WINPEC Working Paper Series No.E1610
Revised, May 2018

Expected Utility Theory with Probability Grids and Incomparabilities

Mamoru Kaneko

Waseda INstitute of Political EConomy
Waseda University
Tokyo,Japan

Expected Utility Theory with Probability Grids and Incomparabilities*†

Mamoru Kaneko‡

May 31, 2018

Abstract

We reformulate expected utility theory, from the viewpoint of bounded rationality, by introducing probability grids and a cognitive bound; that is, we restrict permissible probabilities only to decimal (binary) 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, taking the form of mathematical induction with the measurements of utilities as the induction base and their extensions as the induction step. When a cognitive bound is small, the preference relation involves many incomparabilities, but, as the cognitive bound becomes less restrictive, there are less incomparabilities. In the case of no cognitive bound, our theory is considered classical expected utility theory. The main part of this paper is a study of incomparabilities involved in the preference relation. We give a complete characterization of incomparabilties, and a representation theorem in terms of a two-dimensional vector-valued utility function. 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

1.1 Motivations

We reconsider EU theory primarily from the viewpoint of bounded rationality. We restrict permissible probabilities in EU theory 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 $\Pi_\rho = \Pi_k = \{\frac{0}{10^k}, \frac{1}{10^k}, ..., \frac{10^k}{10^k}\}$. This restriction enables our theory to be constructive and allows us to study preference incomparabilities. When there is no cognitive bound, i.e., $\rho = \infty$, our theory, restricted to the set of exactly measured pure alternatives, gives complete preferences

*This is an entirely new version of the working paper, No.E1610, Waseda Institute of Political Economy.

†The author thanks J. J. Kline, P. Wakker, M. Lewandowski, L. Tang, A. Dominiak, S. Shiba, M. Cohen, and Y. Rebille for helpful comments on earlier versions of this paper. 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)

over the set of lotteries, and the theory is still a proper fragment of classical EU theory. We first disentangle the new components of our theory.

Although Simon's [23] original concept of bounded rationality meant a relaxation of simple utility maximization, the concept itself can be interpreted in many ways such as bounded logical inference or bounded perception ability (cf., Rubinstein [19]). In EU theory, two types of mathematical components are involved; object-components taken by a decision maker and mata-components used by the outside analyst. The former should be primary targets in EU theory, and the latter such as highly complex rational as well as irrational probabilities are added for analytic convenience. This addition leads to Simon's [24] critique of EU theory as a description of super rationality.[1] For example, probability values of the size $\frac{t}{10^2}$ ($t = 0, ..., 10^2$) are quite accurate for ordinary people,[2] but classical EU theory starts with a full real number theory and makes no distinction between components for a decision maker and those for an outside analyst.

The separation of these mathematical components is a problem of degree. To capture this, we introduce the concepts of *probability grids* and *cognitive bound* ρ to EU theory. For a finite $\rho = k$, the set of probability grids (permissible probabilities) is given as $\Pi_k = \{\frac{0}{10^k}, \frac{1}{10^k}, ..., \frac{10^k}{10^k}\}$. The decision maker thinks about his evaluation of preferences with Π_k for a small k to a larger k up to a bound ρ. When there is no cognitive limitation, i.e., $\rho = \infty$, we define $\Pi_\infty = \cup_{k=0}^{\infty} \Pi_k$.

We describe a process of deriving preferences by the decision maker from shallow to deeper probability grids up to a cognitive bound. This approach shares motivations for "constructive decision theory" with Shafer [21], [22] and with Blume *et al.* [3]. These authors study Savage's [20] subjective utility/probability theory so as to introduce certain constructive features for decision making.[3] This paper formulates constructive decision making in an explicit manner while restricting its focus to EU theory with probability grids.

In a broad sense, our treatment of probabilities falls in the field called "imprecise probabilities/similarity" (cf., Augustin *et al.* [2], Rubinstein [18], Tserenjigmid [25]).[4] In our approach, however, each probability in Π_k itself is precise, and its discrete presence is not interpreted as representing "imprecise probabilities". If the probabilities in Π_k are not fine enough for the decision maker to make preference measurements, then he would go to finer probabilities in Π_{k+1} and may repeat the process up to his cognitive bound. Here, "imprecision" may be involved and caused in this process with the cognitive bound.

Now, we consider how to describe our constructive EU theory. Constructiveness needs starting preferences; we take a hint from von Neumann-Morgenstern [26]. They divided the motivating argument into the following two statements, although this separation was not reflected in their mathematical development:

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

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

We formulate these steps as mathematical induction: Step B is the inductive base measuring

[1] We take his critique applied broadly to expected utility theory, while it may refer directly to Savage's [20] theory.

[2] Recall that a significance level for statistical hypothesis testing is typically 5% or 1%.

[3] In [21], [22], when basic probability/utility schemes are given, they are extended to the expected utility form. In [3], propositional logic is adopted to describe decision making (with the use of real number probabilities) and mental states are described in a semantic manner. Either approach takes some constructive aspect, but a decision making process is not explicitly formulated in a constructive manner.

[4] In this literature, imprecision/similarity is given in the mind of a decision maker under the assumption that imprecision/similarity is defined over all real number probabilities.

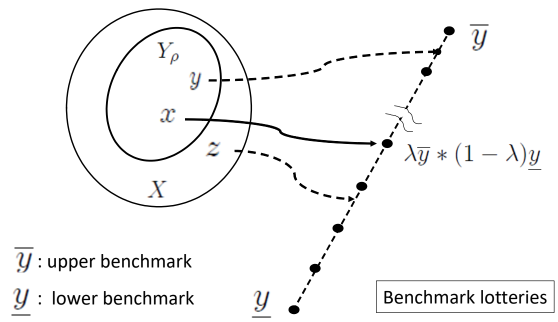

Figure 1: Measurement by the benchmarke scale

utilities from pure alternatives in terms of *upper and lower benchmarks* $\overline{y}$, $\underline{y}$ and permissible probability grids, and Step I is the induction step extending these measurements to lotteries with more risks.[5]

In Fig.1, Step B is depicted; x is measured by a rough scale, y needs a more precise scale, yet both x and y are exactly measured within a given bound ρ. However, z is not exactly measured.[6] In our approach, these measurements are described in terms of axioms. Then, Step I is formulated in a constructive manner from shallower to deeper layers of probability grids; here, it is described in terms of inference rules.

Since preferences are constructed piece by piece, completeness cannot be assumed either in Step B or Step I. As the probability grids are more precise, more preference comparisons become possible. Here, we weaken the standard "independence condition," which acts as a bridge between layers.

Consider 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 $y \sim [\overline{y}, \lambda; \underline{y}]$. Suppose $\lambda = \frac{83}{10^2} \in \Pi_2$. Consider another lottery $d = \frac{25}{10^2} y * \frac{75}{10^2}\underline{y}$. Step B is not applied to this since d includes the third alternative y. However, because of the indifference $y \sim [\overline{y}, \frac{83}{10^2}; \underline{y}]$, we substitute $[\overline{y}, \frac{83}{10^2}; \underline{y}]$ for y in d and calculate the resulting probabilities for the benchmarks $\overline{y}$ and $\underline{y}$:

$$d = \tfrac{25}{10^2} y * \tfrac{75}{10^2}\underline{y} \sim \tfrac{25}{10^2}[\overline{y}, \tfrac{83}{10^2}; \underline{y}] * \tfrac{75}{10^2}\underline{y} = \tfrac{2075}{10^4}\overline{y} * \tfrac{7925}{10^4}\underline{y}. \tag{1}$$

This is the result of the standard "independence condition". We observe two points here. One is a jump from the layer of depth 2 to that of depth 4, and the other is the involvement of quite precise probabilities. To have a meaningful cognitive limitation, we avoid the jump from depth 2 to 4. We formulate a weaker version of the "independence condition", avoiding such a jump, to make a connection between neighboring layers. Also, if 4 is regarded as too deep, we would take a cognitive bound as $\rho = 2$ or 3.

Keeping the above mentioned motivations in mind, we explain the development of our theory

[5] Our method is applied to the subjective probability theory due to Anscombe-Aumann [1]. See Section 8.

[6] This method is dual to with the measurement method in terms of certainty equivalent of a lottery (cf., Kontek-Lewandowski [14] and its references). In our method, the set of benchmark lotteries up to some cognitive bound forms a base scale while in the latter, the set of monetary payments is a base scale.

in Section 1.2.

1.2 Theoretical development

Since Step B may involve probability grids of various depths, the induction base may be scattered over layers. The preference relation $\succsim_{B,k}$appears as a part of the induction base at the layer k in PB in Table 1.1. The main part of induction is a bridge from $\succsim_{k-1}$ to the next $\succsim_k$; this bridge is a formulation of "independence condition". In Table 1.1, each $\succsim_k$ is derived from $\succsim_{k-1}$ and $\succsim_{B,k}$. We emphasize that this process is described from the viewpoint of the decision maker.

Table 1.1

PB: base relation	$\succsim_{B,0}$	$\subseteq$	$\succsim_{B,1}$	$\subseteq \dots \subseteq$	$\succsim_{B,k} = \succsim_{B,\rho}$
	$\downarrow$		$\downarrow$		$\downarrow$
PI: constructed relations	$\succsim_0$	$\rightarrow$	$\succsim_1$	$\rightarrow \dots \rightarrow$	$\succsim_k = \succsim_\rho$

When the cognitive bound ρ is a finite k, the *resultant preference relation* is given as the last $\succsim_\rho = \succsim_k$ and, when $\rho = \infty$, the resultant relation is given as $\succsim_\infty = \cup_{t=0}^{\infty} \succsim_k$. We show that in either case, $\succsim_\rho$ is a well-defined binary relation. In Section 4, we study the relationship of our theory to classical EU theory restricting the pure alternatives to those exactly measured in Step B. We show that the resultant relation $\succsim_\infty$ has an expected utility representation, which implies that $\succsim_\infty$ is complete. Our theory is still considered constructive. We provide a further extension of $\succsim_\infty$ to make a comparison with 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 [24].

The main part of this paper is a study of incomparabilties involved in $\succsim_\rho$ for a given cognitive bound $\rho < \infty$. The first step is to partition the set, $L_\rho(X)$, of all lotteries into the set M_ρ of *measurable lotteries* and the set of *non-measurable lotteries* $L_\rho(X) - M_\rho$. A measurable lottery f has an indifferent benchmark lottery $\lambda\overline{y} * (1-\lambda)\underline{y}$ for some $\lambda \in \Pi_\rho$. We can characterize a condition for a lottery f to belong to M_ρ, in terms of only ρ and the probabilities involved in f, which is given as Theorem 5.2. In the above example with $\overline{y}$, $\underline{y}$, y, When $\rho \geq 4$, the lottery $d = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$ is indifferent to $\frac{2075}{10^4} \overline{y} * \frac{7925}{10^4} \underline{y}$, which implies that d is measurable, but when $\rho = 2$ or 3, the second lottery is not permissible; d is non-measurable.

Preference incomparability involves non-measurable lotteries; incomparability occurs only if at least one lottery is non-measurable. In order to study incomparability, we consider the concepts of *lub* and *glb* (least upper and greatest lower bounds) for each lottery $f \in L_\rho(X)$; the *lub* of f is the least preferred benchmark lottery better than or indifferent to f, and the *glb* is parallelly defined. Theorem 6.1 completely characterizes the incomparability/comparability for two lotteries in terms of their *lub* and *glb*. Theorem 6.2 synthesizes the results of incomparability/comparability so that $\succsim_\rho$ is represented by the two-dimensional vector function consisting of the *lub* and *glb* of each lottery f endowed with the interval order due to Fishburn [6]. This theorem may be interpreted as what von Neumann-Morgenstern [26], p.29 indicated.[7]

[7] Dubra, *et al.* [5] developed a representation theorem 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 disappear when there are no cognitive bounds.

We apply the results on incomparabilities to the Allais paradox, specifically, to an experimental result from Kahneman-Tversky [11]. We show that the paradoxical results remains when the cognitive bound $\rho \geq 4$. However, when $\rho = 2$ or 3, the resultant preference relation $\succsim_\rho$ is compatible with their experimental result.

This paper is organized as follows: Section 2 explains the concept of probability grids. Section 3 gives base preference relations and formulates the derivation process. Section 4 compares our theory with classical EU theory. Section 5 discusses the measurable and nonmeasurable lotteries. Section 6 studies incomparabilties involved in $\succsim_\rho$ for $\rho < \infty$. In Section 7, we exemplify our theory with an experimental result in Kahneman-Tversky [11]. Section 8 concludes this paper with comments on further possible studies.

2 Lotteries with Probability Grids and Preferences

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

$$\Pi_k = \{\tfrac{\nu}{\ell^k} : \nu = 0, 1, ..., \ell^k\} \text{ for any finite } k \geq 0. \tag{2}$$

Here, $\Pi_1 = \{\frac{\nu}{\ell} : \nu = 0, ..., \ell\}$ is the basic set of probability grids, whereas $\Pi_0 = \{0, 1\}$. Each Π_k is a finite set, and let $\Pi_\infty := \cup_{t=0}^{\infty} \Pi_t$, which is countable. We use the standard arithmetic rules over Π_∞; sum and multiplication are needed;[8] for our analysis, we use these calculation rules but, for our axiomatic system itself, they are used in a restricted manner, which we mention in adequate places. 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 $\delta(\frac{20}{10^2}) = \delta(\frac{2}{10}) = 1$.

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

$$\text{either } \lambda > \lambda', \ \lambda = \lambda', \text{ or } \lambda < \lambda'. \tag{3}$$

This is equivalent to that $\geq$ is complete and anti-symmetric.

Now, we show that each element in Π_k is obtained by taking the weighted sums of elements in Π_{k-1} with the equal weights. This is basic for our induction method.

Lemma 2.1 (Decomposition of probabilities). $\Pi_k = \{\sum_{t=1}^{\ell} \frac{1}{\ell}\lambda_t : \lambda_1, ..., \lambda_\ell \in \Pi_{k-1}\}$ for any k $(1 \leq k < \infty)$.

