 to E3 with its base stream $\langle \trianglerighteq_k \rangle_{k<\rho+1}$.

Then, $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ satisfies B0 to B3 by E0. We define $\langle \boldsymbol{u}_k \rangle_{k<\rho+1} = \langle [\overline{u}_k, \underline{u}_k] \rangle_{k<\rho+1}$ as follows: for each $f \in L_k(X)$,

$$\begin{aligned} \overline{u}_k(f) &= \min\{\lambda \in \Pi_k : [\overline{y}, \lambda; \underline{y}] \succsim_k f\}; \\ \underline{u}_k(f) &= \max\{\lambda \in \Pi_k : f \succsim_k [\overline{y}, \lambda; \underline{y}]\}. \end{aligned} \tag{33}$$

Because of Lemma 4.1.(2), these functions are well defined. Let $f, g \in L_k(X)$. Let $\underline{u}_k(f) = \underline{\lambda}_f$ and $\overline{u}_k(g) = \overline{\lambda}_g$. Then, $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$ if and only if $\underline{\lambda}_f \geq \overline{\lambda}_g$. This implies that $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$ if and only if $f \succsim_k g$. Indeed, if $\boldsymbol{u}_k(f) \geq_I \boldsymbol{u}_k(g)$, then $f \succsim_k [\overline{y}, \underline{\lambda}_f; \underline{y}] \succsim_k [\overline{y}, \overline{\lambda}_g; \underline{y}] \succsim_k g$, i.e., $f \succsim_k g$ by E2. Conversely, let $f \succsim_k g$. By (33), $\underline{u}_k(f) \geq \overline{u}_k(g)$.

When $f \in B_k(\overline{y}; \underline{y})$, we have $\overline{u}_k(f) = \lambda_f = \upsilon_k(f)$. Thus, $e0$ holds. Consider $e1$ to $e3$.

$e1$: Let $\widehat{f} = (f_1, ..., f_\ell) \in L_{k-1}(X)^\ell$. By (33) for $k-1$, $\overline{u}_{k-1}(f_t)$ is written as $\lambda_t \in \Pi_{k-1}$ for all $t = 1, ..., \ell$. Let $\widehat{\lambda} = (\lambda_1, ..., \lambda_\ell)$. In this case, $\widehat{e} * \overline{u}_{k-1}(\widehat{f}) = \widehat{e} * \widehat{\lambda} \in \Pi_k$. By (33) for k, it holds that $\widehat{e} * \overline{u}_{k-1}(\widehat{f}) = \widehat{e} * \widehat{\lambda} \geq \overline{u}_k(\widehat{e} * \widehat{f})$. The other assertion that $\underline{u}_k(\widehat{e} * \widehat{f}) \geq \widehat{e} * u_{k-1}(\widehat{f})$ is similarly proved.

$e2$: Let us prove that $\widehat{e} * \overline{u}_{k-1}(\widehat{f}) = \widehat{e} * \widehat{\lambda} = \overline{u}_k(f)$ for some decomposition $\widehat{f}$ of f. The other half can be proved similarly. By (33), we have $\overline{u}_k(f) = \overline{\lambda}_f \in \Pi_k$. Let $h = [\overline{y}, \overline{\lambda}_f; \underline{y}]$. By (33), h is the least preferred among $h' \succsim_k f$, where B1 and E2 are used. By (22), the preference $h \succsim_k f$ is derived by E1, i.e., there are decompositions $\widehat{h}, \widehat{f}$ of h, f such that $\widehat{h} \succsim_{k-1} \widehat{f}$ and $h = \widehat{e} * \widehat{h}$, $f = \widehat{e} * \widehat{f}$. Then, each f_t of $\widehat{f}$ has h'_t such that $\overline{u}_{t-1}(f_t) = u_{t-1}(h'_t) = \lambda'_{f_t}$. By (33) and $\widehat{h} \succsim_{k-1} \widehat{f}$, we have $\widehat{h} \succsim_{k-1} \widehat{h}'$. Since $u_k(h) = \widehat{e} * u_{k-1}(\widehat{h}) \geq \widehat{e} * u_{k-1}(\widehat{h}') = \widehat{e} * \overline{u}_{k-1}(\widehat{f}) \geq u_k(h)$, we have $\overline{u}_k(f) = u_k(h) = \overline{u}_{k-1}(\widehat{f})$. The other half can be proved similarly.

$e3$: By (33), $\overline{u}_k(x) = \lambda \in \Pi_k$ and $\underline{u}_k(x) = \lambda' \in \Pi_k$ for some λ and λ' in Π_k. Hence, $\overline{u}_k(x) = u_k([\overline{y}, \lambda; \underline{y}])$ and $\underline{u}_k(x) = u_k([\overline{y}, \lambda'; \underline{y}])$. This is the conclusion of $e2$.$\blacksquare$

6 Measurability, Comparability, and the EU hypothesis

Our main concern is the behavior of the preference stream $\langle \succsim_k \rangle_{k<\rho+1}$ and utility stream $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ for a finite ρ. Here, we study the concepts of measurable and non-measurable lotteries; incomparabilities are intimately related to non-measurable lotteries. Conversely, comparability and the EU hypothesis hold for measurable lotteries. In this section, ρ is still allowed to be finite or infinite. In the following, we assume that $\langle \succsim_k \rangle_{k<\rho+1}$ satisfies E0 to E3, relative to a base preference stream $\langle \trianglerighteq_k \rangle_{k<\rho+1}$ satisfying B0 to B3.

6.1 Measurable and non-measurable lotteries

We define the set M_k for $k < \rho + 1$ by

$$M_k = \{f \in L_k(X) : f \sim_k g \text{ for some } g = [\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})\}. \tag{34}$$

Each $f \in M_k$ is precisely measured by the benchmark scale $B_k(\overline{y}; \underline{y})$, while measurement of $f \in L_k(X) - M_k$ contains some indeterminacy. We call $f \in M_k$ *measurable* and $f \in L_k(X) - M_k$ *non-measurable*. Here, we study measurability and non-measurability.

Under our axioms, it holds that

$$\text{for each } f \in M_k, \text{ the probability weight } \lambda \text{ with } f \sim_k [\overline{y}, \lambda; \underline{y}] \text{ is unique,} \tag{35}$$

which we denote by λ_f. In M_k, no incomparabilities are observed; that is, if $f, g \in M_k$ with $\lambda_f \geq \lambda_g$, then $f \sim_k [\overline{y}, \lambda_f; \underline{y}] \succsim_k [\overline{y}, \lambda_g; \underline{y}] \sim_k g$. It also holds by Proposition 4.1.(1) that

$$M_k \subseteq M_{k+1} \text{ for all } k < \rho + 1. \tag{36}$$

To analyze the structure of M_k, we define $Y_k = M_k \cap X$ for all $k < \rho + 1$. By (36), we have $Y_k \subseteq Y_{k+1}$ for all $k < \rho + 1$. It follows from E0* that $y \in Y_k$ if and only if y and $[\overline{y}, \lambda_y; \underline{y}]$ are indifferent with respect to $\trianglerighteq_k$; pure alternative y is precisely measure by the benchmark scale. Measurability for a pure alternative is a property of a base preference relation $\trianglerighteq_k$. In Example 3.1, $Y_0 = Y_1 = \{\overline{y}, \underline{y}\}$ and $Y_2 = X = \{\overline{y}, y, \underline{y}\}$ in A and B of Fig.2, but in C, $Y_k = \{\overline{y}, \underline{y}\}$ even when $\rho = \infty$, i.e., y becomes never measurable.

The following lemma is about the structure of M_k.

Lemma 6.1 (1): If $f \in M_k$, then $f(y) = 0$ or 1 for all $y \in Y_k - Y_{k-1}$ and $f(y) = 0$ for all $y \in X - Y_k$, where $Y_{-1} = \emptyset$.

(2): $M_k \subseteq L_k(Y_k)$ for all $k < \rho + 1$.

One implication from Lemma 6.1.(1) is; for any $f \in M_k$ with $\delta(f) > 0$, if $f(x) > 0$, then $x \in Y_{k-1}$, which will be used in the proof of Theorem 6.1. Lemma 6.1.(2) states that we can concentrate on $L_k(Y_k)$ for consideration of M_k.

As indicated in (i) and (ii) in Section 1, each lottery $f \in L_k(Y_k)$ involves two types of depths, i.e., the measurement depth $\delta(\lambda_y)$ of $y \in Y_k$ with $f(y) > 0$ and the depth $\delta(f(y))$ of the probability value $f(y)$. In fact, measurability is characterized by their sum.