Proof. The right-hand set is included in Π_k by (2). Consider the converse. Each $\lambda \in \Pi_k$ is expressed as $\lambda = 1$ or $\lambda = \sum_{t=1}^{k} \frac{\nu_t}{\ell^t}$ where $0 \leq \nu_t < \ell$ for $t = 1, ..., k$. In the case of $\lambda = 1$, let $\lambda_1 = ... = \lambda_\ell = 1 \in \Pi_{k-1}$, and $\lambda = \sum_{t=1}^{\ell} \frac{1}{\ell}\lambda_t$. Consider the second case. Let $\lambda_1 = ... = \lambda_{\nu_1} = 1$, $\lambda_{\nu_1+1} = \frac{\nu_2}{\ell^1} + ... + \frac{\nu_k}{\ell^{k-1}}$, and $\lambda_t = 0$ for $t = \nu_1 + 2, ..., \ell$. This definition is applied when $\nu_1 = 0$. These $\lambda_1, ..., \lambda_\ell$ belong to Π_{k-1} and $\lambda = \sum_{t=1}^{\ell} \frac{1}{\ell}\lambda_t$.$\blacksquare$

The union $\Pi_\infty = \cup_{k=0}^{\infty} \Pi_k$ is a proper subset of the set of all rational numbers. For example, when $\ell = 10$, Π_∞ has no recurring decimals, which are also rationals. We also note that Π_∞

[8] Π_k is a subset of $\{\sum_{t=k_1}^{k_2} \nu_t \cdot \ell^t : -\ell < \nu_t < \ell$ for t with $k_1 \leq t \leq k_2$ and k_1, k_2 are integers with $k_1 < 0 < k_2\}$, which is a ring, i.e., it is closed with respect to the three arithmetic operations $+, -,$ and $\times$. See Mendelson [15], p.95.

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 $\frac{1}{3}$. Nevertheless, Π_∞ is dense in $[0,1]$, which is crucial in comparing our theory with classical EU theory in Section 4.2.

Let X be a set of pure alternatives. For any $k < \infty$, we define $L_k(X)$ by

$$L_k(X) = \{f : f \text{ is a function from } X \text{ to } \Pi_k \text{ with } \textstyle\sum_{x\in X} f(x) = 1\}. \tag{4}$$

Since $\Pi_k \subseteq \Pi_{k+1}$, it holds that $L_k(X) \subseteq L_{k+1}(X)$. We denote $L_\infty(X) = \cup_{k=0}^{\infty} L_k(X)$. 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 will be relevant in Section 5.

We denote a *cognitive bound* by ρ, which is a natural number or infinity ∞. If $\rho = k < \infty$, then $L_\rho(X) = L_k(X)$, and if $\rho = \infty$, then $L_\rho(X) = L_\infty(X) = \cup_{k=0}^{\infty} L_k(X)$. The latter is interpreted as the case of no cognitive limitation.

Example 2.1. Let $X = \{\overline{y}, y, \underline{y}\}$ and $\rho = 2$. Since lottery $d = \frac{25}{10^2} y * \frac{75}{10^2} \underline{y}$ is in $L_2(X) - L_1(X)$, the depth $\delta(d) = 2$; but since $d' = \frac{20}{10^2} y * \frac{80}{10^2} \underline{y} = \frac{2}{10} y * \frac{8}{10} \underline{y} \in L_1(X)$, its depth is $\delta(d') = 1$.

Now, we formulate a connection from $L_{k-1}(X)$ to $L_k(X)$. We say that $\mathbf{f} = (f_1, ..., f_\ell)$ 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

$$f(x) = \textstyle\sum_{t=1}^{\ell} \frac{1}{\ell} \times f_t(x) \text{ for all } x \in X. \tag{5}$$

Let $\mathbf{e} = (\frac{1}{\ell}, ..., \frac{1}{\ell})$, and f is denoted by $\mathbf{e} * \mathbf{f}$ or $\sum_{t=1}^{\ell} \frac{1}{\ell} * f_t$. In other words, when $\mathbf{f} = (f_1, ..., f_\ell)$ is given, $f = \mathbf{e} * \mathbf{f}$ is a composite lottery and is reduced to a lottery in $L_k(X)$ by (5). In our axiomatic system, we use *reduction of a composite lottery* to a lottery only in this form. The next lemma connects $L_{k-1}(X)$ to $L_k(X)$, which facilitates our induction method described in Table 1.1.

Lemma 2.2 (Decomposition of lotteries). Let $1 \leq k < \infty$. Then, $L_k(X) = \{\mathbf{e} * \mathbf{f} : \mathbf{f} \in L_{k-1}(X)^\ell\}$.

The right-hand side of the assertion is the set of composed lotteries from $L_{k-1}(X)$, and is a subset of the left-hand side $L_k(X)$. The converse inclusion $\subseteq$ is essential and means that each lottery in $L_k(X)$ is decomposed to a weighted sum of some $(f_1, ..., f_\ell)$ in $L_{k-1}(X)^\ell$ with the equal weights. A proof is given in the Appendix, which involves precise construction of a decomposition $\mathbf{f} = (f_1, ..., f_\ell)$. This lemma is crucial in our theory.

A remark on the treatment of the domain X for $L_k(X)$ will be relevant in this paper. when we restrict X to a subset X', we define $L_k(X') := \{f \in L_k(X) : f(x) > 0 \text{ implies } x \in X'\}$. Hence, $L_k(X')$ is directly a subset of $L_k(X)$. Lemma 2.2 hods for these $L_k(X')$ and $L_{k-1}(X')$. Also, we call a subset S of X a *support* of $f \in L_k(X)$ iff $f(x) > 0$ implies $x \in S$. Let $f \in L_k(X)$ and S the support of f with $f(x) > 0$ for all $x \in S$. Applying Lemma 2.2 to f and $L_k(S)$, we have a decomposition $\mathbf{f} \in L_{k-1}(S)^\ell$.

The lottery $d = [y, \frac{25}{10^2}; \underline{y}]$ in Example 2.1 has two types of decompositions:

$$\tfrac{5}{10} * [y, \tfrac{5}{10}; \underline{y}] + \tfrac{5}{10} * \underline{y} \;\text{ and }\; \tfrac{2}{10} * y + \tfrac{1}{10} * [y, \tfrac{5}{10}; \underline{y}] + \tfrac{7}{10} * \underline{y}. \tag{6}$$

In the first, a decomposition $\mathbf{f} = (f_1, ..., f_{10})$ is given as $f_1 = ... = f_5 = [y, \frac{5}{10}; \underline{y}]$ and $f_6 = ... = f_{10} = \underline{y}$. In the second, $\mathbf{f}$ is given as $f_1 = f_2 = y$, $f_3 = [y, \frac{5}{10}; \underline{y}]$ and $f_4 = ... = f_{10} = \underline{y}$. We use these short-hand expressions rather than a full specification of $\mathbf{f} = (f_1, ..., f_{10})$. The proof of Lemma 2.2 constructs the second type of a decomposition in a general manner. In Section 3.2, we show that this multiplicity causes no problem in our inductive method.

We assume that a decomposition of $f \in L_k(X)$ takes the form of $\mathbf{f} = (f_1, ..., f_\ell)$ with $f = \sum_{t=1}^{\ell} \frac{1}{\ell} * f_t$. When $\ell = 10$, binary decompositions are not enough for Lemma 2.2, For example, consider the lottery $f = \frac{3}{10}\overline{y} * \frac{3}{10}y * \frac{4}{10}\underline{y}$ in the above example. This f is not expressed by a binary combination in $L_0(X) = X$ with weights in Π_1. Without requiring a decomposition be in the one-level lower layer, this could be possible such as $f = \frac{5}{10}(\frac{6}{10}\overline{y} * \frac{4}{10}y) * \frac{5}{10}(\frac{2}{10}y * \frac{8}{10}\underline{y})$, but the requirement for decompositions to be in one-layer lower is crucial for the constructive argument in our theory.

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

$$\begin{aligned} f \ &\succ \ g \iff f \succsim g \text{ and not } g \succsim f; \\ f \ &\sim \ g \iff f \succsim g \text{ and } g \succsim f; \\ f \ &\bowtie \ g \iff \text{neither } f \succsim g \text{ nor } g \succsim f. \end{aligned} \tag{7}$$

The incomparability relation $\bowtie$ is new and is studied in the subsequent sections. Nevertheless, all the axioms are about the relations $\succsim$, $\succ$, and $\sim$. The relation $\bowtie$ is defined as the residual part of $\succsim$. Although $\sim$ and $\bowtie$ are sometimes regarded as closely related (cf., Shafer [21], p.469), they are well separated in Theorem 6.2.

3 EU Theory with Probability Grids

Our theory has two parts: base preference relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ with four axioms, and a derivation process with three inference rules to derive preference relations $\langle \succsim_k \rangle_{k=0}^{\rho}$. The former describes the benchmark scales and measurements of pure alternatives in terms of benchmark scales. The latter describes derives preferences over lotteries with more risks. The derivation process generates a well-defined binary relation $\succsim_\rho$ uniquely relative to given base relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$. When a finite cognitive bound ρ is given, the process stops at ρ, and its resultant relation is $\succsim_\rho$. When $\rho = \infty$, the resultant relation is $\succsim_\infty = \cup_{k=0}^{\infty} \succsim_k$.

3.1 Base preference relations

The set of pure alternatives X contains two distinguished elements $\overline{y}$ and $\underline{y}$, which we call the *upper* and *lower benchmarks*. Let $k < \infty$. We call a lottery 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 $\mathbf{B}_k(\overline{y}; \underline{y}) := \{[\overline{y}, \lambda; \underline{y}] : \lambda \in \Pi_k\}$. In particular, $\mathbf{B}_0(\overline{y}; \underline{y}) = \{\overline{y}, \underline{y}\}$. The dots in Fig.1 express the benchmark lotteries. We define $\mathbf{B}_\infty(\overline{y}; \underline{y}) = \cup_{k=0}^{\infty} \mathbf{B}_k(\overline{y}; \underline{y})$.

Let $0 \leq k < \rho + 1$.[9] Let $\succsim_{B,k}$ be a subset of

$$D_k = \mathbf{B}_k(\overline{y}; \underline{y})^2 \cup \{(x, g), (g, x) : x \in X \text{ and } g \in \mathbf{B}_k(\overline{y}; \underline{y})\}. \tag{8}$$

[9]Stipulating $\infty + 1 = \infty$, we express the two statements "$k \leq \rho$ if $\rho < \infty$" and "$k < \rho$ if $\rho = \infty$" as "$k < \rho + 1$".

The first part is used to describe the preferences over the benchmark scale $\mathbf{B}_k(\overline{y};\underline{y})$. We call this the *scale part*, which is uniquely determined by Axiom B1. The second part is to describe the measurement of each pure alternative $x \in X$ in terms of the benchmark scale, which we call the *measurement part*. For example, if $(x,g) \in \succsim_{B,k}$ but $(g,x) \notin \succsim_{B,k}$, we have a strict preference $x \succ_{B,k} g$, and if $(x,g) \notin \succsim_{B,k}$ and $(g,x) \notin \succsim_{B,k}$, then x and g are incomparable. We allow $\succsim_{B,k}$ to be very partial.

We require all pure alternatives be between the upper and lower benchmarks $\overline{y}$ and $\underline{y}$.

Axiom B0 (Benchmarks): $\overline{y} \succ_{B,0} \underline{y}$ and $\overline{y} \succsim_{B,0} x \succsim_{B,0} \underline{y}$ for all $x \in X$.

The next states that $\mathbf{B}_k(\overline{y};\underline{y})$ is used as the base scale for depth k.

Axiom B1 (Benchmark scale): For $\lambda, \lambda' \in \Pi_k$, $\lambda \geq \lambda' \Longleftrightarrow [\overline{y},\lambda;\underline{y}] \succsim_{B,k} [\overline{y},\lambda';\underline{y}]$.

It follows from this axiom that for $\lambda, \lambda' \in \Pi_k$,

$$\lambda = \lambda' \Longleftrightarrow [\overline{y},\lambda;\underline{y}] \sim_{B,k} [\overline{y},\lambda';\underline{y}], \text{ and } \lambda > \lambda' \Longleftrightarrow [\overline{y},\lambda;\underline{y}] \succ_{B,k} [\overline{y},\lambda';\underline{y}]. \tag{9}$$

Thus, $\succsim_{B,k}$ is a complete relation over $\mathbf{B}_k(\overline{y};\underline{y})$ by (3). This is the scale part of $\succsim_{B,k}$, and is precise up to the cognitive bound ρ.

The measurement part of $\succsim_{B,k}$ is consistent with the scale part in the sense of no reversals with Axiom B1.

Axiom B2 (Non-reversal): For all $x \in X$ and $\lambda, \lambda' \in \Pi_k$, $[\overline{y},\lambda;\underline{y}] \succsim_{B,k} x$ and $x \succsim_{B,k} [\overline{y},\lambda';\underline{y}] \Longrightarrow \lambda \geq \lambda'$.

If we assume transitivity for $\succsim_{B,k}$ over D_k, B2 could be derived from B1, but we adopt B2 instead of transitivity, since this is more basic.

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

Axiom B3 (Preservation): For all $f, g \in D_k$, $f \succsim_{B,k} g \Longrightarrow f \succsim_{B,k+1} g$.

Axioms B2 and B3 are used only in the proof of Theorem 3.1 in this paper. Since Theorem 3.1 is the foundation of our theory, Axioms B2 and B3 are very basic.

The above axioms still allow great freedom for base preference relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$. To see this fact, we consider the following examples.

Example 3.1. Let $X = \{\overline{y}, y, \underline{y}\}$. Consider two examples for $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ satisfying Axioms B0 to B3. For $k = 0$, the scale part is $\{(\overline{y},\underline{y})\}$ and the measurement part may be given as $\{(\overline{y},y),(y,\underline{y})\}$, meaning $\overline{y} \succ_{B,0} y$ and $y \succ_{B,0} \underline{y}$. In Fig.2, $\overline{y}$ and $\underline{y}$ are located at the points of probability 1 and 0 at $k = 0$, and y is between them.

Let $k \geq 1$. In Fig.2, the thin solid lines give the upper and lower bounds for y: In both (*1*) and (*2*), y is strictly between $[\overline{y},\frac{9}{10};\underline{y}]$ and $[\overline{y},\frac{7}{10};\underline{y}]$ at $k = 1$. The scale part is: for $\nu, \nu' = 0, ..., 10$, $[\overline{y},\frac{\nu}{10};\underline{y}] \succsim_{B,1} [\overline{y},\frac{\nu'}{10};\underline{y}] \Longleftrightarrow \nu \geq \nu'$. The measurement part is given as $[\overline{y},\frac{9}{10};\underline{y}] \succ_{B,1} y \succ_{B,1} [\overline{y},\frac{7}{10};\underline{y}]$. The solid lines merge at $k = 2$ in (*1*), meaning that y becomes indifferent to $[\overline{y},\frac{83}{10^2};\underline{y}]$. Axiom B3 implies that $[\overline{y},\frac{9}{10};\underline{y}] \succ_{B,2} y \succ_{B,2} [\overline{y},\frac{7}{10};\underline{y}]$ remain. The lines do not merge in (*2*) and y is between $[\overline{y},\frac{9}{10};\underline{y}]$ and $[\overline{y},\frac{8}{10};\underline{y}]$ for $k \geq 2$; there are no indifferent benchmark lotteries to y at any $k \geq 0$.

The decision maker reflects upon his mind to look for his preferences $\succsim_{B,k}$ by a thought experiment with the probability grids Π_k. From the viewpoint of "bounded rationality", we are

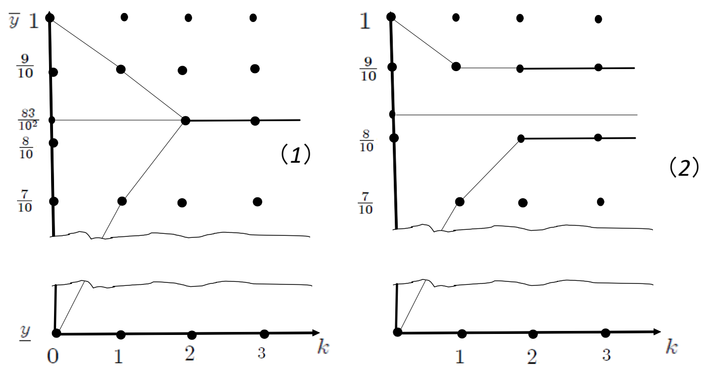

Figure 2: Base preference relations

inclined to have the view that he may stop his search for base preferences when he is happy enough for them (even though some reservations remain for him). In (*2*) of Example 3.1, he stops his search at depth 2.

3.2 Derivation process

Let a cognitive bound $\rho \leq \infty$ be given. We consider an extension from the base relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ to preferences over $L_k(X)$ for $k < \rho + 1$. We start with $\succsim_0$, and derive $\succsim_k$ from $\succsim_{k-1}$ and $\succsim_{B,k}$, provided that $\succsim_{k-1}$ is already defined.

The first inference rule is to derive preferences $\succsim_k$ from the basic preferences $\succsim_{B,k}$, which is the inductive base. In Table 1.1, the vertical arrows indicate to this derivation.

Inference Rule C0: For any $f, g \in D_k$, if $f \succsim_{B,k} g$, then $f \succsim_k g$, including the strict preference case.

Now, we consider the inductive step. Let $1 \leq k < \rho+1$. For $\mathbf{f} = (f_1, ..., f_\ell)$ and $\mathbf{g} = (g_1, ..., g_\ell)$, we write $\mathbf{f} \succsim_k \mathbf{g}$ iff $f_t \succsim_k g_t$ for all $t = 1, ..., \ell$. Recall that $\mathbf{f}$ is a decomposition of $f \in L_k(X)$ when $\mathbf{f} \in L_{k-1}(X)^\ell$ and $f = \mathbf{e} * \mathbf{f}$. The connection from $\succsim_{k-1}$ to $\succsim_k$ is given as follows:

Inference Rule C1: Let $f \in L_k(X),\ g \in \mathbf{B}_k(\overline{y}; \underline{y})$, or $f \in \mathbf{B}_k(\overline{y}; \underline{y}),\ g \in L_k(X)$. Suppose that $\mathbf{f}$, $\mathbf{g}$ are decompositions of f, g with $\mathbf{f} \succsim_{k-1} \mathbf{g}$. Then $f \succsim_k g$.
In addition, if $f_t \succ_{k-1} g_t$ for some $t = 1, ..., \ell$, then $f \succ_k g$.

In Table 1.1, the horizontal arrows indicate this derivation. Here, reduction of compound lotteries is used only for $f = \mathbf{e} * \mathbf{f}$ and $g = \mathbf{e} * \mathbf{g}$. It is a very weak version of the independence axiom since it connects one layer to the next only. In Section 4, we compare Inference Rule C1 with the independence condition in classical EU theory.

The last inference rule is transitivity: Let $1 \leq k < \rho + 1$.

Inference Rule C2 (Transitivity): For any $f, g, h \in L_k(X)$, if $f \succsim_k g$ and $g \succsim_k h$, then $f \succsim_k h$.

This suggests to take the transitive closure of preferences at layer k. This rule implies strict preference versions so that if at least one preference in the premise is replaced by a strict preference, then the conclusion is also strict.

Now, we have the derivation process for $\langle \succsim_k \rangle_{k=0}^{\rho}$.

Derivation process (abbreviated as DP) up to a cognitive bound ρ:

Step 0: For any $f, g \in L_0(X)$, if $f \succsim_{B,0} g$, then $f \succsim_0 g$ (including strict preferences[10]).

Step k $(1 \leq k < \rho + 1)$:

k0**:** for any $f, g \in L_k(X)$, if $f \succsim_{B,k} g$, then $f \succsim_k g$ (including strict preferences).

k1: $f \succsim_k g$ is derived by C1 (including strict preferences).

k2: $f \succsim_k g$ is derived by C2.

Smallest requirement (SM) for $\langle \succsim_k \rangle_{k=0}^{\rho}$: each $\succsim_k$ is obtained by a finite number of applications of Step 0 to k.

The smallest requirement means that each $\succsim_k$ is obtained only by repeating Step 0 - Step k $(k < \rho + 1)$ finitely many times. Although $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ is allowed to have some arbitrariness, $\langle \succsim_k \rangle_{k=0}^{\rho}$ is uniquely determined relative to a given $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$.

We have one basic problem to see whether each $\succsim_k$ is a binary relation over $L_k(X)$. The question is whether it is a subset of $L_k(X)^2$, i.e., for any $f, g \in L_k(X)$, exactly one of $(f, g) \in \succsim_k$ and $(f, g) \notin \succsim_k$ holds. In our system, the negation $(f, g) \notin \succsim_k$ may be derived as a part of strict preferences $g \succ_k f$. Therefore, it suffices to show that for any $f, g \in L_k(X)$,

$$f \succsim_k g \implies \text{not } (g \succ_k f). \tag{10}$$

Once this is proved, $\succsim_k$ is a well-defined binary relation over $L_k(X)$. A proof is given in the Appendix. As stated above, we have the uniqueness of a sequence $\langle \succsim_k \rangle_{k=0}^{\rho}$ by the Smallest requirement.

Theorem 3.1 (Well-definedness). The DP generates a unique sequence of binary relations $\langle \succsim_k \rangle_{k=0}^{\rho}$, provided that $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ is given.

We are interested in the *resultant preference relation*; recall that for $\rho < \infty$, it is the last relation $\succsim_\rho$ of $\langle \succsim_k \rangle_{k=0}^{\rho}$, and for $\rho = \infty$, $\succsim_\rho = \cup_{k=0}^{\infty}(\succsim_k)$.

By Theorem 3.1, we have the set-theoretical description of the generated sequence $\langle \succsim_k \rangle_{k=0}^{\rho}$:

$$\succsim_0 \;=\; \succsim_{B,0}; \text{ and } \succsim_k \;= [(\succsim_{k-1})^{\mathrm{C1}} \cup (\succsim_{B,k})]^{tr} \text{ for each } k \ (1 \leq k < \rho + 1), \tag{11}$$

where $(\succsim_{k-1})^{\mathrm{C1}}$ is the set of preferences derived from $\succsim_{k-1}$ by C1. Then we take the transitive closure of $(\succsim_{k-1})^{\mathrm{C1}} \cup (\succsim_{B,k})$; we denote, by $\succsim^{tr}$, the *transitive closure* of $\succsim$, defined by: $f \succsim^{tr} g \iff f = h_0 \succsim h_1 \succsim ... \succsim h_m = g$ for some $h_0, ..., h_m$.

Axioms B2 and B3 for $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ are basic for (the proof of) Theorem 3.1, but once $\langle \succsim_k \rangle_{k=0}^{\rho}$ is well defined, these are inherited by $\langle \succsim_k \rangle_{k=0}^{\rho}$ in the following manner. Thus, Axioms B2 and B3 themselves will not be referred in the following.

Lemma 3.1. Let $1 \leq k < \rho + 1$.

(1): For all $f \in L_k(X)$ and $\lambda, \lambda' \in \Pi_k$, $[\overline{y}, \lambda; \underline{y}] \succsim_k f$ and $f \succsim_k [\overline{y}, \lambda'; \underline{y}] \Longrightarrow \lambda \geq \lambda'$.

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

Proof.(1): Let $[\overline{y}, \lambda; \underline{y}] \succsim_k f$ and $f \succsim_k [\overline{y}, \lambda'; \underline{y}]$. By C2 (transitivity), we have $[\overline{y}, \lambda; \underline{y}] \succsim_k [\overline{y}, \lambda'; \underline{y}]$. If $\lambda < \lambda'$, then $[\overline{y}, \lambda'; \underline{y}] \succ_{B,k} [\overline{y}, \lambda; \underline{y}]$ by B1; by C0, $[\overline{y}, \lambda'; \underline{y}] \succ_k [\overline{y}, \lambda; \underline{y}]$, i.e., not

[10] That is, if $f \succ_{B,0} g$, then $f \succ_0 g$.

$[\overline{y}, \lambda; \underline{y}] \succsim_k [\overline{y}, \lambda; \underline{y}]$, a contradiction. Hence, $\lambda \geq \lambda'$.

(2): Let $f \in L_{k-1}(X)$ and $g \in \mathbf{B}_{k-1}(\overline{y}; \underline{y})$. 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 $f = \sum_{t=1}^{\ell} \frac{1}{\ell} * g_t$. Hence, $(f_1, ..., f_\ell)$ and $(g_1, ..., g_\ell)$ are decompositions of f and g. By C1, we have $f \succsim_k g$. In the case where $g \succsim_{k-1} f$, we prove $g \succsim_k f$ similarly.

Now, let $f, g \in L_{k-1}(X)$. By (11) for k, there are a finite sequence of lotteries in $h_0, h_1, ..., h_l$ in $L_{k-1}(X)$ such that $f = h_0 \succsim_{k-1} h_1 \succsim_{k-1} ... \succsim_{k-1} h_l = g$ and each $h_t \succsim_{k-1} h_{t+1}$ is derived by C1 and B0 to B3; thus, at least one of h_t, h_{t+1} belongs to $\mathbf{B}_{k-1}(\overline{y}; \underline{y})$. By the conclusion of the first paragraph, it holds that $h_t \succsim_k h_{t+1}$ for $t = 0, ..., l-1$. By C2, $f = h_0 \succsim_k h_l = g$.■

The following sets of pure alternatives play an important role in subsequent studies:

$$Y_k = \{x \in X : x \sim_{B,k} [\overline{y}, \lambda; \underline{y}] \text{ for some } \lambda \in \Pi_k\} \text{ for } k \ (0 \leq k < \rho + 1). \tag{12}$$

This is the set of pure alternatives exactly measured by the benchmark scale $\mathbf{B}_k(\overline{y}; \underline{y})$, $k < \rho+1$. Since $\Pi_{k-1} \subseteq \Pi_k$, we have, by B0, $\{\overline{y}, \underline{y}\} \subseteq Y_0 \subseteq Y_1 \subseteq ...$, but $Y_{k+1} - Y_k$ may be empty. In *(1)* of Example 3.1, $Y_0 = Y_1 = \{\overline{y}, \underline{y}\}$ and $Y_k = \{\overline{y}, y, \underline{y}\}$ for $k \geq 2$, and in *(2)* $Y_k = \{\overline{y}, \underline{y}\}$ for $k \geq 0$. We let $Y_\rho = \cup_{k=0}^{\rho} Y_k$; in particular, when $\rho = \infty$, we denote Y_ρ simply by Y. In Section 4, we make comparisons between our theory and classical EU theory, in which the union $Y = \cup_{k=0}^{\infty} Y_k$ plays a crucial role.

Lemma 3.2.(1): For each $y \in Y_k$, a probability $\lambda_y \in \Pi_k$ with $y \sim_{B,k} [\overline{y}, \lambda_y; \underline{y}]$ is unique. In particular, $\lambda_{\overline{y}} = 1$ and $\lambda_{\underline{y}} = 0$.

(2): The transitive closure $\succsim_{B,k}^{tr}$ of $\succsim_{B,k}$ is complete over $Y_k \cup \mathbf{B}_k(\overline{y}; \underline{y})$.

Proof. (1): Let $y \sim_{B,k} [\overline{y}, \lambda_y; \underline{y}] \sim_{B,k} [\overline{y}, \lambda'_y; \underline{y}]$. By (9), $\lambda_y = \lambda'_y$. The remaining follows from B0 and (9).