Theorem 6.1 (Measurability criterion). Let $k < \rho + 1$ and $f \in L_k(X)$. Let $k_f = \max\{\delta(\lambda_y) + \delta(f(y)) : f(y) > 0\}$. Then,

$$f \in M_k \text{ if and only if } k_f \leq k. \tag{37}$$

We can read (37) in two ways. One is to fix a lottery $f \in L_k(Y_k)$ but to change (increase) k. Any lottery f in $L_\infty(Y) = \cup_{k<\infty} L_k(Y) = \cup_{k<\infty} L_k(Y_k)$ becomes measurable when k is large enough. For example, when $f = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$ and $y \sim_2 [\overline{y}, \frac{83}{10^2}; \underline{y}]$, we have $k_f = \delta(\frac{83}{10^2}) + \delta(\frac{25}{10^2}) = 4$; by (37), $f \in M_k$ if and only if $k \geq 4$. The other reading of (37) is to fix a k and to change f. If $\delta(\lambda_y) > 0$ for some $y \in Y_k$, there is an $f \in L_k(Y_k)$ such that $\delta(\lambda_y) + \delta(f(y)) > k$; thus $f \notin M_k$ by (37). Thus, non-measurable lotteries exist as long as $\{\overline{y}, \underline{y}\} \subsetneq Y_k$.

Incomparability $\bowtie_k$ and indifference $\sim_k$ may appear similar: indeed, Shafer [28], p.469, discussed whether $\bowtie_k$ and $\sim_k$ could be defined together and pointed out a difficulty from the constructive point of view. Theorem 6.2 gives a clear distinction between $\sim_k$ and $\bowtie_k$. By E2, $\sim_k$ is transitive, but $\bowtie_k$ is not; indeed, we have distinct $f, h \in M_k$ with $f \sim_k h$, but by Theorem 6.2, for any $g \notin M_k$, $f \bowtie_k g$ and $g \bowtie_k h$. Also, reflexivity holds only for the measurable domain M_k.

Theorem 6.2. Let $f, g \in L_k(X)$.

(1) (No indifferences outside M_k): If $f \notin M_k$, then $f \nsim_k g$.

(2) (Reflexivity): $f \sim_k f$ if and only if $f \in M_k$.

6.2 EU hypothesis for measurable lotteries

Our theory is closely related to the expected utility hypothesis. It is explicitly assumed for the benchmark scale, i.e., B1 and b1. For the other part, it is only partially observed by looking at conditions $e1, e2$ for $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ as well as Axiom E1 for preference relations $\langle \succsim_k \rangle_{k<\rho+1}$. In fact, the EU hypothesis holds for the measurable domain M_k, which is now shown.

Let $\langle \boldsymbol{u}_k \rangle_{k<\rho+1}$ be the extended utility stream satisfying $e0$ to $e3$, given Theorem 5.1, relative to a base utility stream $\langle \boldsymbol{\upsilon}_k \rangle_{k<\rho+1}$. It follows from (34) and (25) that for any $f \in L_k(X)$,

$$\overline{u}_k(f) = \underline{u}_k(f) \text{ if and only if } f \in M_k. \tag{38}$$

Following our convention, we drop the upper and lower bars and write $u_k(f)$ for $f \in M_k$. In fact, $u_k(f)$ is expressed as the expected utility value of the base utility function υ_k; recall $\upsilon_k = \overline{\upsilon}_k = \underline{\upsilon}_k$ over Y_k because $Y_k = M_k \cap X$.

Theorem 6.3 (**EU hypothesis in the measurable domain**). For each $k < \rho + 1$, $u_k(f) = E_f(\upsilon_k)$ for all $f \in M_k$.

Thus, the EU hypothesis holds for measurable lotteries, which gives a simple method of calculation of $u_k(f)$. On the other hand, by (34) and (25), it holds that

$$\overline{u}_k(f) > \underline{u}_k(f) \text{ if and only if } f \in L_k(X) - M_k. \tag{39}$$

Thus, the EU hypothesis does not hold in the simple form for non-measurable lotteries.

6.3 Proofs

Proof of Lemma 6.1. We show (1) and (2) by induction on $k \geq 0$. Let $k = 0$. Since $Y_0 = M_0$, we have $f \in M_0 = Y_0 = L_0(Y_0)$, which implies (1) and (2). Suppose the induction hypothesis that (1) and (2) hold for k. Now, we take any $f \in M_{k+1}$.

Suppose, on the contrary, that $0 < f(y_o) < 1$ for some $y_o \in Y_{k+1} - Y_k$ or $0 < f(y_o) \leq 1$ for some $y_o \in X - Y_{k+1}$. If $f(y_o) = 1$ and $y_o \in X - Y_{k+1}$, by (35), there is no $g \in B_{k+1}(\overline{y}; \underline{y})$ such that $f \nsim_{k+1} g$, a contradiction to $f \in M_{k+1}$. Hence, we can assume $0 < f(y_o) < 1$. Since $y_o \in Y_{k+1} - Y_k$ or $y_o \in X - Y_{k+1}$, y_o differs from $\overline{y}$ and $\underline{y}$. Hence, $f \notin B_{k+1}(\overline{y}; \underline{y})$.

By $f \in M_{k+1}$, we have a $g \in B_{k+1}(\overline{y}; \underline{y})$ with $f \sim_{k+1} g$. Since $0 < f(y_o) < 1$, it holds that $f \in L_{k+1}(X) - X$. Hence, E3 is applied to $f \sim_{k+1} g$ with the middle $h = g$; we have decompositions $\widehat{f}, \widehat{g}$ of f, g with $\widehat{f} \succsim_k \widehat{g}$. If one preference was strict, then $f \succ_{k+1} g$ by E1*, impossible; hence $\widehat{f} \sim_k \widehat{g}$. By the induction hypothesis, we have $\widehat{f} \in L_k(Y_k)^\ell$. Thus, by E1, $f = \widehat{e} * \widehat{f} \in L_{k+1}(Y_k)$. Hence, we have the assertion (1) for $k+1$. This implies (2) for $k+1$; that is, any $f \in M_{k+1}$ has a support in Y_{k+1}.■

Proof of Theorem 6.1. We prove (37) by induction on $k \geq 0$. Let $k = 0$. Since $Y_0 = L_0(Y_0) = M_0$, it holds that $\delta(\lambda_f) = \delta(f(y)) = 0$ for all $f \in L_0(Y_0) = M_0$. Thus, (37) holds for $k = 0$. Now, suppose the induction hypothesis that (37) holds for k. We prove (37) for $k+1$. In the following, let $f \in L_{k+1}(Y_{k+1})$.

Let $f(y) > 0$ for some $y \in Y_{k+1} - Y_k$. By Lemma 6.1.(1), $f(y) = 1$, i.e., $f = y$ and $\delta(f(y)) = 0$. Since $y \in Y_{k+1} - Y_k$, we have $f = y \in Y_{k+1} \subseteq M_{k+1}$ and $\delta(\lambda_y) = k + 1$. In this sense, $f \in M_{k+1}$ $\Longleftrightarrow \delta(\lambda_y) + \delta(f(y)) = k + 1$, i.e., (37) holds for $k+1$.

Now, we take any $f \in L_{k+1}(Y_{k+1})$ satisfying

$$f(y) = 0 \text{ for any } y \in Y_{k+1} - Y_k. \tag{40}$$

We prove (a): $k_f \leq k + 1 > 0 \Rightarrow f \in M_{k+1}$; and (b): its converse.

(a): Let $k_f \leq k + 1$. Let $k^* = \max\{\delta(f(y)) : y \in Y_{k+1}\}$. Then, $k^* \leq k + 1$. Let $k^* = 0$. Then, $f = y$ for some $y \in Y_{k+1}$; hence, $f = y \in Y_{k+1} \subseteq M_{k+1}$.

Let $k_f < k+1$. By the induction hypothesis, we have $f \in M_k$. Now, let $k_f = k+1$. Then, by Lemma 2.1, we have a decomposition $\widehat{f}$ of f such that $\widehat{f} \in L_k(Y_{k+1})$ and for all $t \leq \ell$, $\delta(f_t(x)) < \delta(f(x))$ for all $x \in Y_{k+1}$ with $f(x) > 0$. This implies $k_{f_t} < k_f$ for $t = 1, ..., \ell$. Thus, by the induction hypothesis, we have $f_t \in M_k$ for $t = 1, ..., \ell$. Thus, we have $g_t \in B_k(\overline{y}; y)$ with $f_t \sim_k g_t$ for $t = 1, ..., m$. By E1, we have $f = \widehat{e}{*}\widehat{f} \sim_{k+1} \widehat{e}{*}\widehat{g} \in B_{k+1}(\overline{y}; y)$. This means $f \in M_{k+1}$.