(2): When $f, g \in \mathbf{B}_k(\overline{y}; \underline{y})$, the assertion holds by (9). Consider $f = y \in Y_k$ and $g \in \mathbf{B}_k(\overline{y}; \underline{y})$. Then, by (12), $f = y \sim_{B,k} h$ for some $h = [\overline{y}, \lambda; \underline{y}] \in \mathbf{B}_k(\overline{y}; \underline{y})$. Then, by (9), $h \succsim_{B,k} g$ or $g \succsim_{B,k} h$. Thus, $f \succsim_{B,k}^{tr} g$ or $g \succsim_{B,k}^{tr} f$. The case where $f, g \in Y_k$ is similar.■

By Lemma 3.2, we define the utility function $u_o : Y_k \to \Pi_k$ by

$$u_o(y) = \lambda_y \text{ for all } y \in Y_k. \tag{13}$$

This u_o represents the relation $\succsim_{B,k}^{tr}$ over Y_k, that is, for any $x, y \in Y_k$, $u_o(x) \geq u_o(y) \iff x \succsim_{B,k}^{tr} y$. This $u_o(\cdot)$ is extended to $Y_k \cup \mathbf{B}_k(\overline{y}; \underline{y})$ by $u_o([\overline{y}, \lambda; \underline{y}]) = \lambda$, which fully represents $\succsim_{B,k}^{tr}$ over $Y_k \cup \mathbf{B}_k(\overline{y}; \underline{y})$. We discuss a further extension to $L_\infty(Y)$ in Section 4.

Here, we give one remark on base preference relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$; they were given before the derivation process of preferences $\langle \succsim_B \rangle_{k=0}^{\rho}$. In fact, each $\succsim_{B,k}$ of $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ can be formulated as a process along the derivation process, rather than the whole relations are given before the DP. In the beginning of Step k, in addition to the relation $\succsim_{B,k-1}$ made in Step $k-1$, the decision maker evaluates relevant pure alternatives $x \in X$, and then, he prepares (a relevant part of) $\succsim_{B,k}$ and goes to the DP. This process can be formulated in various forms; for example, the decision maker evaluates only the pure alternatives relevant for given lotteries $f, g \in L_k(X)$. We will give a brief discussion on relationships on this process to Simon's [23] satisficing/aspiration argument in Section 8.

4 Relationship to Classical EU Theory

Before our study of the behavior of $\succsim_\rho$ for $\rho < \infty$, we look at the relationship between our theory and classical EU theory (cf., Herstein-Milnor [9], Fishburn [7]). Our theory with $\rho = \infty$ allows the expected utility hypothesis over the lotteries in $L_\infty(Y)$. The main difference is that in our theory, permissible probabilities are from $\Pi_\infty = \cup_{k=0}^{\infty}\Pi_k$, while all probabilities from $[0,1]$ are allowed in classical theory. Our resultant relation $\succsim_\infty = \cup_{k=0}^{\infty} \succsim_k$ is uniquely extended to a relation in the sense of classical theory, but it involves non-constructive components.

4.1 Expected utility hypothesis in case with no cognitive bounds

Let $\langle \succsim_k \rangle_{k=0}^{\infty}$ be the preference relations constructed by the DP from base relations $\langle \succsim_{B,k} \rangle_{k=0}^{\infty}$. We restrict our attention to the set of pure alternatives $Y = \cup_{k=0}^{\infty} Y_k$, where Y_k is given in (12). Consider the following extension of the function u_o given by (13) to $L_\infty(Y)$ and the induced preference relation $\succsim_e$;

$$u_e(f) = \textstyle\sum_{y \in Y} f(y) u_o(y) \text{ for any } f \in L_\infty(Y); \tag{14}$$

$$\text{for any } f, g \in L_\infty(Y),\ f \succsim_e g \Longleftrightarrow u_e(f) \geq u_e(g). \tag{15}$$

The value $u_e(f)$ belongs to Π_∞ for each $f \in L_\infty(Y)$, since u_o is a function from Y to Π_∞.

We have the following equivalence between $\succsim_\infty$ and $\succsim_e$, which is proved below.

Theorem 4.1 (1): $f \sim_\infty [\overline{y}, u_e(f); \underline{y}]$ for all $f \in L_\infty(Y)$.

(2)(Expected utility hypothesis): For any $f, g \in L_\infty(Y)$, $f \succsim_\infty g \Longleftrightarrow f \succsim_e g$.

Thus, the resultant relation $\succsim_\infty$ is complete over $L_\infty(Y)$. The equivalence result (2) hods for $\succsim_\infty$ and $\succsim_e$. However, the restrictions on $L_k(Y)$ do not necessarily enjoy this equivalence; for a given $k < \infty$, the relation $\succsim_{e,k} := \succsim_e \cap L_k(Y)^2$ is already complete over $L_k(Y)$, but the relation $\succsim_k$ may not be complete over $L_k(Y)$.

Theorem 4.1.(2) is the expected utility hypothesis, though Inference Rule C1 is a weak version of independence with a specific form of reduction of compound lotteries, that is, (5). Assume the full form of reduction: for any $f, g \in L_\infty(Y)$ and $\lambda \in \Pi_\infty$, we define $\lambda f * (1-\lambda) g \in L_\infty(Y)$ by

$$(\lambda f * (1-\lambda) g)(x) = \lambda f(x) + (1-\lambda) g(x) \text{ for all } x \in X. \tag{16}$$

Then, it follows from Theorem 4.1.(2) that the preference relation $\succsim_\infty$ satisfies the full independence condition: for any $f, g \in L_\infty(Y)$ and $\lambda \in \Pi_\infty$ $(0 < \lambda)$,

ID1$_\infty$: $f \succ_\infty g \Longrightarrow \lambda f * (1-\lambda) h \succ_\infty \lambda g * (1-\lambda) h$;

ID2$_\infty$: $f \sim_\infty g \Longrightarrow \lambda f * (1-\lambda) h \sim_\infty \lambda g * (1-\lambda) h$.

Then, preferences are freely carried over from shallow layers to deeper layers. When $\rho < \infty$, the reduced compound lottery $\lambda f * (1-\lambda) h$ may go beyond ρ. The classical version of these conditions are discussed in Section 4.2.

Proof of Theorem 4.1: First, we show that (2) is an immediate consequence of (1). Let $f, g \in L_\infty(Y)$. By (1), for some k_o, $f \sim_k [\overline{y}, u_e(f); \underline{y}]$ and $g \sim_k [\overline{y}, u_e(g); \underline{y}]$ for all $k \geq k_o$. Since $\succsim_\infty = \cup_{k=0}^{\infty} \succsim_k$, we can take a $k \geq k_o$ so that $f \succsim_\infty g \Longleftrightarrow f \succsim_k g$. By B1, (15), and C2, $f \succsim_k g$ $\Longleftrightarrow [\overline{y}, u_e(f); \underline{y}] \succsim_k [\overline{y}, u_e(g); \underline{y}] \Longleftrightarrow u_e(f) \geq u_e(f) \Longleftrightarrow f \succsim_e g$. In sum, $f \succsim_\infty g \Longleftrightarrow f \succsim_e g$.

Now, let us prove (1). For $k < \infty$, let $L_k(Y_k) = \{f \in L_k(X) : f$ has a support S in $Y_k\}$, and

$$L_k^*(Y_k) = \{f \in L_k(Y_k) : \text{for } t = 0, ..., k,\ f(y) \in \Pi_{k-t} \text{ if } y \in Y_t - Y_{t-1}\}, \tag{17}$$

where $Y_{-1} = \emptyset$. We have the following two assertions:

$$\text{for any } k \geq 0, \text{ if } y \in Y_k - Y_{k-1} \text{ and } f \in L_k^*(Y_k), \text{ then } f(y) = 0 \text{ or } 1; \tag{18}$$

$$L_\infty(Y) = \cup_{k=0}^{\infty} L_k^*(Y_k). \tag{19}$$

First, we see (18); let $f \in L_k^*(Y_k)$. If $y \in Y_k - Y_{k-1}$, then $f(y) \in \Pi_0$, i.e., $f(y) = 0$ or 1. Next, let us see (19). The inclusion $\subseteq$ is essential. Let $f \in L_\infty(Y)$. This f has a finite support S, and there is a k' such that for each $y \in S$, $y \in Y_t - Y_{t-1}$ for some $t \leq k'$ and also $f(y) \in \Pi_{k'}$. Let $k = 2k'$. Then, for any $y \in S$, if $y \in Y_t - Y_{t-1}$, then $f(y) \in \Pi_{k'} \subseteq \Pi_{k-t}$. Thus, $f \in L_k^*(Y_k)$.

Now, we show by induction over $k = 0, ...$ that

$$f \sim_k [\overline{y}, u_e(f); \underline{y}] \ \text{ for all } f \in L_k^*(Y_k). \tag{20}$$

Once this is proved, we complete the proof of (1); indeed, taking any $f \in L_\infty(Y)$, by (19), we have $f \in L_k^*(Y_k)$ for some $k < \infty$, and thus, $f \sim_k [\overline{y}, u_e(f); \underline{y}]$ by (20).

Let us prove (20). Let $k = 0$. For any $y \in L_0^*(Y_0), y \sim_0 \overline{y}$ or $y \sim_0 \underline{y}$ by (12) for $k = 0$ and DP.(0), i.e., $y \sim_0 [\overline{y}, 1; \underline{y}]$ and $y \sim_0 [\overline{y}, 0; \underline{y}]$. Suppose that (20) holds for k. Take any $f \in L_{k+1}^*(Y_{k+1})$. If $f(y) = 1$ for $y \in Y_{k+1} - Y_k$, then $f = y \sim_{k+1} [\overline{y}, \lambda_y; \underline{y}]$ by (12), which implies $f = y \sim_{k+1} [\overline{y}, u_e(f); \underline{y}]$ because $\lambda_y = u_e(y)$.

Suppose that $f(y) = 0$ for all $y \in Y_{k+1} - Y_k$. Then, f has a support in Y_k. By Lemma 2.2, f has a decomposition $\mathbf{f} = (f_1, ..., f_l)$. Hence, if $y \in Y_t - Y_{t-1}$, then $f_l(y) = 0$ for each $l = 1, ..., \ell$, which means that each f_l belongs to $L_k^*(Y_k)$. By the induction hypothesis, there is a $\mathbf{g} = (g_1, ..., g_\ell)$ such that $\mathbf{f} \sim_k \mathbf{g}$ and $g_t = [\overline{y}, u_e(f_l); \underline{y}]$ for $l = 1, ..., \ell$. Applying C1, we have $f = \mathbf{e} * \mathbf{f} \sim_{k+1} \mathbf{e} * \mathbf{g}$, and this becomes $f \sim_{k+1} \sum_{t=1}^{\ell} \frac{1}{\ell} * g_t = \sum_{t=1}^{\ell} \frac{1}{\ell} * [\overline{y}, u_e(f_t); \underline{y}] = [\overline{y}, \sum_{t=1}^{\ell} \frac{1}{\ell} u_e(f_t); \underline{y}] = [\overline{y}, u_e(f); \underline{y}]$; the last equality follows from (15).$\blacksquare$

4.2 Extension to classical EU theory

Our theory with $\rho = \infty$ and $Y = \cup_{k=0}^{\infty} Y_k$ still differs from classical EU theory where all probabilities from $[0, 1]$ are permissible. In fact, our resultant relation $\succsim_\infty$ can be uniquely extended to a preference relation in the sense of classical EU theory with the set of pure alternatives Y. Nevertheless, the extension involves non-constructive components.

We first give a summary of classical EU theory. Let $L_{[0,1]}(Y) = \{f : f : Y \to [0, 1]$ is a lottery with a finite support $S \subseteq Y\}$. Here, a lottery f can take any real value in $[0, 1]$ but $\sum_{y \in S} f(y) = 1$ for some finite subset S of Y. Since $\Pi_k \subseteq [0, 1]$ for all $k < \infty$, $L_\infty(Y) = \cup_{k=0}^{\infty} L_k(Y)$ is a subset of $L_{[0,1]}(Y)$; the lotteries taking values in $[0, 1] - \Pi_\infty$ are newly included in $L_{[0,1]}(Y)$. The set $L_{[0,1]}(Y)$ is an uncountable set, but we show that $L_\infty(Y)$ is dense in $L_{[0,1]}(Y)$. We assume the reduction of compound lotteries: for any $f, g \in L_{[0,1]}(Y)$ and $\lambda \in [0, 1]$, $\lambda f * (1 - \lambda) g$ is defined by (16).

We adopt the following axiomatic system, which is one among various equivalent systems. Let $\succsim_E$ be a binary relation over $L_{[0,1]}(Y)$; and we assume NM0 to NM2 on $\succsim_E$.

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

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

The difference between **ID1 - ID2 and ID1$_\infty$ - ID2$_\infty$** is only the domains of lotteries and permissible probabilities. Under these axioms, there is a utility function $U_E : L_{[0,1]}(Y) \to R$ so that

$$\text{for any } f, g \in L_{[0,1]}(Y),\ f \succsim_E g \Longleftrightarrow U_E(f) \geq U_E(g), \tag{21}$$

$$U_E(f) = \Sigma_{y \in S}\ f(y)u_E(y) \text{ for each } f \in L_{[0,1]}(Y), \tag{22}$$

where S is a finite support of f and u_E is the restriction of U_E on Y. The converse holds; if $\succsim_E$ is given by (21) and (22), then, NM0 to NM2 hold for $\succsim_E$.

Now, we connect this theory to our EU theory with probability grids. Let $\succsim_\infty = \cup_{k=0}^{\infty} \succsim_k$ be the resultant relation generated by DP from base relations $\langle \succsim_{B,k} \rangle_{k=0}^{\infty}$. Then, we have a unique extension $\succsim_E$ of $\succsim_\infty$ having NM0 to NM2.

Theorem 4.2 (Unique extension). There is a unique binary relation $\succsim_E$ over $L_{[0,1]}(Y)$ such that for any $f, g \in L_\infty(Y)$, $f \succsim_\infty g \Longrightarrow f \succsim_E g$ and NM0 to NM2 hold for $\succsim_E$.

This theorem is proved by denseness of $L_\infty(Y)$ in $L_{[0,1]}(Y)$ and continuity of $U_E(\cdot)$ with respect to point-wise convergence. Denseness is a direct consequence from the denseness of $\Pi_\infty = \cup_{k=0}^{\infty} \Pi_k$ in $[0,1]$. Continuity of $U_E(\cdot)$ means that for any sequence $\{f^\nu\}$ in $L_\infty(Y)$, if $f^\nu(y) \to f(y)$ for each $y \in Y$, then $\lim_{\nu \to \infty} U_E(f^\nu) = U_E(f)$. The function U_E given by (22) is continuous.

The proofs of Lemma 4.1 and Theorem 4.2 may appear to be constructive following $\succsim_\infty$. Indeed, as long as $f \in L_\infty(Y)$, the involved probabilities in each f are described by a finite list of natural numbers. Thus, our theory is constructive up to $\succsim_\infty$, but the last extension step to $\succsim_E$ to $L_{[0,1]}(Y)$ 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.[11]

Lemma 4.1. $L_\infty(Y)$ is a dense subset of $L_{[0,1]}(Y)$.

Proof. It suffices to show that $L_\infty(Y)$ is dense in $L_{[0,1]}(Y)$. Take any $f \in L_{[0,1]}(Y)$. This f has a finite support $S = \{y_0, y_1, ..., y_m\}$ in Y. 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 $\nu \to \infty$.

For any natural number ν, let $z_{\nu,t} = \min\{\pi_t \in \Pi_\nu : \pi_t \geq f(y_t)\}$ for all $t = 0, ..., m$. Then, 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, $u_{\nu,t} \in \Pi_\nu$ for all $t \leq m-1$, and $1 \geq 1 - \sum_{t<m} z_{\nu,t} = u_{\nu,m}$. Since m is fixed, we can take some ν_o so that $\frac{m}{\ell^\nu} < \frac{\ell^\nu - 1}{\ell^\nu}$ for any $\nu \geq \nu_o$. For any $\nu \geq \nu_o$, $1 - \sum_{t<m} z_{\nu,t} \geq 1 - \frac{m}{\ell^\nu} > \frac{1}{\ell^\nu}$. Hence,

[11] We adopt Axiom NM1 avoiding the use of a topology. 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, the above NM1 allows to restrict it to $L_{[0,1]\cap Q}(Y)$, where Q is the set of rationals. In this case, the extension result given in Theorem 4.2 is regarded as approximately constructive.

$u_{\nu,m} \in \Pi_\nu$. Also, we have $\sum_{t=0}^{m} u_{\nu,t} = 1$. Now, 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 $L_\nu(Y)$. When $\nu \to +\infty$, $g^\nu(y) \to f(y)$ over Y.■

Proof of Theorem 4.2. The relation $\succsim_\infty$ is a binary relation over $L_\infty(Y)$. Recall the $u_e : L_\infty(Y) \to R$ was given by (14). We extend this to $U_E : L_{[0,1]}(Y) \to R$ by (22) with $u_E = u_o$ given in (13) and define $\succsim_E$ by (21). Then, $U_E(\cdot)$ coincides with u_e over $L_\infty(Y)$. Hence, $\succsim_E$ is an extension of $\succsim_\infty$. It is easy to see that $\succsim_E$ satisfies NM0-NM2.

Finally, we show that the extension $\succsim_E$ is unique. Suppose that $\succsim'_E$ is an extension of $\succsim_\infty$ and satisfies NM0 to NM2. Hence, there is a $U'_E : L_{[0,1]}(Y) \to R$ satisfying (21) and (22). As stated above, U'_E is continuous. Since $\succsim'_E$ is an extension of $\succsim_\infty$, it holds that for all $f, g \in L_\infty(Y)$,

$$u_e(f) \succsim_\infty u_e(g) \Longleftrightarrow U'_E(f) \geq U'_E(g).$$

By Lemma 4.1, there are sequences $\{f^\nu\}$ and $\{g^\nu\}$ in $L_\infty(Y)$ such that they converge to f and g. Then, $U_E(f) = \lim_{\nu\to\infty} u_e(f^\nu) \geq \lim_{\nu\to\infty} u_e(g^\nu) = U_E(g) \Longleftrightarrow U'_E(f) = \lim_{\nu\to\infty} U'_E(f^\nu) \geq \lim_{\nu\to\infty} U'_E(g^\nu) = U'_E(g)$. This means that $f \succsim_E g$ is determined uniquely by $U_E(\cdot)$.■

5 Measurable and Nonmeasurable Lotteries

Our main concern is the behavior of the resultant relation $\succsim_\rho$ for a finite cognitive bound ρ. In particular, we are interested in incomparabilities involved in $\succsim_\rho$, which we study in Section 6. Here, we prepare the concepts of measurable and nonmeasurable lotteries. Incomparabilities are closely related to nonmeasurable lotteries; when f and g are incomparable, at least one of them is nonmeasurable. In this section, ρ is allowed to be finite or infinite.

Let $\langle \succsim_k \rangle_{k=0}^{\rho}$ be the preference relations derived by DP from given base relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$. We define the domain M_k for $k < \rho + 1$ by

$$M_k = \{f \in L_k(X) : f \sim_k g \text{ for some } g = [\overline{y}, \lambda; \underline{y}] \in \mathbf{B}_k(\overline{y}; \underline{y})\}. \tag{23}$$

That is, $f \in M_k$ is exactly measured by the benchmark scale $\mathbf{B}_k(\overline{y}; \underline{y})$. This is a direct generalization of (11), which divides the set of pure alternatives X into the set of measurable pure alternatives Y_k and its complement $X - Y_k$. It holds that $Y_k \subseteq M_k$. When $k = 0$, we have $M_0 = Y_0$ since $\mathbf{B}_0(\overline{y}; \underline{y}) = \{\overline{y}, \underline{y}\}$.

We call $f \in M_k$ *measurable*, and $f \in L_k(X) - M_k$ *nonmeasurable*. By B1 and C2,

$$\text{for each } f \in M_k, \text{ the probability weight } \lambda \text{ with } f \sim_k [\overline{y}, \lambda; \underline{y}] \text{ is unique,} \tag{24}$$

which we denote by λ_f. It holds by Lemma 3.1.(2) that $M_k \subseteq M_{k+1}$ for all $k < \rho+1$. For $\rho = \infty$, let $M_\infty = \cup_{k=0}^{\infty} M_k$.

The following lemma is about a structure of M_k. For this lemma, it is used that $Y_0 = M_0$.

Lemma 5.1.(1): Let $k < \rho+1$. 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 $k < \rho + 1$.

Proof. 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 $0 < f(y) < 1$ for some $y \in Y_{k+1} - Y_k$ or $0 < f(y) \leq 1$ for some $y \in X - Y_{k+1}$. We denote this y by y_o. If $f(y_o) = 1$ and $y_o \in X - Y_{k+1}$, by (12), there is no $g \in \mathbf{B}_{k+1}(\overline{y}; \underline{y})$ 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, $f \notin \mathbf{B}_{k+1}(\overline{y}; \underline{y})$.

By $f \in M_{k+1}$, we have a $g \in \mathbf{B}_{k+1}(\overline{y}; \underline{y})$ with $f \sim_{k+1} g$. By (11), there are $h_0 = f, h_1, ..., h_m = g$ such that $(h_{t-1}, t_t) \in (\succsim_k)^{\text{C1}} \cup (\succsim_{B,k+1})$ for $t = 1, ..., m$. This implies that $h_0 = f, h_1, ..., h_m = g$ are all indifferent with respect to $\succsim_{k+1}$. Since $0 < f(y_o) < 1$ and $f \notin \mathbf{B}_{k+1}(\overline{y}; \underline{y})$, we have $(f, h_1) \notin (\succsim_{B,k+1})$ by (8), which implies $(f, h_1) \in (\succsim_k)^{\text{C1}}$. Since $h_1 \sim_{k+1} g$ and $h_1, g \in \mathbf{B}_{k+1}(\overline{y}; \underline{y})$, h_1 and g are identical by (9). Hence, $(f, g) \in (\succsim_k)^{\text{C1}}$. This implies that f and g have decompositions $\mathbf{f} = (f_1, ..., f_\ell)$ and $\mathbf{g} = (g_1, ..., g_\ell)$ such that $f_t \in L_k(Y_{k+1})$, $g_t \in \mathbf{B}_k(\overline{y}; \underline{y})$ for all $t = 1, ..., \ell$, and $\mathbf{f} \sim_k \mathbf{g}$. Since $0 < f(y_o) < 1$, there is some f_t among $f_1, ..., f_\ell$ such that $f_t(y_o) > 0$. Since $y_o \in Y_{k+1} - Y_k$ or $y_o \in X - Y_{k+1}$, it holds that $f_t \notin L_k(Y_k)$. On the other hand, since $g_t \in \mathbf{B}_k(\overline{y}; \underline{y})$, we have $f_t \in M_k$; by the induction hypothesis, we have $f_t \in M_k \subseteq L_k(Y_k)$, a contradiction. Hence, we have the assertion (1) for $k + 1$. This implies (2) for $k + 1$.■

It holds by Lemma 5.1.(2) and Theorem 4.1 that $M_\infty = L_\infty(Y)$. Conversely, we can restrict the statements of Theorem 4.1 to M_k. That is, the expected utility hypothesis holds for the measurable domain M_k. Recall that λ_f, $u_e(f)$, and $\succsim_e$ are defined, respectively, in (24), (14), and (15). An implication is that no incomparabilities are observed in M_k.

Theorem 5.1 (Case C: expected utility over M_k). For each $k < \rho + 1$,

(1): $\lambda_f = u_e(f)$ for any $f \in M_k$;

(2): for any $f, g \in M_k$, $f \succsim_k g \Longleftrightarrow f \succsim_e g$.

Proof. **(1)**: By Theorem 4.1, $f \sim_{k'} [\overline{y}, u_e(f); \underline{y}]$ for some k'. By Lemma 3.1, we can assume $k' \geq k$. Using Lemma 3.1 again, we have $[\overline{y}, \lambda_f; \underline{y}] \sim_{k'} f \sim_{k'} [\overline{y}, u_e(f); \underline{y}]$. Then, we have $\lambda_f = u_e(f)$ by (9).
(2): Let $f \succsim_k g$. By (1) of this theorem, $[\overline{y}, u_e(f); \underline{y}] \sim_k f \succsim_k g \sim_k [\overline{y}, u_e(g); \underline{y}]$. This is equivalent to $u_e(f) \geq u_e(g)$, which is further equivalent to $f \succsim_e g$. The converse is obtained by tracing back this argument.■

Thus, λ_f is the same as the expected utility value $u_e(f)$, and $\succsim_k$ coincides with the expected utility preferences $\succsim_e$ over the measurable domain M_k. Theorem 5.2 gives a condition for a lottery $f \in L_k(Y_k)$ to be in M_k, using only the depth data included in f; for each $y \in Y_k$, f has two types of depths, i.e., the measurement depth $\delta(\lambda_y)$ of y and the depth $\delta(f(y))$ of the value $f(y)$, and their sum should be smaller than or equal to k for $f \in M_k$. Theorem 5.1 implies that as long as $f, g \in L_k(Y_k)$ satisfy this condition, f and g are comparable by the expected utility preferences $\succsim_e$.

Theorem 5.2 (Measurability criterion): Let $k < \rho + 1$ and $f \in L_k(Y_k)$. Then,

$$f \in M_k \Longleftrightarrow \max\{\delta(\lambda_y) + \delta(f(y)) : f(y) > 0\} \leq k. \tag{25}$$

Proof. We prove (25) 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, (25) holds for $k = 0$. Now, suppose the induction hypothesis that (25) holds for k. We prove (25) for $k + 1$.

Consider $f \in L_{k+1}(Y_{k+1})$. Let $y \in Y_{k+1} - Y_k$ with $f(y) > 0$. By Lemma 5.1.(1), $f(y) = 1$, i.e., $f = y$ and $\delta(f(y)) = 0$. Since $y \in Y_{k+1} - Y_k$, we have $\delta(\lambda_y) = k+1$. Hence, $y \in M_{k+1} \Longleftrightarrow \delta(\lambda_y) = k+1$.

Now, let $f(y) = 0$ for any $y \in Y_{k+1} - Y_k$. Suppose $f \in M_{k+1}$. If $f \in M_k$, we have the right-hand side of (25) by the induction hypothesis. Hence, we can suppose $f \in M_{k+1} - M_k$. Since $f \in L_{k+1}(Y_{k+1})$ and $f(y) = 0$ for any $y \in Y_{k+1} - Y_k$, we have $0 < f(y) \leq 1$ for some $y \in Y_k$. If $f(y) = 1$ and $y \in Y_k$, then $f \in M_k$, a contradiction to $f \in M_{k+1} - M_k$. Hence, $0 < f(y) < 1$ for some $y \in Y_k$; thus, $f \notin \mathbf{B}_{k+1}(\overline{y}; \underline{y})$.

Then, since $f \in M_{k+1} - M_k$, we have $f \sim_{k+1} g$ for some $g \in \mathbf{B}_{k+1}(\overline{y}; \underline{y})$. By (11), there are $h_0 = f, h_1, ..., h_m = g$ such that $h_t \sim_{k+1} h_{t+1}$ and $(h_t, h_{k+1}) \in (\succsim_k)^{\mathrm{C1}} \cup (\succsim_{B,k+1})$ for all $t = 0, ..., m-1$. Since $0 < f(y) < 1$ for some $y \in Y_k$ and $f \notin \mathbf{B}_{k+1}(\overline{y}; \underline{y})$, we have $(f, h_1) = (h_0, h_1) \notin \succsim_{B,k+1}$; thus, $(f, h_1) \in (\succsim_k)^{\mathrm{C1}}$ and $h_1 \in \mathbf{B}_{k+1}(\overline{y}; \underline{y})$. By the indifference of $\sim_{k+1}$, we have $h_1 \sim_{k+1} g$; in fact, they are identical by B1. In sum, $(f, g) \in (\succsim_k)^{\mathrm{C1}}$.

It follows from $(f, g) \in (\succsim_k)^{\mathrm{C1}}$ that there are decompositions $\mathbf{f} = (f_1, ..., f_\ell)$ and $\mathbf{g} = (g_1, ..., g_\ell)$ of f and g so that $\mathbf{f} \sim_k \mathbf{g}$. For $t = 1, ..., \ell$, since $g_t \in \mathbf{B}_k(\overline{y}; \underline{y})$, we have $f_t \in M_k$. By the induction hypothesis, we have $\delta(\lambda_y) + \delta(f_t(y)) \leq k$ for all $y \in Y_k$ and $t = 1, ..., \ell$. Since $f = \mathbf{e} * \mathbf{f}$, it holds that $\delta(f(y)) \leq \max_t \delta(f_t(y)) + 1$ for all $y \in Y_k$. Since $\delta(f(y)) = 0$ for all $y \in Y_{k+1} - Y_k$, we have $\delta(\lambda_y) + \delta(f(y)) \leq k+1$ for all $y \in Y_{k+1}$ with $f(y) > 0$.