(b): Let $f \in M_{k+1}$. If $f \in M_k$, we have the right-hand side of (37) by the induction hypothesis. Hence, we can suppose $f \in M_{k+1} - M_k$. Also, we can assume $f \notin B_{k+1}(\overline{y}; \underline{y})$; indeed, if $f \in B_{k+1}(\overline{y}; \underline{y})$, i.e., $f = [\overline{y}, \lambda_f; \underline{y}]$ for some $\lambda_f \in \Pi_{k+1}$, we have $\max\{\delta(\lambda_x) + \delta(f(x)) : f(x) > 0\} = \delta(\lambda_f) \leq k+1$. Also, it holds that $0 < f(y) < 1$ for some $y \in Y_k$. Indeed, by (40), we have $0 < f(y)$ for some $y \in Y_k$ and $f(y) < 1$ by $f \in M_{k+1} - M_k$.

Since $f \in M_{k+1} - M_k$, we have $f \sim_{k+1} g$ for some $g \in B_{k+1}(\overline{y}; \underline{y})$. By E3, there are decompositions $\widehat{f}$ and $\widehat{g}$ of f, g, and $\widehat{f} \sim_k \widehat{g}$. Thus, $\widehat{f} \in (M_k)^\ell$. By the induction hypothesis, we have $\delta(\lambda_y) + \delta(f_t(y)) \leq k$ for all $y \in Y_k$ with $f_t(y) > 0$ and $t = 1, ..., \ell$. Since $f = \widehat{e}{*}\widehat{f}$, it holds that $\delta(f(y)) \leq \max_{t \leq \ell} \delta(f_t(y)) + 1$ for all $y \in Y_k$ with $f_t(y) > 0$. Since $\delta(f(y)) = 0$ for all $y \in Y_{k+1} - Y_k$ by (40), we have $\delta(\lambda_y) + \delta(f(y)) \leq k+1$ for all $y \in Y_{k+1}$ with $f(y) > 0$.■

Proof of Theorem 6.2. **(1)**: Suppose that $f \notin M_k$ and $g \in M_k$. Then, $g \sim_k [\overline{y}, \lambda_g; \underline{y}]$. If $f \sim_k g$, then $f \in M_k$ by E2, a contradiction. Hence, $f \nsim_k g$. Now, let $f, g \notin M_k$. Suppose $f \sim_k g$. By Lemma 5.1, $f \succsim_k h \succsim_k g$ for some $h \in B_k(\overline{y}; \underline{y})$ but $g \succsim_k h' \succsim_k f$ for some $h' \in B_k(\overline{y}; \underline{y})$. By E2, this implies $f \sim_k h \sim_k g$ and $h = h'$. This is impossible since $f, g \notin M_k$. Hence, $f \nsim_k g$.

(2): The *if* part is by (34) and E2. The *only-if* part (contrapositive) follows from (1).■

Proof of Theorem 6.3. When $f \in B_k(\overline{y}; \underline{y}) \cup X$, we have, by $e0$, $\overline{u}_k(f) = \underline{u}_k(f) = \upsilon_k(f)$; if $f = [\overline{y}, \lambda; \underline{y}] \in B_k(\overline{y}; \underline{y})$, then, by $b1$, $\upsilon_k(f) = \lambda\upsilon_k(\overline{y}) + (1-\lambda)\upsilon_k(\underline{y}) = E_f(\upsilon_k)$; and if $f = x \in X$, then $\upsilon_k(f) = 1 \times \upsilon_k(x) = E_f(\upsilon_k)$. Now, we show the assertion by induction on $k < \rho + 1$. The case $k = 0$ is included in the case $f \in B_k(\overline{y}; \underline{y}) \cup X$, $k < \rho + 1$.

Suppose that the assertion holds for $k - 1$. Let $f \in M_k$. We assume that $f \notin B_k(\overline{y}; \underline{y}) \cup X$. Then, $\delta(f) > 0$; as remarked after Lemma 6.1, if $f(x) > 0$, then $x \in Y_{k-1}$, so, $\upsilon_k(x) = \upsilon_{k-1}(x)$. Since $f \in M_k$, we have $f \sim_k h$ for some $h \in B_k(\overline{y}; \underline{y})$. By (22), this $f \sim_k h$ is derived by E1 from the decompositions $\widehat{f}$ and $\widehat{h}$ of f and h. By E1*, it holds that $\widehat{f} \sim_{k-1} \widehat{h}$. Hence, $\widehat{f} \in (M_{k-1})^\ell$. By the induction hypothesis, we have $\overline{u}_{k-1}(f_t) = \underline{u}_{k-1}(f_t) = E_{f_t}(\upsilon_{k-1})$ for $t \leq \ell$. Thus, by $e1$, $\widehat{e} * \overline{u}_{k-1}(\widehat{f}) \geq \overline{u}_k(f) \geq \underline{u}_k(f) \geq \widehat{e} * \underline{u}_{k-1}(\widehat{f})$; thus these are all equal. Now, we have $u_k(f) = \widehat{e} * \underline{u}_{k-1}(\widehat{f}) = \sum_t \frac{1}{\ell} E_{f_t}(\upsilon_{k-1}) = E_f(\upsilon_k)$.■

7 Toward the Classical EU Theory

Up to Section 4, we have focussed on the development of our theory from the constructive point of view. Theorem 4.1 reflects this constructiveness, which is extracted by Axiom E3 as well as condition $b3$. These are constraints on depths and interact on a finite cognitive bound. When we delete these constraints, we go to the classical EU theory. Nevertheless, we have two steps to the classical EU theory; the first is to go to the case $\rho = \infty$ with $Y = \cup_{k<\infty} Y_k$, where $Y_k = M_k \cap X$ for $k < \infty$. This Y is typically a proper subset of X so that each $y \in Y$ is exactly measured at some finite k. The second is to allow all real number probabilities for lotteries, i.e., we take lotteries in $L_{[0,1]}(Y)$. In this section, we focus on the set Y and give only a remark on the case of X.

7.1 Two steps to the classical EU theory

Let $\langle \precsim_k \rangle_{k<\infty}$ be a preference stream satisfying E0 to E2, relative to a base preference stream $\langle \trianglerighteq_k \rangle_{k<\infty}$ satisfying B0 to B3. Note that Axiom E3 is not assumed in this section. Also, let $\langle \boldsymbol{\upsilon}_k \rangle_{k<\infty}$ be a base utility stream satisfying (18) of Theorem 3.1. Note that $\langle \boldsymbol{\upsilon}_k \rangle_{k<\infty}$ satisfies conditions $b0$ to $b3$.

The limit preference relation of $\langle \succsim_k \rangle_{k<\infty}$ is defined to be $\succsim_\infty = \cup_{k<\infty} \succsim_k$; that is, the decision maker can go to a layer of any depth k for his preference comparisons. Let $x \in Y = \cup_{k<\infty} Y_k$, where $Y_k = \{x \in X : x$ and $[\overline{y}, \lambda; \underline{y}]$ are indifferent with respect to $\trianglerighteq_k$ for some $\lambda \in \Pi_k\}$ for each $k < \infty$. By Theorem 3.2 (uniqueness), we normalize $\langle \boldsymbol{\upsilon}_k \rangle_{k<\infty}$ so that for each $x \in Y$,

$$x \sim_k [\overline{y}, \upsilon_k(x); \underline{y}] \text{ for some } k < \infty. \tag{41}$$

This holds over Y. Now, we define, for each $x \in X$,

$$\overline{\upsilon}_\infty(x) = \lim_{k\to\infty} \overline{\upsilon}_k(x) \text{ and } \underline{\upsilon}_\infty(x) = \lim_{k\to\infty} \underline{\upsilon}_k(x). \tag{42}$$

These are well defined by $b3$. In particular, when $x \in Y$, there is a k_x such that $\overline{\upsilon}_k(x) = \underline{\upsilon}_k(x)$ for all $k \geq k_x$. For a large enough k, $\overline{\upsilon}_k(x) = \underline{\upsilon}_k(x)$. Hence, we can write $\overline{\upsilon}_\infty(x) = \underline{\upsilon}_\infty(x) = \upsilon_\infty(x)$ for $x \in Y$. Note that these definitions do not need Axiom E3 at all.

We have the following theorem.

Theorem 7.1 (EU hypothesis without cognitive bounds). For all $f, g \in L_\infty(Y)$,

$$f \succsim_\infty g \text{ if and only if } E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty). \tag{43}$$

Thus, without the cognitive restriction, the EU hypothesis holds for the limit relation $\succsim_\infty$ over the set of lotteries $L_\infty(Y) = \cup_{k<\infty} L_k(Y) = \cup_{k<\infty} L_k(Y_k)$. It is important to notice that for each pair $f, g \in L_\infty(Y)$, the equivalence (43) holds for a large enough k, i.e., (43) is written as $f \succsim_k g$ if and only if $E_f(\upsilon_k) \geq E_g(\upsilon_k)$. When Y is a finite set, the EU hypothesis holds for large k uniform over Y. This theorem is proved without a cognitive bound ρ; Axiom E3 and condition $e3$ are dropped. Thus, this theorem differs from Theorem 6.3.