Conversely, let $\delta(\lambda_y) + \delta(f(y)) \leq k+1$ for all $y \in Y_{k+1}$ with $f(y) > 0$. 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, $f = y \in Y_{k+1} \subseteq M_{k+1}$. Suppose $0 < k^* \leq k+1$. Then, $f \in L_{k^*}(Y_{k+1})$. By Lemma 2.2, we have a decomposition $\mathbf{f}$ of f. Then, $f_t \in L_{k^*-1}(Y_{k+1})$ for $t = 1, ..., \ell$. Then, $\delta(f_t(y)) \leq k^* - 1$ for all $y \in Y_{k+1}$. Since $\delta(\lambda_y) + \delta(f(y)) \leq k+1$ for all $y \in Y_{k+1}$ with $f(y) > 0$, it holds for $t = 1, ..., \ell$ that $\delta(\lambda_y) + \delta(f_t(y)) \leq k^* - 1 \leq k$ for all $y \in Y_{k+1}$ with $f_t(y) > 0$. By the inductive hypothesis for k, we have $f_t \in M_k$ for $t = 1, ..., \ell$. Thus, we have $g_t \in \mathbf{B}_k(\overline{y}; y)$ with $f_t \sim_k g_t$ for $t = 1, ..., m$. By C1, we have $f = \mathbf{e} * \mathbf{f} \sim \mathbf{e} * \mathbf{g} \in \mathbf{B}_{k+1}(\overline{y}; y)$. This means $f \in M_{k+1}$.■

The assertion (25) is read in two ways. One is to fix a lottery $f \in L_k(Y_k)$ but to change k. Let $k_f = \max\{\delta(\lambda_y) + \delta(f(y)) : f(y) > 0\}$. Then, (25) is written as

$$f \in M_k \Longleftrightarrow k_f \leq k. \tag{26}$$

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 $k_f = \delta(\frac{83}{10^2}) + \delta(\frac{25}{10^2}) = 4$. Hence, (26) implies that $f \in M_k \Longleftrightarrow k \geq 4$. This means that any lottery f in $L_\infty(Y_\infty)$ becomes measurable when k is large enough. The other reading of (25) is to fix a k and to change f. There is an $f \in L_k(Y_k)$ such that $\delta(f(y)) = k$. As long as $\delta(\lambda_y) > 0$ for some $y \in Y_k$, there is an $f \in L_k(Y_k)$ such that $\delta(\lambda_y) + \delta(f(y)) > k$, i.e., $f \notin M_k$. Thus, nonmeasurable lotteries exist as long as $\{\overline{y}, \underline{y}\} \subsetneq Y_k$. When Y_k becomes constant after some k, the set of nonmeasurable lotteries $L_k(Y_k) - M_k$ does not grow after k.

We mention the following theorem for the indifferences $\sim_k$ over $L_k(X)$; the indifference relation $\sim_k$ occurs only in M_k and, in particular, reflexivity holds only in M_k. This is needed for Theorem 6.2.

Theorem 5.3. Let $k < \rho + 1 \leq \infty$ and $f, g \in L_k(X)$.

(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 $f \in M_k$.

Proof (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 C2, a contradiction. Hence, $f \nsim_k g$. Now, let $f, g \notin M_k$. Suppose $f \sim_k g$. By (11), there are $f = h_0, h_1, ..., h_m = g \in L_k(X)$ such that $h_l \sim_k h_{l+1}$ and at least one of h_l and h_{l+1} belongs to $\mathbf{B}_k(\overline{y}; \underline{y})$ for $l = 1, ..., m-1$. Since $h_0 = f \notin M_\rho$ and $h_1 \in \mathbf{B}_k(\overline{y}; \underline{y}) \subseteq M_\rho$, we have, by the first case, $h_0 \nsim_k h_1$, a contradiction. Hence, $f \nsim_k g$.

(2): The *if* part is by the definition of M_k, B1, and C2. The *only-if* part (contrapositive) follows from (1) of this theorem.■

6 Incomparabilities and their Characterization

Now, we study how the resultant preference relation $\succsim_\rho$ involves incomparabilities for $\rho < \infty$. We will give a comment on the case $\rho = \infty$ in the end of this section. We characterize incomparabilities by the concepts *lub* and *glb* of $f \in L_\rho(X)$. Unifying this characterization with the comparability result (Theorem 5.1), we obtain the representation of $\succsim_\rho$ over $L_\rho(X)$ in terms of the interval order introduced by Fishburn [6]. Throughout this section, we assume $\rho < \infty$.

First, we show the following lemma.

Lemma 6.1. For each $k \geq 0$, $\overline{y} \succsim_k f \succsim_k \underline{y}$ for any $f \in L_k(X)$.

Proof. We show the assertion by induction over $k \geq 0$. Let $f \in L_0(X) = X$. By B0 and $\succsim_0 = \succsim_{B,0}$, we have the assertion for $k = 0$. Suppose the induction hypothesis that $\overline{y} \succsim_k f \succsim_k \underline{y}$ for any $f \in L_k(X)$. Consider $f \in L_{k+1}(X)$. Then, by Lemma 2.2, there is a vector $\mathbf{f} = (f_1, ..., f_\ell) \in L_k(X)^\ell$ such that $f = \mathbf{e} * \mathbf{f}$. By the induction hypothesis, $\overline{y} \succsim_k f_t \succsim_k \underline{y}$ for any $t = 1, ..., \ell$. By Rule C1, we have $\overline{y} = \mathbf{e} * \overline{y} \succsim_{k+1} f = \mathbf{e} * \mathbf{f} \succsim_{k+1} \mathbf{e} * \underline{y} = \underline{y}$.■

Lemma 6.1 guarantees that every $f \in L_\rho(X)$ has upper and lower bounds in $\mathbf{B}_\rho(\overline{y}; \underline{y})$. We can define the *lub* $\overline{\lambda}_f$ and *glb* $\underline{\lambda}_f$ of each $f \in L_\rho(X)$ by

$$\begin{aligned} \overline{\lambda}_f &= \min\{\lambda : [\overline{y}, \lambda; \underline{y}] \in \mathbf{B}_\rho(\overline{y}; \underline{y}) \text{ with } [\overline{y}, \lambda; \underline{y}] \succsim_\rho f\}; \\ \underline{\lambda}_f &= \max\{\lambda : [\overline{y}, \lambda; \underline{y}] \in \mathbf{B}_\rho(\overline{y}; \underline{y}) \text{ with } f \succsim_\rho [\overline{y}, \lambda; \underline{y}]\}. \end{aligned} \tag{27}$$

In general, it holds that $\overline{\lambda}_f \geq \underline{\lambda}_f$. If $\overline{\lambda}_f$ and $\underline{\lambda}_f$ coincide, then f belongs to M_ρ and is exactly measured; and if they differ, then f belongs to $L_\rho(X) - M_\rho$. These observations are described as follows: for any $f \in L_\rho(X)$,

$$\overline{\lambda}_f = \underline{\lambda}_f = \lambda_f \iff f \in M_\rho; \text{ and } \overline{\lambda}_f > \underline{\lambda}_f \iff f \in L_\rho(X) - M_\rho. \tag{28}$$

We remark that the *lub* and *glb* are defined in terms of the weak relation $\succsim_\rho$ in (27), but when $f \in L_\rho(X) - M_\rho$, these are strict relation $\succ_\rho$ by Theorem 5.3.(1). Here, we show only the direction $\Longleftarrow$ of the second of (28). Let $f \in L_\rho(X) - M_\rho$. Then, $[\overline{y}, \overline{\lambda}_f; \underline{y}] \succ_\rho f \succ_\rho [\overline{y}, \underline{\lambda}_f; \underline{y}]$. By C2, we have $[\overline{y}, \overline{\lambda}_f; \underline{y}] \succ_\rho [\overline{y}, \underline{\lambda}_f; \underline{y}]$; so $\overline{\lambda}_f > \underline{\lambda}_f$ by (9).

Consider the following mutually exclusive and exhaustive cases:

C: $f, g \in M_\rho$; and

IC: $f \in L_\rho(X) - M_\rho$ or $g \in L_\rho(X) - M_\rho$.

In case **C**, comparability is shown in Theorem 5.1. Consider case **IC**. Let $f, g \in L_\rho(X) - M_\rho$. Then, we have the gaps, $\overline{\lambda}_f > \underline{\lambda}_f$ and $\overline{\lambda}_g > \underline{\lambda}_g$. When these gaps are separated, e.g.,

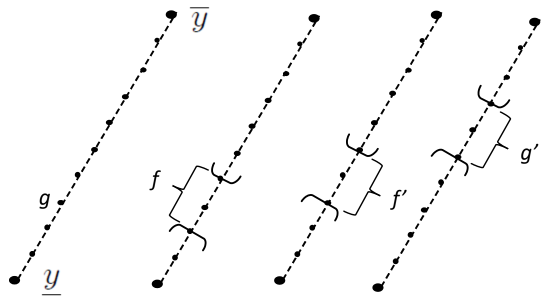

Figure 3: Incomparabilities

$\overline{\lambda}_f > \underline{\lambda}_f \geq \overline{\lambda}_g > \overline{\lambda}_g$, it holds that $f \succ_\rho g$; indeed, $f, g \in L_\rho(X) - M_\rho$ and $\underline{\lambda}_f \geq \overline{\lambda}_g$ imply $f \succ_\rho [\overline{y}, \underline{\lambda}_f; \underline{y}] \succsim_\rho [\overline{y}, \underline{\lambda}_g; \underline{y}] \succ_\rho g$ by Theorem 5.3.(1). In this case, f and g are comparable. Incomparability occurs in the other case where the gaps intersect, i.e., $\overline{\lambda}_f > \overline{\lambda}_g > \underline{\lambda}_f > \underline{\lambda}_g$. These comparisons become degenerate when $f \in L_\rho(X) - M_\rho$ and $g \in M_\rho$. In Fig.3, the second and fourth (first and fourth) lines show the separable cases, where comparability holds, and the second and third (first and second) lines show nonseparable cases, where incomparability holds.

The above observations hold in a general manner, which are stated in Theorem 6.1, where $f \bowtie_\rho g$ is defined in (7), meaning that f and g are incomparable with respect to $\succsim_\rho$.

Theorem 6.1 (Case IC). Suppose that at least one of f, g is in $L_\rho(X) - M_\rho$. Then,

$$f \succ_\rho g \Longleftrightarrow \underline{\lambda}_f \geq \overline{\lambda}_g; \tag{29}$$

$$f \bowtie_\rho g \Longleftrightarrow \overline{\lambda}_f > \underline{\lambda}_g \text{ and } \overline{\lambda}_g > \underline{\lambda}_f. \tag{30}$$

Proof. First, suppose that $f \notin M_\rho$ and $g \in M_\rho$. Then, $f \nsim_\rho g$ by Theorem 5.3.(1). Consider the equivalence of (29). If $f \succ_\rho g$, then $f \succ_\rho g \sim_\rho [\overline{y}, \lambda_g; \underline{y}]$, which implies $\underline{\lambda}_f \geq \lambda_g$ by (27) and B1. Conversely, if $\underline{\lambda}_f \geq \lambda_g$, then $f \succ_\rho [\overline{y}, \underline{\lambda}_f; \underline{y}] \succsim_\rho [\overline{y}, \lambda_g; \underline{y}] \sim_\rho g$ by (27) and B1, which implies $f \succ_\rho g$ by C2. The equivalence of (30) is similar. We can prove similarly these assertions in the case $f \in M_\rho$ and $g \notin M_\rho$.

Now, suppose that $f, g \notin L_\rho(X) - M_\rho$. Since $f \nsim_\rho g$ by Theorem 5.3.(1), we have $f \bowtie_\rho g \Longleftrightarrow$ neither $f \succ_\rho g$ nor $g \succ_\rho f$. Hence, (30) follows from (29). Now, we prove (29); specifically,

(i): $f \succ_\rho g \Longleftrightarrow f \succ_\rho h \succ_\rho g$ for some $h \in \mathbf{B}_\rho(\overline{y}; \underline{y})$;

(ii): $f \succ_\rho h \succ_\rho g$ for some $h \in \mathbf{B}_\rho(\overline{y}; \underline{y}) \Longleftrightarrow \underline{\lambda}_f \geq \overline{\lambda}_g$.

These imply (29). Let us see (ii). Suppose that $f \succ_\rho h \succ_\rho g$ for some $h \in \mathbf{B}_\rho(\overline{y}; \underline{y})$. Since $\underline{\lambda}_f$ is the *lub* of f, by (27), $\underline{\lambda}_f \geq \lambda_h$. Similarly, $\lambda_h \geq \overline{\lambda}_g$. Thus, $\underline{\lambda}_f \geq \overline{\lambda}_g$. Conversely, let $\underline{\lambda}_f \geq \overline{\lambda}_g$. Then, by (27), $f \succ_\rho [\overline{y}, \overline{\lambda}_f; \underline{y}] \succsim_\rho [\overline{y}, \underline{\lambda}_g; \underline{y}] \succ_\rho g$. Hence, we can adopt $[\overline{y}, \overline{\lambda}_f; \underline{y}]$ for h.