Now, we compare directly the pair $(L_\infty(Y), \succsim_\infty)$ with the classical EU theory with Axioms NM0 to NM2 in Section 2.1. First, we need to replace the entire set X of pure alternatives by the set Y of measurable pure alternatives. Then, Axiom NM0 (completeness and transitivity) follows from (43). Axiom NM1 should be weakened, since Π_∞ is not closed with division;

NM1^o : for any $f \in L_\infty(Y)$, there is a $\lambda \in \Pi_\infty$ such that $f \sim_\infty [\overline{y}, \lambda; \underline{y}]$.

That is, for any $f \in L_\infty(Y)$, there is some $k < \infty$ such that $f \sim_k [\overline{y}, \lambda; \underline{y}]$ for some $\lambda \in \Pi_k$. This is, more or less, the definition of measurability (34). Axiom NM2 needs to restrict the set of scalars to $\Pi_\infty = \cup_{k<\infty} \Pi_k$. We summarize this observation.

Theorem 7.2 (Axioms NM0, NM1^o, NM2 for $(L_\infty(Y), \succsim_\infty)$). The system $(L_\infty(Y), \succsim_\infty)$ satisfies Axioms NM0, NM1^o, and NM2 (with $\Pi_\infty = \cup_{k<\infty} \Pi_k$).

This is interpreted as meaning that the axiomatic system NM0, NM1^o, NM2 are an abbreviated fragment of our theory.

The next step is to jump to $L_{[0,1]}(Y)$ and to extend the relation $\succsim_\infty$ to $L_{[0,1]}(Y)$. The extension is uniquely determined and it is a relation in the classical theory. However, this extension involves non-constructive components; $L_{[0,1]}(Y)$ is uncountable but $L_\infty(Y)$ is countable. First, we have the following lemma.

Lemma 7.1. $L_\infty(Y)$ is a dense subset of $L_{[0,1]}(Y)$.

Now, we define a binary relation $\succsim_E$ over $L_{[0,1]}(Y)$ by: for any $f, g \in L_{[0,1]}(Y)$,

$$f \succsim_E g \text{ if and only if } E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty). \tag{44}$$

We have the following theorem.

Theorem 7.3 (Unique extension). The relation $\succsim_E$ defined by (44) is a unique extension of $\succsim_\infty$ to $L_{[0,1]}(Y)$ with NM0 to NM2; that is,

(1): for any $f, g \in L_\infty(Y)$, $f \succsim_\infty g$ if and only if $f \succsim_E g$.

(2): $\succsim_E$ satisfies NM0 to NM2.

This theorem is proved by the denseness of $L_\infty(Y)$ in $L_{[0,1]}(Y)$ and the continuity of $E_f(\upsilon_\infty)$ with respect to f relative to point-wise convergence, where $E_f(\upsilon_\infty)$ is *continuous* iff for any sequence $\{f^\nu\}$ in $L_{[0,1]}(Y)$ and $f \in L_{[0,1]}(Y)$, if $f^\nu(y) \to f(y)$ for each $y \in Y$, then $\lim_{\nu\to\infty} E_{f^\nu}(\upsilon_\infty) = E_f(\upsilon_\infty)$. The proof of the theorem may appear to be constructive, but the last extension step to $\succsim_E$ is non-constructive, since probabilities newly involved in $f \in L_{[0,1]}(Y) - L_\infty(Y)$ may be given only in a nonconstructive manner.[8]

In (42), the limit utility functions $\overline{\upsilon}_\infty$ and $\underline{\upsilon}_\infty$ are defined over X. The set X is divided into $X_E = \{x \in X : \overline{\upsilon}_\infty(x) = \underline{\upsilon}_\infty(x)\}$ and $X_S = \{x \in X : \overline{\upsilon}_\infty(x) > \underline{\upsilon}_\infty(x)\}$. The set $Y = \cup_{k<\infty} Y_k$ may be a proper subset of X_E. We may ask whether the results given above could hold for X_E. In the above results, the limit can be regarded as large finite, but $X_E - Y$ and X_S may not enjoy such finite approximation of the limit. For example, it would be possible $\overline{\upsilon}_k(x) > \underline{\upsilon}_k(x)$ for all $k < \infty$ but $\overline{\upsilon}_\infty(x) = \underline{\upsilon}_\infty(x)$; here, comparability may appear suddenly in limit. Hence, this situation differs from $Y = \cup_{k<\infty} Y_k$.

We may have more subtle relations from large finite worlds to the limit. Since the convergences in X_E and X_S are monotone but arbitrary, the full real number theory appear here (see Mendelson [23], p.217). This is far from our original motivation of bounded rationality. For the entire understanding, however, it would be helpful to see how our theory behaves in X_E and X_S for $\rho = \infty$. This is an open problem.

7.2 Proofs

Proof of Theorem 7.1. Let Y' be any finite subset of Y. We prove this by induction on $k < \infty$ that

$$\text{for any } f \in L_k(Y'),\ f \sim_\infty [\overline{y}, E_f(\upsilon_\infty); \underline{y}] \tag{45}$$

Let $f \in L_\infty(Y) = \cup_{k<\infty} L_k(Y)$. Since f to has a finite support S, f belongs to $L_k(Y')$ for some finite subset Y' of Y. Hence, by (45), $f \sim_\infty [\overline{y}, E_f(\upsilon_\infty); \underline{y}]$. Now, recall that for some k_o, $E_f(\upsilon_\infty) = E_f(\upsilon_k)$ for all $k \geq k_o$. Let $f, g \in L_\infty(Y)$. For large enough k, $f \succsim_\infty g$ if and only if $[\overline{y}, E_f(\upsilon_\infty); \underline{y}] \sim_\infty f \succsim_\infty g \sim_\infty [\overline{y}, E_f(\upsilon_\infty); \underline{y}]$ if and only if $[\overline{y}, E_f(\upsilon_k); \underline{y}] \sim_k f \succsim_k$

[8] We avoid the use of a topology for Axiom NM1. This does not change the content of classical EU theory as long as the set of lotteries is given as $L_{[0,1]}(Y)$. However, NM1 allows to restrict it to $L_{[0,1]\cap\mathbb{Q}}(Y)$. In this case, the extension result given in Theorem 7.3 is regarded as approximately constructive in the theoretical sense.

$g \sim_k [\overline{y}, E_f(\upsilon_\infty); \underline{y}]$ if and only if $[\overline{y}, E_f(\upsilon_k); \underline{y}] \succsim_k [\overline{y}, E_f(\upsilon_k); \underline{y}]$ if and only if $E_f(\upsilon_k) \geq E_f(\upsilon_k)$ if and only if $E_f(\upsilon_\infty) \geq E_f(\upsilon_\infty)$. Here, we use B1 and E2.

Now, we show (45) by induction on $k < \infty$. For $k = 0$, $f \in L_0(Y')$ is a pure alternative $x \in Y'$. Hence, $x \sim_k [\overline{y}, \upsilon_k(x); \underline{y}]$ for some k, i.e., $x \sim_k [\overline{y}, E_x(\upsilon_k); \underline{y}] = [\overline{y}, E_x(\upsilon_\infty); \underline{y}]$. Suppose the induction hypothesis that (45) holds for k. Let $f \in L_{k+1}(Y') - L_k(Y')$. By Lemma 2.1, there is a decomposition $\widehat{f} = (f_1, ..., f_n)$ of f. Since each f_t is in $L_k(Y')$, it holds that $f_t \sim_\infty [\overline{y}, E_{f_t}(\upsilon_\infty); \underline{y}]$. This is written as: for some $k < \infty$, $f_t \sim_k [\overline{y}, E_{f_t}(\upsilon_k); \underline{y}]$ for all $t = 1, ..., \ell$. Thus, by E1, we have $f = \widehat{e} * \widehat{f} \sim_k \widehat{e} * ([\overline{y}, E_{f_1}(\upsilon_k); \underline{y}], ..., [\overline{y}, E_{f_\ell}(\upsilon_k); \underline{y}]) = [\overline{y}, \sum_t \frac{1}{\ell} E_{f_t}(\upsilon_k); \underline{y}] = [\overline{y}, E_{\widehat{e} * \widehat{f}}(\upsilon_k); \underline{y}] = [\overline{y}, E_f(\upsilon_k); \underline{y}]$. Hence, $f \sim_\infty [\overline{y}, E_f(\upsilon_\infty); \underline{y}]$.■

Proof of Theorem 7.2. By (43), Axiom NM0 holds for $(L_\infty(Y), \succsim_\infty)$. As mentioned, NM1^o follows from (34). We can prove ID1 of NM2 by Theorem 7.1 that if $f, g, h \in L_\infty(Y)$ and $\alpha \in \Pi_\infty$, if $f \succ_\infty g$, then $\alpha f * (1-\alpha) h \succ_\infty \alpha g * (1-\alpha) h$. ID2 is similar. ■

Proof of Lemma 7.1. Take any $f \in L_{[0,1]}(Y)$. This f has a finite support $S = \{y_0, y_1, ..., y_m\}$ in Y with $f(y_t) > 0$ for $t = 0, ..., m$. We construct a sequence $\{g^\nu\}_{\nu=\nu_o}^\infty$ so that $g^\nu \in L_\infty(Y)$ for $\nu \geq \nu_0$, and for each $y \in Y$, $g^\nu(y) \to f(y)$ as $\nu \to \infty$. When $m = 0$, it suffices to let $g^\nu = f$ for all $\nu \geq 0$. In the following, we assume $m \geq 1$.

For any natural number ν, let $z_{\nu,t} = \max\{\pi_t \in \Pi_\nu : \pi_t \leq f(y_t)\}$ for all $t = 0, ..., m$. Since S is fixed and finite, there is a ν_o such that for all $\nu \geq \nu_o$, $\frac{1}{\ell^\nu} \leq z_{\nu,t} \leq 1 - \frac{1}{\ell^\nu}$ for all $t = 0, ..., m-1$ and $\frac{1}{\ell^\nu} \leq 1 - \sum_{t<m} z_{\nu,t} \leq 1 - \frac{m}{\ell^\nu}$. Also, we define $u_{\nu,0}, ..., u_{\nu,m}$ by

$$u_{\nu,t} = \begin{cases} z_{\nu,t} & \text{if } t < m \\ 1 - \sum_{t<m} z_{\nu,t} & \text{if } t = m. \end{cases}$$

Then, $\sum_{t \leq m} u_{\nu,t} = 1$ and $u_{\nu,t} \in \Pi_\nu$ for all $t \leq m-1$. Since $\frac{1}{\ell^\nu} \leq 1 - \sum_{t<m} z_{\nu,t} = u_{\nu,m} \leq 1 - \frac{m}{\ell^\nu}$, we have $u_{\nu,m} \in \Pi_\nu$.

We define $\{g^\nu\}_{\nu=\nu_o}^\infty$ by

$$g^\nu(y) = \begin{cases} 0 & \text{if } y \in Y - S \\ u_{\nu,t} & \text{if } y = y_t \in S. \end{cases}$$

Then, each g^ν belongs to $L_\nu(Y)$. For each $t \leq m-1$, since $g^\nu(y_t) = u_{\nu,t} \leq f(y_t) < u_{\nu,t} + \frac{1}{\ell^\nu} = g^\nu(y_t) + \frac{1}{\ell^\nu}$ for all $\nu \geq \nu_o$, we have $\lim_{\nu \to \infty} g^\nu(y_t) = f(y_t)$. Since $g^\nu(y_m) - \frac{m}{\ell^\nu} = 1 - \sum_{t<m} g^\nu(y_t) - \frac{m}{\ell^\nu}$ $\leq 1 - \sum_{t<m} f(y_t) = f(y_m) \leq 1 - \sum_{t<m} g^\nu(y_t) = g^\nu(y_m)$ for all $\nu \geq \nu_o$. Thus, $\lim_{\nu \to \infty} g^\nu(y_m) = f(y_m)$.■

Proof of Theorem 7.3.(1): Let $f, g \in L_\infty(Y)$. Then, if k is large enough, then $f, g \in L_k(Y_k)$ and $\upsilon_k(x) = \upsilon_\infty(x)$ for all $x \in Y_k$. Now, suppose $f \succsim_\infty g$. Then $f \succsim_k g$, equivalently, $E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty)$, which implies $f \succsim_E g$. Conversely, if $f \succsim_E g$, then $E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty)$, equivalently, $f \succsim_k g$ for a large enough k. Thus, $f \succsim_\infty g$.

(2): The relation $\succsim_E$ is a complete preordering, i.e., it satisfies NM0. Let us see NM1; let $f \succsim_E h \succsim_E g$. Then, $E_f(\upsilon_\infty) \geq E_h(\upsilon_\infty) \geq E_g(\upsilon_\infty)$. Choose a $\lambda \in [0,1]$ so that $E_h(\upsilon_\infty) = \lambda E_f(\upsilon_\infty) + (1-\lambda) E_g(\upsilon_\infty)$. Then, $E_{\lambda f + (1-\lambda) g}(\upsilon_\infty) = \lambda E_h(\upsilon_\infty) = (1-\lambda) E_g(\upsilon_\infty) = E_h(\upsilon_\infty)$. Finally, we can see NM2.ID1: let $f \succsim_E g$, i.e., $E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty)$. Hence, for any $\lambda \in [0,1]$ and $h \in L_{[0,1]}(Y)$, we have $E_{\lambda f + (1-\lambda) h}(\upsilon_\infty) = \lambda E_f(\upsilon_\infty) + (1-\lambda) E_h(\upsilon_\infty) \geq \lambda E_g(\upsilon_\infty) + (1-\lambda) E_h(\upsilon_\infty) = E_{\lambda g + (1-\lambda) h}(\upsilon_\infty)$, i.e., $\lambda f + (1-\lambda) h \succsim_E \lambda g + (1-\lambda) h$. Similarly, we can verify NM2.ID2.

Finally, we show that $\succsim_E$ is uniquely determined. Suppose that $\succsim'_E$ is an extension of $\succsim_\infty$ in the sense of (1) and satisfies NM0 to NM2. Then, for any $f, g \in L_\infty(Y)$, $f \succsim_E g$ if and only if $f \succsim_\infty g$,and by the supposition, $f \succsim'_\infty g$ if and only if $E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty)$. Hence, for any

$f, g \in L_\infty(Y)$, $f \succsim'_E g$ if and only if $E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty)$.

Now, let $f, g \in L_{[0,1]}(Y)$ with $f \succsim'_E g$. By Lemma 7.1, there are sequences $\{f^\nu\}$ and $\{g^\nu\}$ in $L_\infty(Y)$ such that they point-wise converge to f and g. As stated above, $E_h(\upsilon_\infty)$ is continuous with respect to h. Then, $E_f(\upsilon_\infty) = \lim_{\nu\to\infty} E_{f^\nu}(\upsilon_\infty) \geq \lim_{\nu\to\infty} E_{g^\nu}(\upsilon_\infty) = E_g(\upsilon_\infty)$. We have shown that that for any $f, g \in L_{[0,1]}(Y)$, $E_f(\upsilon_\infty) \geq E_g(\upsilon_\infty)$ if and only if $f \succsim'_E g$. Thus, $f \succsim_E g$ if and only if $f \succsim'_E g$.■

8 An Application to a Kahneman-Tversky Example

We apply our theory to an experimental result reported in Kahneman-Tversky [15]. The experimental instance is formulated as Examples 3.1 and 5.1, and the relevant lotteries are $c = [\overline{y}, \frac{2}{10}; \underline{y}]$ and $d = \frac{25}{100}y * \frac{75}{100}\underline{y}$, which are incomparable for people with $\rho = 2$. It is the key how the observed behaviors are connected to the incomparabilities predicted in our theory. First, we look at the Kahneman-Tversky example, and then we make a certain postulate to have such a connection.

In the Kahneman-Tversky example, 95 subjects were asked to choose one from lotteries a and b, and one from c and d. In the first problem, 20% chose a, and 80% chose b. In the second, 65% chose c; and the remaining chose d.

$$\begin{array}{rcll} a &=& [4000, \frac{80}{10^2}; 0] \ (20\%) \ \text{ vs. } & b = 3000 \text{ with probability 1 } (80\%) \\ c &=& [4000, \frac{20}{10^2}; 0] \ (65\%) \ \text{ vs. } & d = [3000, \frac{25}{10^2}; 0] \qquad\qquad (35\%). \end{array}$$

The case of modal choices, denoted by $b \wedge c$, contradicts the classical EU theory. Indeed, these choices are expressed in terms of expected utilities as:

$$\begin{array}{rcl} 0.80u(4000) + 0.20u(0) &<& u(3000) \\ 0.20u(4000) + 0.80u(0) &>& 0.25u(3000) + 0.75u(0). \end{array} \tag{46}$$

Normalizing $u(\cdot)$ with $u(0) = 0$, and multiplying 4 to the second inequality, we have the opposite inequality of the first, a contradiction. The other case violating the classical EU theory is $a \wedge d$. It predicts the outcomes $a \wedge c$ and $b \wedge d$, depending upon the value $u(3000)$. This is a variant of "common ratio effect" discussed in the literature, which is briefly discussed in Remark 8.1.