Consider (i): The direction $\Longleftarrow$ is obtained by C2. Now, suppose $f \succ_\rho g$. By (11), there is a finite sequence of distinct $g_0, ..., g_m$ such that $f = g_m \succsim_\rho ... \succsim_\rho g_0 = g$ and at least one of each adjacent pair g_l, g_{l+1} belongs to $\mathbf{B}_\rho(\overline{y}; \underline{y})$ for each $l = 0, ..., m-1$. Hence, g_1 belongs to $\mathbf{B}_\rho(\overline{y}; \underline{y})$, since $f = g_m$ and $g_0 = g$ are in $L_\rho(\overline{Y}) - M_\rho$. Since $f \succsim_\rho g_1 \succsim_\rho g$ and $g_1 \in \mathbf{B}_\rho(\overline{y}; \underline{y})$, we have $f \succ_\rho g_1 \succ_\rho g$ by Theorem 5.3.(1).■

Theorem 6.1 provides a complete characterization of incomparabilities involved in the relation $\succsim_\rho$. In order to synthesize this result and Theorem 5.1 for the measurable domain M_ρ, we consider the vector-valued function $\Lambda(f) = (\overline{\lambda}_f, \underline{\lambda}_f)$ for any $f \in L_\rho(X)$ with the binary relation $\geq$ over $\Pi_\rho \times \Pi_\rho$ given by

$$(\xi_1, \xi_2) \geq (\eta_1, \eta_2) \Longleftrightarrow \xi_2 \geq \eta_1. \tag{31}$$

This relation is Fishburn's [6] *interval order*. Using the function $\Lambda(\cdot)$ and $\geq$, we synthesize the results for cases **C** and **IC**.

Theorem 6.2 (Representation). For any $f, g \in L_\rho(X)$, $f \succsim_\rho g \Longleftrightarrow \Lambda(f) \geq \Lambda(g)$.

Proof. Consider case **C** : $f, g \in M_\rho$. Since $\Lambda(f) = (\lambda_f, \lambda_f)$ and $\Lambda(g) = (\lambda_g, \lambda_g)$, the right-hand side of (31) is $\lambda_f \geq \lambda_g$. Thus, the assertion follows from Theorem 5.1. Consider case **IC** that at least one of f, g belongs to $L_\rho(X) - M_\rho$. Theorem 6.1 states that $f \succ_\rho g \Longleftrightarrow \underline{\lambda}_f \geq \overline{\lambda}_g$ and $g \succ_\rho f \Longleftrightarrow \underline{\lambda}_g \geq \overline{\lambda}_f$. Since $f \not\sim_\rho g$ by Theorem 5.3.(1), we have the assertion of the theorem.■

The relation $\geq$ is transitive, and also anti-symmetric, i.e., $(\xi_1, \xi_2) \geq (\eta_1, \eta_2)$ and $(\eta_1, \eta_2) \geq (\xi_1, \xi_2) \Longrightarrow (\xi_1, \xi_2) = (\eta_1, \eta_2)$ (and $\xi_1 = \xi_2$). It is reflexive only for (ξ_1, ξ_2) with $\xi_1 = \xi_2$. Thus, $\geq$ is weaker than a partial ordering. The relation $\succsim_\rho$ is transitive but is neither anti-symmetric nor reflexive.

Incomparability $\bowtie_\rho$ and indifference $\sim_\rho$ may appear similar: indeed, Shafer [21], p.469, discussed whether $\bowtie_\rho$ and $\sim_\rho$ could be defined together and pointed out a difficulty from the constructive point of view. Theorem 6.2 gives a clear distinction between $\sim_\rho$ and $\bowtie_\rho$; by Rule C2, $\sim_\rho$ is transitive, but $\bowtie_\rho$ not. In Fig.3, $g \bowtie_\rho f$ and $f \bowtie_\rho f'$ but $f' \succ_\rho g$.

There are two sources for incomparabilities; $X - Y = X - \cup_{k=0}^{\infty} Y_k$ and a finite ρ. Lemma 5.1.(2) implies that when $f(x) > 0$ for some $x \in X - Y$, this f does not belong to M_k for any $k < \infty$. On the other hand, even when the support of f is included in Y, f does not belong to M_k for $k < k_f$, where k_f is given in (25). Thus, $f, g \in L_\infty(Y)$ are possibly incomparable with respect to $\succsim_\rho \Longleftrightarrow \rho < \max\{k_f, k_g\}$.

Theorem 6.2 corresponds to von Neumann-Morgenstern's [26], p.29, indication of a possibility of a representation of a preference relation involving incomparabilities in terms of a many-dimensional vector-valued function. Our result shows a specific form of their indication. If there are multiple pairs of different benchmarks, our representation theorem may be stated by a higher-dimensional vector-valued function; this will be briefly discussed in Section 8.

Dubra *et al.* [5] obtained the representation result in the form that an incomplete preference relation is represented by a class of expected utility functions. Here, available probabilities are given as arbitrary real numbers in the interval $[0, 1]$ and incompleteness is allowed. Perhaps, this literature is closely related to our consideration of classical EU theory in Section 4.2. When we restrict our attention to the set of pure alternatives to $Y = \cup_{k=0}^{\infty} Y_k$,, we have a complete preference relation, but when we consider the entire set X, the derived preference relation and its extension to $L_{[0,1]}(X)$ could be incomplete. Then, it is an open problem whether Theorem 6.2 can be extended or we need a representation theorem in the form of [5]. Nevertheless, some non-constructive elements are involved here, as pointed out in Section 4.2, and the restriction $\rho < \infty$ is natural from the viewpoint of our motivation of bounded rationality.

For further developments of our theory including practical applications, it is crucial to study properties of the *lub* $\overline{\lambda}_f$ and *glb* $\underline{\lambda}_g$ of $f \in M_\rho$ and/or a general method of calculating them. This exceeds the scope of the present paper; we point out that Theorem 5.2 will be crucial for this development. Here, we consider how to calculate the *lub* $\overline{\lambda}_d$ and *glb* $\underline{\lambda}_d$ of lottery $d = \frac{25}{100} y * \frac{75}{100} \underline{y}$ in Example 3.1 and an additional case. These calculation results will be used in Section 7.

Example 6.1 (Example 3.1 continued). Recall $X = \{\overline{y}, y, \underline{y}\}$, and $d = \frac{25}{100} y * \frac{75}{100} \underline{y}$. Here, we consider two cases

$$\begin{aligned}(A) \quad &: \quad [\overline{y}, \tfrac{9}{10}; \underline{y}] \succ_{B,1} y \succ_{B,1} [\overline{y}, \tfrac{7}{10}; \underline{y}] \text{ and } y \sim_{B,2} [\overline{y}, \tfrac{77}{10^2}; \underline{y}]; \\ (B) \quad &: \quad [\overline{y}, \tfrac{9}{10}; \underline{y}] \succ_{B,1} y \succ_{B,1} [\overline{y}, \tfrac{7}{10}; \underline{y}] \text{ and } y \sim_{B,2} [\overline{y}, \tfrac{83}{10^2}; \underline{y}]\end{aligned} \tag{32}$$

Case (B) is (1) of Fig.3.1, and (A) is additional. These are the cases of risk-averse and risk-lover. Since $d \in L_2(X) - L_1(X)$, we assume $\rho \geq 2$. The *lub* and *glb* of d are given in Table 6.1:

Table 6.1: the *lub* and *glb* of d

	(A)	(B)
$\rho = 2$	$\overline{\lambda}_d = 25/10^2$ & $\underline{\lambda}_d = 0$	the same
$\rho = 3$	$\overline{\lambda}_d = 225/10^3$ & $\underline{\lambda}_d = 175/10^3$	the same
$\rho \geq 4$	$\overline{\lambda}_d = \underline{\lambda}_d = 1925/10^4$	$\overline{\lambda}_d = \underline{\lambda}_d = 2075/10^4$

These results will be calculated in the Appendix. The cases (A) and (B) differ only for $\rho \geq 4$. By (28), $d = \frac{25}{100} y * \frac{75}{100} \underline{y}$ is measurable if and only if $\rho \geq 4$. To calculate the *lub* $\overline{\lambda}_d$ and *glb* $\underline{\lambda}_d$ of d, we need to substitute upper and lower bounds given in (32) for y in $d = \frac{25}{100} y * \frac{75}{100} \underline{y}$. Since the depth $\delta(d)$ is already 2, this substitution with the indifference in the latter in (32) requires $\rho \geq 4$. Indeed, when $\rho \geq 4$, the difference between $y \sim_{B,2} [\overline{y}, \frac{77}{10^2}; \underline{y}]$ in (A) and $y \sim_{B,2} [\overline{y}, \frac{83}{10^2}; \underline{y}]$ in (B) leads to the difference in Table 6.1. When $\rho = 2$ or 3, only the same preferences in (32) can be used for the calculations of $\overline{\lambda}_d$ and $\underline{\lambda}_d$; so, the results given in Table 6.1 are the same.

Consider comparability/incomparability between d and $c := [\overline{y}, \frac{2}{10}; \underline{y}]$. When $\rho \geq 4$, d is directly comparable with $c = [\overline{y}, \frac{2}{10}; \underline{y}]$ by Theorem 5.1, i.e., $c \succ_\rho d$ in (A) and $d \succ_\rho c$ in (B). When $\rho = 2$ or 3, there are gaps between $\overline{\lambda}_d$ and $\underline{\lambda}_d$. Since c is in these gaps; by Theorem 6.1, d and c are incomparable.

7 An Application to a Kahneman-Tversky Example

We apply our theory to an experimental result reported in Kahneman-Tversky [11]. The experimental instance is formulated as Example 6.1, and the relevant lotteries are $c = [\overline{y}, \frac{2}{10}; \underline{y}]$ and $d = \frac{25}{100} y * \frac{75}{100} \underline{y}$; the *lub* and *glb* of d are given in Table 6.1. It is the key how the observed behaviors in the experiment 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 interpret the choice behaviors of subjects who are predicted to show incomparabilties.

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.

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

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

$$\begin{aligned} 0.80u(4000) + 0.20u(0) \quad &< \quad u(3000) \\ 0.20u(4000) + 0.80u(0) \quad &> \quad 0.25u(3000) + 0.75u(0). \end{aligned} \tag{33}$$

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 contradicting classical EU theory is $a \wedge d$. EU theory itself predicts the outcomes $a \wedge c$ and $b \wedge d$, depending upon the value $u(3000)$. This is a variant of many experiments reported.[12]

In [11], no more information is mentioned other than the percentages mentioned above. Consider three possible distributions of the answers in terms of percentages over the four cases, described in Table 7.1: 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 percentage for $b \wedge c$, which implies 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 7.1

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

Let $\overline{y} = 4000$, $\underline{y} = 0$, $y = b = 3000$, and $\rho \geq 2$. Consider the two cases (A) and (B) given in (32). In comparisons between lotteries a and b, the theory predicts, independent of ρ, the choice a (or b) in case (A) (or (B)). Since (A) (or (B)) is the case of risk-lover (or risk-averse), the choice a from a, b is expected to be less frequent than b.

Comparisons between lotteries c and d depend upon ρ. In case $\rho \geq 4$, it follows from Table 6.1 that in (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 (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, in case $\rho \geq 4$, the theory predicts only the diagonal cells $a \wedge c$ and $b \wedge d$ would happen depending upon cases (A) and (B), which are the same as the predictions of classical EU theory. Thus, if all subjects have their cognitive bounds $\rho \geq 4$, the theory is inconsistent with the experimental result.

Consider case $\rho = 3$. Table 6.1 states $\overline{\lambda}_d = \frac{225}{10^3} > \lambda_c = \frac{2}{10} > \underline{\lambda}_d = \frac{175}{10^3}$ in either case (A) or (B); thus, by Theorem 6.1, c and d are incomparable for a subject. When $\rho = 2$, Table 6.1 states $\overline{\lambda}_d = \frac{25}{10^2} > \lambda_c = \frac{2}{10} > \underline{\lambda}_d = 0$ in either (A) or (B) and c and d are also incomparable. Here, notice that $\overline{\lambda}_d = \frac{25}{10^2}$ is much closer to $\lambda_c = \frac{2}{10}$ than $\underline{\lambda}_d = 0$.

Here, we find a conflict in the sense that every subject chose one lottery in each choice problem in the experiment while our theory states that c and d are incomparable under some parameter values. 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.[13] Here, we assume that even when the lotteries are incomparable to a subject, he is forced to make a decision in some arbitrary manner: we specify the

[12] This type of an anomaly is called the "common ratio effect" and has been extensively studied both theoretically and experimentally; typically, the independence axiom is weakened while keeping the probability space as a continuum (cf., Prelec [17] and its references).

[13] It could be difficult for people to show incapability to answer a question if it appears linguistically and logically.clear-cut. The present author knows only one person consciously to refuse to answer such a question. Davis-Maschler [4], Sec.6, it is 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, only 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.

following postulate for his behavior:

Postulate BH: each subject makes a random choice between c and d following the probabilities proportional to the distances from λ_c to $\underline{\lambda}_d$ and from λ_c to $\overline{\lambda}_d$.

Table 7.2 summarizes the above calculated results based on BH. In the case $\rho = 3$, BH implies that the probabilities for the choices c and d are equal in either case (A) or (B), i.e., $\frac{2}{10} - \frac{175}{10^3} : \frac{225}{10^3} - \frac{2}{10} = 1 : 1$. In the case $\rho = 2$, the ratio for the choices c and d becomes $20 : 5$.

Table 7.2

	(A)	(B)
$\rho = 2$	$c : d = 20 : 5$	the same as in (A)
$\rho = 3$	c & d are equally possible	the same as in (A)
$\rho \geq 4$	$\overline{\lambda}_d = \underline{\lambda}_d = 1925/10^4$: c is chosen	$\overline{\lambda}_d = \underline{\lambda}_d = 2075/10^4$: d is chosen

To see the relationship between Table 7.1 and Table 7.2, we specify the distribution between (A) and (B); and that for ρ. We consider the distributions: $r_A : r_B = 2 : 8$ and $r_2 : r_3 : r_{4+} = 5 : 4 : 1$, where r_{4+} is the ratio of subjects with $\rho \geq 4$. The first ratio is based on the idea that more people are risk-averse, and the second that more people are bounded rational. Then, the expected percentage of $a \wedge c$ is calculated as $100 \times \frac{2}{10} \times (\frac{5}{10} \times \frac{20}{25} + \frac{4}{10} \times \frac{1}{2} + \frac{1}{10}) = 14\%$. Calculating the corresponding percentages in the other cases, we obtain Table 7.3:

Table 7.3

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

Table 7.4

	$c : 56\%$	$d : 44\%$
$a : 20\%$	$a \wedge c : 14.4$	$a \wedge d : 5.6$
$b : 80\%$	$b \wedge c : 41.6$	$b \wedge d : 38.4$

The ratio $20 : 80$ for a vs. b is taken from Table 7.2 and is assumed; so it is the same in Table 7.3. The ratio $62 : 38$ for c vs. d in Table 7.3 slightly differs from $65 : 35$ in Table 7.2. This is resulted from our theory and the specification $r_2 : r_3 : r_{4+} = 5 : 4 : 1$ together with the other parameter values. Incidentally, if we specify $r_2 : r_3 : r_{4+} = 4 : 4 : 2$, i.e., more people are less boundedly rational, we have Table 7.4.