In [15], no more information is mentioned other than the above percentages. Consider three possible distributions of the answers in terms of percentages over the four cases. In Table 8.1, the first, second, or third entry in each cell is the percentage derived by assuming 65%, 52%, or 45% for $b \wedge c$. The first 65% is the maximum possibility for $b \wedge c$, which leads to 0% for $a \wedge c$, and these determine the 20% for $a \wedge d$ and 15% for $b \wedge d$. The second entries are based on the assumption that the choices of b and c are stochastically independent, for example, $52 = (0.80 \times 0.65) \times 100$ for $b \wedge c$. In the third entries, 45% is the minimum possibility for $b \wedge c$. We interpret this table as meaning that each cell was observed at a significant level.

Table 8.1

	$c : 65\%$	$d : 35\%$
$a : 20\%$	$a \wedge c$: EU: 0 //**13**//20	$a \wedge d$: paradox: 20//**7**// 0
$b : 80\%$	$b \wedge c$: paradox: 65//**52**//45	$b \wedge d$: EU: 15//**28**//35

Let $\overline{y} = 4000$, $\underline{y} = 0$, $y = b = 3000$, and $\rho \geq 2$. Consider two cases *A:* $\boldsymbol{v}_2(y) = \boldsymbol{u}_2(y) = [\frac{77}{10^2}, \frac{77}{10^2}]$ and *B*: $\boldsymbol{v}_2(y) = \boldsymbol{u}_2(y) = [\frac{83}{10^2}, \frac{83}{10^2}]$ in Example 3.1, and recall that $\boldsymbol{u}_2(a) = \boldsymbol{u}_2([\overline{y}, \frac{80}{10^2}; \underline{y}])$

$= [\frac{80}{10^2}, \frac{80}{10^2}]$. Our theory predicts, independent of ρ, the choice a (or b) in case A (B). We assume that the distribution of subjects over A and B is the same as that given in Table 8.1, i.e.,

$$A : B = 20\% : 80\%. \tag{47}$$

We calculate the distribution of choices c and d based upon (47) and the distribution of ρ.

Comparisons between c and d depend upon ρ. In case $\rho \geq 4$, it follows from the calculation results in Example 5.1 that in case A, $c = [\overline{y}, \frac{2}{10}; \underline{y}] \succ_4 [\overline{y}, \frac{1925}{10^4}; \underline{y}] \sim_4 [y, \frac{25}{10^2}; \underline{y}] = d$; so c is chosen, and in case B, $d = [y, \frac{25}{10}; \underline{y}] \sim_4 [\overline{y}, \frac{2075}{10^4}; \underline{y}] \succ_4 [\overline{y}, \frac{20}{10^2}; \underline{y}] = c$; so d is chosen. In sum, our theory predicts only the diagonal cells $a \wedge c$ and $b \wedge d$ for cases A and B, which are the same as the predictions of the classical EU theory. Thus, if all subjects have their cognitive bounds $\rho \geq 4$, our theory is inconsistent with the experimental result.

Let $\rho = 3$. In case A, (30) states $\boldsymbol{u}_3(c) = [\frac{2}{10}, \frac{2}{10}] \geq_I \boldsymbol{u}_3(d) = [\frac{199}{10^3}, \frac{189}{10^3}]$, and in case B, (31) states $\boldsymbol{u}_3(d) = [\frac{211}{10^3}, \frac{201}{10^3}] \geq_I \boldsymbol{u}_3(c) = [\frac{2}{10}, \frac{2}{10}]$. Hence, people with $\rho = 3$ behave in the same manner as those with $\rho \geq 4$, though d is non-measurable.

In case $\rho = 2$. (29) states that people in cases A and B show the same base utility evaluation of d, i.e., $\boldsymbol{u}_2(d) = [\frac{23}{10^2}, \frac{14}{10^2}]$. Since $\boldsymbol{u}_2(c) = [\frac{2}{10}, \frac{2}{10}]$, c and d are incomparable for these people.

Here, we find a conflict between our theory and the reported experimental result in that every subject chose one lottery in each of the above choice problems, while our theory states that c and d are incomparable for people with $\rho = 2$. The issue is how a subject behaves for the choice problem when the lotteries are incomparable for him. In such a situation, a person would typically be forced (e.g., following social customs) to make a choice.[9] Here, we assume the following postulate for choice behavior for a subject having incomparabilities:

Postulate BH: each subject makes a random choice between c and d, following the probabilities proportional to the distances from $u_2(c)$ to $\underline{u}_2(d)$ and from $\overline{u}_2(d)$ to $u_2(c)$.

Since $\boldsymbol{u}_2(d) = [\frac{23}{10^2}, \frac{14}{10^2}]$ and $\boldsymbol{u}_2(c) = [\frac{2}{10}, \frac{2}{10}]$, the probabilities for the choices c and d are $\frac{2}{10} - \frac{14}{10^2} : \frac{23}{10^2} - \frac{2}{10} = 2 : 1$.

Table 8.2

	A	B
$\rho = 2$	$c : d = 2 : 1$	$c : d = 2 : 1$
$\rho \geq 3$	$c : d = 1 : 0$	$c : d = 0 : 1$

Table 8.2 summarizes the above calculated results. To see the relationship between Table 8.1 and Table 8.2, we specify the distribution of people over $\rho = 2, 3, ...$ We consider two distributions of ρ

$$r_2 : r_{+3} = 9 : 1 \text{ and } r_2 : r_{+3} = 8 : 2,$$

where r_{3+} is the ratio of subjects with $\rho \geq 3$. These are adopted based on the idea that $\rho = 3$ is already quite precise, and the portion of people with $\rho \geq 3$ is already small.

[9] It may be difficult for people to show incapability of answering a question if it appears linguistically and logically.clear. The present author knows only one person in our profession to refuse consciously to answer such a question. Davis-Maschler [5], Sec.6 reported that when a number of game theorists/economists were asked about their predictions about choices in a specific example in a cooperative game theory, Martin Shubik refused to answer a questionnaire. It was his reason that the specification in terms of cooperative game is not enough to have a precise prediction for the question. Usually, people answer such a question, often unconsciously by filling up gaps.

In the case $r_2 : r_{+3} = 9 : 1$, the percentage of the choices $a \wedge c$ is calculated as $100 \times \frac{2}{10} \times (\frac{9}{10} \times \frac{2}{3} + \frac{1}{10} \times 1) = 14\%$. The corresponding percentages $b \wedge c$ is calculated as $100 \times \frac{8}{10} \times (\frac{9}{10} \times \frac{2}{3} + \frac{2}{10} \times 0) = 48\%$. Thus, we obtain Table 8.3. Table 8.4 is based on $r_2 : r_{+3} = 8 : 2$.

Table 8.3: $r_2 : r_{+3} = 9 : 1$

	$c : 62\%$	$d : 38\%$
$a : 20\%$	$a \wedge c : 14$	$a \wedge d : 6$
$b : 80\%$	$b \wedge c : 48$	$b \wedge d : 32$

Table 8.4: $r_2 : r_{+3} = 8 : 2$

	$c : 55\%$	$d : 45\%$
$a : 20\%$	$a \wedge c : 12$	$a \wedge d : 8$
$b : 80\%$	$b \wedge c : 43$	$b \wedge d : 37$

The results in Tables 8.3 and 8.4 are quite compatible to Table 8.2. Perhaps, we should admit that this is based upon our specifications of parameter values as well as Postulate BH. To make stronger assertions, we need to think about more cases of parameter values and different forms of BH. Nevertheless, this study may lead to observations on new aspects on bounded rationality that ρ seems quite small.

Remark 8.1 (Common ratio effect). The anomaly mentioned in (46) is often called the "common ratio effect" (cf. Prelec [25], van de Kuilen-Wakker [22], and their references). It refers to the observation such as the fact that the opposite of the second inequality in (46) is obtained from the first with multiplication of $1/4 = 25/10^2$. In our theory for case B with $\rho = 2$, b is strictly preferred to a, but c and d, which are obtained by the multiplication, are incomparable, and the independence condition, NM2, is violated. We made the additional postulate BH to connect incomparability to the observed behavior in the experiment. The postulate shows a bigger tendency to choose c. In this sense, our result shows the "common ratio effect". However, Postulate BH does not directly take depths for the choice behavior of agents. Perhaps, there are different postulates taking depths of lotteries to explain the "common ratio effect" more directly. This is an open problem (see Section 9, [c]).

9 Conclusions

We developed the EU theory with probability grids and preference formation. The permissible probabilities are restricted to the form of ℓ-ary fractions up to a given cognitive bound ρ. We divide the argument into the measurement step of preferences (utilities) on pure alternatives in terms of the benchmark scale and the extension step to lotteries with more risks. We have taken the constructive point of view of the decision maker for our theory. The development includes the approach in terms of vector-valued utilities with the interval order due to Fishburn [8]. The connections between these two approaches are shown to be equivalent in Sections 3 to 5. These approaches are complementary; each may give better interpretations as well as some technical merits over the other.

When the cognitive bound ρ is finite, the resultant preference relation $\succsim_\rho$ over $L_\rho(X)$ is incomplete. We divided $L_\rho(X)$ into the set M_ρ of measurable lotteries and its complement $L_\rho(X) - M_\rho$. The resultant $\succsim_\rho$ is complete over M_ρ, while it involves incomparabilities in $L_\rho(X) - M_\rho$. In Section 6, we studied the relationship between non-measurability and incomparability. When there is no cognitive bound, our theory gives a complete preference relation over $L_\infty(Y)$, enjoying the expected utility hypothesis. However, our main concern is still the bounded case $\rho < \infty$.

In Section 8, we applied the incomparability results to the Allais paradox, specifically, to an experimental example in Kahneman-Tversky [15]. We showed that the prediction of our theory is compatible with their experimental result; incomparabilities involved for $\rho = 2$ are crucial in interpreting their result.

We have succeeded in considering aspects of bounded rationality in terms of probability grids and cognitive bounds for EU theory. Although our theory allows us to consider cases from very shallow depths to the case of no cognitive bounds, the aspects of bounded rationality are more suitably seen with shallow ρ. When, however, we consider a specific decision problem, other aspects of bounded rationality may manifest themselves. We should have more researches on the aspects of bounded rationality in various directions. Here, we give a few possible research agenda.

The first three are related to bounded rationality.

[a] **Constructive method of particular preferences:** We presented our theory following Table 1.1 to derive all the preferences in a layer from the previous layer. However, the decision maker may think about his preferences more locally focussing only on the target lotteries and involved pure alternatives and relevant probabilities. This question could enable us to think about complexities of preference formation. It may give a better understanding of how much bounded rationalities are involved when only target lotteries are concerned.

[b] **Preference formation in inductive game theory (IGT)**: This theory studies experiential sources for individual knowledge/belief about the structure of the society (cf. Kaneko-Matsui [18]). Our approach has some parallelism to the constructive approach to IGT, due to Kline *et.al* [20]. In particular, Kaneko-Kline [17] studies the other person's preferences from experiences of the other's position through role-switching. However, since it is assumed that experiences include numerical utility values, their treatment does not capture the partial understanding/non-understanding of the other's preferences/desires. Perhaps, lack of full experiences is closely related to incomparability in our theory.

This is also related to the case-based decision theory by Gilboa-Schmeidler [12] as well as to the frequentist interpretation of probability in the context of the EU theory (cf., Hu [14]). The former concerns evaluations of probabilities for causality (course-effect) from experiences, and the latter is about probability as frequency of an event. Bounded memory capacity of a person is relevant for both. Our theory with probability grids and cognitive bounds may give a suggestion to analyze such problems.

[c]**: Behavior under incomparability**: When two lotteries are incomparable, our theory is silent about a choice by the decision maker. In Section 8, we adopted postulate BH for choices by subjects for incomparable lotteries c and d. These lotteries have different depths, i.e., $\delta(c) = 1$ and $\delta(d) = 2$. BH did not directly take depths into account. A different postulate should take depths into account. Then, we may discuss "common ratio effect" (Remark 8.1) in a more direct manner and possibly Ellesberg's paradox, too. This remains an open problem.

The other three comments are on possible generalizations of our theory.

[d]**: Extensions of choices of benchmarks**: In this paper, the benchmarks $\overline{y}$ and $\underline{y}$ are fixed. The choice of the lower $\underline{y}$ could be natural, for example, the status quo. The choice of $\overline{y}$ may be more temporary in nature. In general, there could be different benchmarks than the given ones. We could consider two possible extensions of choices of the benchmarks.

One is a *vertical extension*: we take another pair of benchmarks $\overline{\overline{y}}$ and $\underline{\underline{y}}$ such as $\overline{\overline{y}} \trianglerighteq_0 \overline{y} \trianglerighteq_0 \underline{y} \trianglerighteq_0 \underline{\underline{y}}$. The new set of pure alternatives is given as $X(\overline{\overline{y}}; \underline{\underline{y}})$. The relation between the original system and the new system is not simple. In the case of measurement of temperatures, the grids for the Celsius system do not exactly correspond to those in the Fahrenheit system. We may need multiple bases ℓ for probability grids, and may have multiple preference systems even for similar target problems.

The other extension is *horizontal*: For example, $\underline{y}$ is the present status quo for a student facing a choice problem between the alternative $\overline{y}$ of going to work for a large company and the alternative $\overline{\overline{y}}$ of going to graduate school. He may not be able to make a comparison between $\overline{y}$ and $\overline{\overline{y}}$, while he can make a comparison between detailed choices after the choice of $\overline{y}$ or $\overline{\overline{y}}$. This involves incomparabilities different from those considered in this paper. These possible extensions are open problems of importance.

[e]**: Extensions of the probability grids** Π_ρ: The above extensions may require more subtle treatments of probability grids. A possibility is to extend Π_ρ to $\cup_{\ell=2}^{\overline{\ell}}\Pi_\ell$; that is, probability grids having the denominators $\ell \leq \overline{\ell}$ are permissible. Then, the Celsius and Fahrenheit systems of measuring temperatures are converted from each to the other. A question is how large $\overline{\ell}$ is required for such classes of problems.

[f]**: Subjective probability**: Our theory is almost directly applied to Anscombe-Aumann's [1] theory of subjective probability and subjective utility. An event E such as tomorrow's weather is evaluated asking an essentially the same question as (1) in Section 1. We could have an extension of our theory including the subjective probability theory. It could be difficult to have an extension corresponding to Savage [27], since no benchmark scale is assumed; perhaps, Savage's theory is not in our scope.

Thinking about these problems and extensions makes more progress on our expected utility theory with probability grids and preference formation.

10 Appendix

We prepare the extension δ^* of the depth measure δ to $\Pi_k^* = \{\pi : \pi = \nu/\ell^k$ for some nonnegative integer $\nu\}$, where ν may be larger than ℓ^k. For $\pi \in \Pi_k^*$, we define $\delta^*(\pi) = k^*$ iff $k^* \in \Pi_{k^*}^* - \Pi_{k^*}^*$. Then, $\delta^*(\pi) = \delta(\pi)$ if $\pi \in \Pi_k$. The following facts will be used in the proof of Lemma 2.1:

$$\text{if } \pi = s + \nu'/\ell^k \text{ for an integer } s \text{ and } \nu' < \ell^k, \text{ then } \delta^*(\pi) = \delta^*(\nu'/\ell^k); \tag{48}$$

$$\text{if } \delta^*(\pi), \delta^*(\pi') \leq k, \text{then } \delta(\pi + \pi') \leq k. \tag{49}$$

Proof of Lemma 2.1.[10] Let $k \geq 1$. We show that if $f \in L_k(X)$, then $f = \widehat{e} * \widehat{f}$ for some $\widehat{f} \in L_{k-1}(X)^\ell$ with the depth constraint $\delta(f(x)) > \delta(f_t(x))$ for all $t \leq \ell$ and $x \in X$ with $\delta(f(x)) > 0$. We assume $\delta(f) = k$. Let $\{x_1, ..., x_m\}$ be the support of f with $f(x_t) > 0$ for $t = 1, ..., m$ and $m \geq 2$.

Notice that f can be regarded as the list $(x_1, \nu_1/\ell^k, ..., x_m, \nu_m/\ell^k)$ for some $x_t \in X$ and $0 \leq \nu_t < \ell^k$ for all $t \leq m$ with $\sum_{t=1}^m \nu_t = \ell^k$. This is expressed as:

$$f = (\overbrace{x_1, 1/\ell^k, ..., x_1, 1/\ell^k}^{\nu_1 \text{ times}},\ ...,\ \overbrace{x_m, 1/\ell^k, ..., x_m, 1/\ell^k}^{\nu_m \text{ times}}), \tag{50}$$

i.e., each x_t occurs ν_t times with the same weight $1/\ell^k$. Since $\ell^k = \ell \times \ell^{k-1}$, the list $[x_1, ..., x_1, ..., x_m, ..., x_m]$ of the length ℓ^k can be rewritten as the concatenation of ℓ sublists of length ℓ^{k-1} :

$$[[y_1^1, ..., y_{\ell^{k-1}}^1], ..., [y_1^\ell, ..., y_{\ell^{k-1}}^\ell]]. \tag{51}$$

[10] The author is indebted to a referee for this proof, which was much shorter than the original proof.

Associating weight $1/\ell^{k-1}$ to each element, we regard these as lotteries $f_1, ..., f_\ell$ in $L_{k-1}(X)$:

$$f_1 = [y_1^1, 1/\ell^{k-1}, ..., y_{\ell^{k-1}}^1, 1/\ell^{k-1}], \quad ..., \quad f_\ell = [y_1^\ell, 1/\ell^{k-1}, ..., y_{\ell^{k-1}}^\ell, 1/\ell^{k-1}]. \tag{52}$$

Then, it holds that $\widehat{e} * (f_1, ..., f_\ell) = f$; thus, $f \in L_k(X)$ decomposes into $(f_1, ..., f_\ell) \in L_{k-1}(X)^\ell$.

To show the depth constraint (11), we need a few concepts; first, we denote the list $[x_1, ..., x_1, x_2, ..., x_2, ..., x_m, ..., x_m]$ as $Z = [z_\xi : 1 \leq \xi \leq \ell^k]$ and $\overline{\nu}_t = \sum_{s \leq t} \nu_s$ for $t = 1, ..., m$. Then, $f(x_t)$ is regarded as a segment of Z, given as

$$[z_\xi : \overline{\nu}_{t-1} < \xi \leq \overline{\nu}_t] \text{ with weight } 1/\ell^k. \tag{53}$$

This corresponds to the fragment with x_t in (50). Thus, Z is partitioned in two ways: $[[y_1^1, ..., y_{\ell^{k-1}}^1], ..., [y_1^\ell, ..., y_{\ell^{k-1}}^\ell]]$ of (51) and $[[z_\xi : \overline{\nu}_{t-1} < \xi \leq \overline{\nu}_t] : t = 1, ..., m]$ of (53). The fragments of these partitions may have three types of (nonempty) intersections:

(1): $[y_1^{t'}, ..., y_{\ell^{k-1}}^{t'}]$ is a subfragment of $[z_\xi : \overline{\nu}_{t-1} < \xi \leq \overline{\nu}_t]$;

(2): $[y_1^{t'}, ..., y_{\ell^{k-1}}^{t'}]$ is a superfragment of $[z_\xi : \overline{\nu}_{t-1} < \xi \leq \overline{\nu}_t]$;

(3): one of them starts in the other but continues to the next fragment.

In (1), $f_{t'}(x_t) = 1$. Hence, $\delta(f_{t'}(x_t)) = 0 < \delta(f(x_t))$. In (2), $f_{t'}(x_t) = \nu_t/\ell^{k-1}$ and $f(x_t) = \nu_t/\ell^k$; so, $\delta(\nu_t/\ell^{k-1}) < \delta(\nu_t/\ell^k)$, i.e., $\delta(f_{t'}(x_t)) < \delta(f(x_t))$.

Consider (3). Suppose that $[z_\xi : \overline{\nu}_{t-1} < \xi \leq \overline{\nu}_t]$ ends in $[y_1^{t'}, ..., y_{\ell^{k-1}}^{t'}]$ but it starts in a previous fragment. Then, $f_{t'}(x_t) = \nu'_t/\ell^{k-1}$ for some ν'_t with $0 < \nu'_t < \ell^{k-1}$. In fact, $\sum_{s \leq t} \nu_s = \nu'_t + (t'-1)\ell^{k-1}$. By (48) and (49),

$$\delta^*(\nu_t/\ell^{k-1}) \geq \delta^*(\textstyle\sum_{s \leq t} \nu_s/\ell^{k-1}) = \delta^*(\nu'_t/\ell^{k-1} + (t'-1)) = \delta^*(\nu'_t/\ell^{k-1}).$$

Hence, $\delta(f_{t'}(x_t)) = \delta(\nu'_t/\ell^{k-1}) \leq \delta(\nu_t/\ell^{k-1}) < \delta(\nu_t/\ell^k) = \delta(f(x_t))$. The other case of (3) where $[z_\xi : \overline{\nu}_{t-1} < \xi \leq \overline{\nu}_t]$ starts in $[y_1^{t'}, ..., y_{\ell^{k-1}}^{t'}]$ but ends in a later fragment is similar.■

References

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

[2] Augustin, A., F. P. A. Coolen, G. de Cooman, M. C. T. Troffaes, (2014), *Introduction to Imprecise Probabilities*, Wiley, West Sussex.

[3] Aumann, R. J., (1962), Utility theory without the completeness axiom, *Econometrica* 30, 445-462.

[4] Blume, L. D. Easley, and J. Y. Halpern, (2013), Constructive Decision Theory, Working Paper.

[5] Davis, M., and M. Maschler, (1965), The Kernel of a Cooperative Game, *Naval Research Logistics Quarterly* 12, 223-259.

[6] Dubra, J., and E. A. Ok, (2002), A Model of Procedural Decision Making in the Presence of Risk, *International Economic Review* 43, 1053-1080.

[7] Dubra, J., F. Maccheroni, and E. A. Ok, (2004), Expected Utility Theory without the Completeness Axiom, *Journal of Economic Theory* 115, 118-133.

[8] Fishburn, P., (1970), Intransitive indifference with unequal indifference intervals, *J. Math. Psychology* 7, 144-140.

[9] Fishburn, P., (1971), One-way expected utility with finite consequence space, *Annals of Mathematical Statistics 42, 572-577.*

[10] Fishburn, P., (1972), Alternative axiomatizations of one-way expected utility, *Annals of Mathematical Statistics 43, 1648-1651.*

[11] Fishburn, P., (1982), *The Foundations of Expected Utility*, D. Reidel Publishing Co. London.

[12] Gilboa, I., and D. Schmeidler (2001), *A theory of case-based decisions*, Cambridge University Press.

[13] Herstein, I. B. and J. Minor (1953), An Axiomatic Approach to Measurable Utility, *Econometrica* 21, 291-297.

[14] Hu, T., (2013), Expected Utility Theory from the Frequentist Perspective, *Economic Theory* 53, 9–25.

[15] Kahneman, D., and A. Tversky, (1979), Prospect Theory: An Analysis of Decision under Risk, *Econometrica* 47, 263-292.

[16] Kaneko, M., (1992), The Ordered Field Property and a Finite Algorithm for the Nash Bargaining Solution, *International Journal of Game Theory* 20, 227–236.

[17] Kaneko, M., and J. J. Kline, (2015), Understanding the Other through Social Roles, *International Game Theory Review* 17, No.1, 1540005-1-31, DOI:10.1142/S0219198915400058.

[18] Kaneko, M., and A. Matsui, (1999), Inductive Game Theory: Discrimination and Prejudices, *Journal of Public Economic Theory* 1, 101-137.

[19] Kaneko, M., and K. Nakamura, (1979), The Nash social welfare function, *Econometrica* 47, 423–435.

[20] Kline, J. J., T. Lavendhomme, and S. Waltener, (2018), From Memories to Inductively Derived Views: a Constructive Approach. To appear in *Economic Theory.*

[21] Kontek, K. and M. Lewandowski, (2017), Range-Dependent Utility, *Management Science* 33, 1367 -1382.

[22] van de Kuilen, G., and P. Wakker, (2006), Learning in the Allais Paradox, *Journal of Risk and Uncertainty* 33, 155-164.

[23] Mendelson, E., (1973), *Numbers Systems and Foundations of Analysis*, Krieger, Malabar.

[24] Nash, J. F., (1950), The Bargaining Problem, *Econometrica* 18, 155-162.

[25] Prelec, D. (1998), The Probability Weighting Function, *Econometrica* 60, 497-528.

[26] Rubinstein, A., (1988), Similarity and Decision Making under Risk: Is there a Utility Theory Resolution to the Allais Paradox, *Journal of Economic Theory* 46, 145-153.

[27] Savage, L. J., (1954), *The Foundations of Statistics*, John Wiley and Sons, New York.

[28] Shafer, G., (1986), Savage revisited, *Statistical Science* 1, 463-501.

[29] Shafer, G., (2015), Constructive Decision Theory, *International Journal of Approximate Reasoning* 79, 45-62.

[30] Simon, H. A., (1956), Rational Choice and the Structure of the Environment, *Psychological Review* 63, 129–138.

[31] Simon, H. A., (1983), *Reason in Human Affairs*, Stanford University Press, Stanford.

[32] Turing, A., (1936), On Computable Numbers, with an Application to the Entscheidungsproblem, *Proceedings of London Mathematical Society* 12, 230-265.

[33] Von Neumann, J., and O. Morgenstern, (1944), *Theory of Games and Economic Behavior*, 2nd ed. 1947, Princeton University Press, Princeton, N.J.