The results in Table 7.3 look close to the reported percentages in Table 7.2. 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. This study may lead to observations on new aspects on bounded rationality.

8 Concluding Remarks

We have developed EU utility theory with probability grids and incomparabilities. The permissible probabilities are restricted to the form of ℓ-ary fractions up to a given cognitive bound ρ. The theory is constructive in that it starts with given base preference relations $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ and proceeds with the derivation process (DP) from one layer to the next. When there is no cognitive bound, our theory gives a complete preference relation over $L_\infty(Y)$, and it is a fragment of classical EU theory. However, our main concern was the bounded case $\rho < \infty$.

When $\rho < \infty$, the resultant preference relation $\succsim_\rho$ over $L_\rho(X)$ is not complete as long as X contains a pure alternative y strictly between the benchmarks $\overline{y}$ and $\underline{y}$. 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 gave a complete characterization of incomparabilities and also the representation theorem on $\succsim_\rho$ over $L_\rho(X)$ in terms of the two-dimensional vector-valued function, utilizing Fishburn's [6] interval order. This is interpreted as corresponding to the indication of von Neumann-Morgenstern [26], p.29. In Section 7, we applied the incomparability results to the Allais paradox, specifically, to an experimental instance given in Kahneman-Tversky [11]. We showed that the prediction of our theory is compatible with their experimental result; incomparabilties involved for $\rho = 2$ and $\rho = 3$ are crucial in interpreting their result.

The main part of our theory is about deductive reasoning for decision making, which is bounded with probability grids and a cognitive bound. The derivation of preferences $\langle \succsim_k \rangle_{k=0}^{\rho}$ is formulated as a mathematical induction from base preferences $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$. The source for $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ is based on his inner psychological factors and his past experiences. This part is related to inductive game theory (cf., Kaneko-Matsui [12]). Our approach has some parallelism to the constructive approach, due to Kline *et.al* [13], to an inductive derivation, which includes some deductive part.[14]

Here, we give a few comments on further developments of our theory.

(1) Simon's [23] satisficing/aspiration: As remarked in the end of Section 3.2, $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ can be formulated as a process along the DP. This process can be viewed from Simon's *satisficing/aspiration.* In Step k, the decision maker evaluates a given pure alternative y, which is not yet exactly measured before Step k. Suppose that $[\overline{y}, \overline{\lambda}_y; \underline{y}] \succ_{B,k-1} y \succ_{B,k-1} [\overline{y}, \underline{\lambda}_y; \underline{y}]$, where $\overline{\lambda}_y$ and $\underline{\lambda}_y$ are the best evaluations of y at Step $k-1$. Now, he evaluates each $\lambda \in \Pi_k - \Pi_{k-1}$ with $\overline{\lambda}_y > \lambda > \underline{\lambda}_y$, by considering the propositions:

$$\text{“}[\overline{y}, \lambda; \underline{y}] \succsim_{B,k} y\text{” and/or “}y \succsim_{B,k} [\overline{y}, \lambda; \underline{y}]\text{”}. \tag{34}$$

If he thinks that both hold, then his evaluation of y is determined to be λ, but if this is not the case (there are several subcases to be considered but we ignore the details here), he may improve $\overline{\lambda}_y$ and $\underline{\lambda}_y$, and goes to Step $k+1$ with the improved $\overline{\lambda}_y$ and $\underline{\lambda}_y$.

Simon's satisficing/aspiration suggests that even if the decision maker does not find an exact λ for y at Step k, he may take some λ to be the exact value of y because he gets tired and gives up more thinking. This argument is intimately related to our cognitive bound ρ, but there are still two possible interpretations. In the first interpretation, for each $y \in X$, there is a k such that the decision maker takes λ as good enough for his exact evaluation. Allowing k to depend upon y, the set $Y = \cup_{k=0}^{\infty} Y_k$ coincides with the set of all pure alternatives X. When $\rho < \infty$, it may be the case that $Y = \cup_{k=0}^{\rho} Y_\rho$ coincides with X. In the second interpretation, on the other hand, when he stops at Step k, he may keep different upper $\overline{\lambda}_y$ and lower $\underline{\lambda}_y$. These two cases are different, and we need to study backgrounds and implications for them carefully.

(2): Subjective probability: The above argument 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 (34), i.e., "$[\overline{y}, \lambda; \underline{y}] \succsim_{B,k}$

[14] Broadly speaking, our theory is related to the case-based decision theory by Gilboa-Schmeidler [8], and the frequentist interpretation of probability (cf., Hu [10]). The former may be regarded as evaluations of probabilities for causality (course-effect) from experiences, and the latter concerns about the probability concept itself as frequencies of events. The former is more closely related to inductive game theory, and the latter suggests a possible interpretation of probability, particularly, probability grids in our theory. The concept of probability grids can help us avoid too much freedom of possible probabilities.

$[\overline{y}, E; \underline{y}]$" and/or "$[\overline{y}, E; \underline{y}] \succsim_{B,k} [\overline{y}, \lambda; \underline{y}]$", where $[\overline{y}, E; \underline{y}]$ means that if it is fine tomorrow, the decision maker would get $\overline{y}$ and $\underline{y}$ otherwise; comparing $[\overline{y}, E; \underline{y}]$ with the benchmark lottery $[\overline{y}, \lambda; \underline{y}]$, he evaluate the event E. Thus, we could have an extension of our theory including the subjective probability theory. It would be more difficult to have an extension corresponding to Savage [20], since all probabilities are derived in his theory.

(3): Calculations of the *lub*'s and *glb's* of lotteries $f \in L_\rho(X)$: The *lub* $\overline{\lambda}_f$ and *glb* $\underline{\lambda}_f$ of each lottery f play a crucial role for the characterization of incomparabilities. Calculations for $\overline{\lambda}_f$ and $\underline{\lambda}_f$ for $f \in L_k(X)$ are not straightforward, since minimization and maximization are involved in their definitions; Example 5.1 is about a very simple lottery but still needs complicated calculations. We need a general theory of the properties and calculations of the *lub* and *glb* of lotteries. For this, the condition (25) in Theorem 5.2 dividing between the measurable and nonmeasurable lotteries will be crucial.

(4): Extensions of choices of benchmarks: In the present paper, the benchmarks $\overline{y}$ and $\underline{y}$ are given. 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, benchmarks $\overline{y}$ and $\underline{y}$ are not really fixed; there are different benchmarks than the given ones. We consider two possible extensions of choices of the benchmarks.

One possibility is a *vertical extension*: we take another pair of benchmarks $\overline{\overline{y}}$ and $\underline{\underline{y}}$ such as $\overline{\overline{y}} \succsim_{B,0} \overline{y} \succ_{B,0} \underline{y} \succsim_{B,0} \underline{\underline{y}}$. 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 as long as the permissible grids are different. We may need multiple bases ℓ for probability grids, and may have multiple preference systems even for similar target problems.

Another possibility is a *horizontal extension*: 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.

(5): Extensions of the probability grids Π_ρ: The above extensions may require more subtle treatments of probability grids. This is also related to the other problems such as Nash's [16] bargaining theory to be considered from the viewpoint of bounded rationality. 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 each other. A question is how large $\overline{\ell}$ is required for such classes of problems.

Over all, thinking about these questions make good progress on our expected utility theory with probability grids and incomparabilties. Many more problems are waiting for our further studies.

Appendix

Proof of Lemma 2.2. Let $k \geq 1$. The essential part of Lemma 22 is the direction that if $f \in L_k(X)$, then $f = \mathbf{e} * \mathbf{f}$ for some $\mathbf{f} \in L_{k-1}(X)^\ell$. If $\delta(f) \leq k-1$, then it suffices to take $\mathbf{f} = (f, ..., f) \in L_{k-1}(X)^\ell$ and $f = \mathbf{e} * \mathbf{f}$. Now, we assume $\delta(f) = k$. By $\delta(f) = k \geq 1$, for each $x \in X$, $f(x)$ is expressed as $\sum_{m=1}^{k} \frac{v_m(x)}{\ell^m}$, where $0 \leq v_m(x) < \ell$ for all $m \leq k$. Since the construction of $\mathbf{f}$ takes various steps, we start with its sketch.

We "partition" the total sum $\sum_{x\in X} f(x) = \sum_{x\in X}\sum_{m=1}^{k} \frac{v_m(x)}{\ell^m} = \sum_{m=1}^{k}\sum_{x\in X} \frac{v_m(x)}{\ell^m} = 1$ into ℓ portions so that each has the sum $\frac{1}{\ell}$. However, this may not be directly possible; for example, if $\ell = 10$, $v_1(x) = 2$, and $v_1(x') = 8$, then the sum $\frac{v_1(x)}{\ell^1} + \frac{v_1(x')}{\ell^1} = \frac{2}{10} + \frac{8}{10}$ is not partitioned into 10 portions with $\frac{1}{\ell} = \frac{1}{10}$. To avoid this difficulty, we again "partition" $\frac{2}{10} + \frac{8}{10}$ into the sum of ten 1's with weight $\frac{1}{10}$. In general, for each $m = 1, ..., k$, we represent $\sum_{x\in X} \frac{v_m(x)}{\ell^m} = \frac{1}{\ell^m}\sum_{x\in X} v_m(x)$ by a set natural numbers D_m having the cardinality $\sum_{x\in X} v_m(x)$ with weight $\frac{1}{\ell^m}$ for each element in D_m. We take another partition $\{I_1, ..., I_\ell\}$ of $D_1 \cup ... \cup D_k$ once more so that the sum of weights over I_t is $\frac{1}{\ell}$ for $t = 1, ..., \ell$. Based on this partition, we define a decomposition $\mathbf{f}$ of f.

Formally, let $\tau_0 = 0$ and $\tau_m = \sum_{t=1}^{m}\sum_{x\in X} v_t(x)$ for $m = 1, ..., k$. Then, let $I = \{1, ..., \tau_k\}$, and $D_m = \{\tau_{m-1} + 1, ..., \tau_m\}$ for $m = 1, ..., k$. We define $w(i) = \frac{1}{\ell^m}$ for each $i \in D_m$. Then,

$$\sum_{i\in D_m} w(i) = \sum_{x\in X} \frac{v_m(x)}{\ell^m} \text{ for each } m = 1, ..., k, \tag{35}$$

This implies $\sum_{m=1}^{k}\sum_{i\in D_m} w(i) = \sum_{m=1}^{k}\sum_{x\in X} \frac{v_m(x)}{\ell^m} = \sum_{x\in X} f(x) = 1$. Since each $i \in D_m$ is regarded as coming from one term $v_m(x)$ in $\sum_{x\in X} v_m(x)$, we can define $\varphi(i) = x$. Then, it holds that for each $m = 1, ..., k$,

$$v_m(x) = |\{i \in D_m : \varphi(i) = x\}| \quad \text{for each } x \in X. \tag{36}$$

Note that $v_m(x) = 0$ for any $x \in X$ with $\varphi(i) \neq x$ for any $i \in D_m$. The function φ is used in the final stage of defining a decomposition $\mathbf{f} = (f_1, ..., f_\ell)$ of f.

Now, we show that there is a partition $\{I_1, ..., I_\ell\}$ of $I = \{1, ..., \tau_k\} = D_1 \cup ... \cup D_m$ such that $\sum_{i\in I_t} w(i) = \frac{1}{\ell}$ for $t = 1, ..., \ell$. For this, we define the function W over I by: $W(j) = \sum_{i\leq j} w(i)$ for any $j \in I$; that is, it is the sum of $w(i)$'s over the initial segment of I up to j. Then, $W(\tau_k) = \sum_{m=1}^{k}\sum_{i\in D_m} w(i) = 1$. We have the following "continuity": for any $j \in I$ and $t = 1, ..., \ell$,

$$\tfrac{t-1}{\ell} < W(j) < \tfrac{t}{\ell} \Longrightarrow W(j+1) \leq \tfrac{t}{\ell}. \tag{37}$$

Indeed, let $\frac{t-1}{\ell} < W(j) < \frac{t}{\ell}$ with $j \in D_m$. Then, $m \geq 2$. Then $W(j)$ is expressed as $\frac{s}{\ell^m}$ for some positive integer $s < \ell^m$. Then, $W(j+1) = W(j) + \frac{1}{\ell^{m'}}$ for some $m' \geq m$. Thus, $W(j+1) \leq \frac{t}{\ell}$.

If $W(1) = w(1) = \frac{1}{\ell}$, then, let $I_1 = \{1\}$. Suppose $W(1) = w(1) < \frac{1}{\ell}$. Then, since $W(j)$ is increasing with $W(\tau_k) = 1 > \frac{1}{\ell}$, we find a $\sigma_1 > 1$ by (37) so that $W(\sigma_1) = \frac{1}{\ell}$. Thus, $\sum_{i\in I_1} w(i) = \frac{1}{\ell}$. Similarly, we can construct $I_2, ..., I_\ell$ so that for $t = 2, ..., \ell$, $I_t = \{\sigma_{t-1}, ..., \sigma_t\}$ and

$$\textstyle\sum_{i\in I_t} w(i) = \frac{1}{\ell}. \tag{38}$$

Now, we define functions $f_1, ..., f_\ell$ by: for $t = 1, ..., \ell$,

$$f_t(x) = \sum_{m=1}^{k} \frac{|\{i \in I_t \cap D_m : \varphi(i) = x\}|}{\ell^{m-1}} \quad \text{for each } x \in X. \tag{39}$$

We show that these form a decomposition of f.

Let us see hat $f_t \in L_{k-1}(X)$. If $|\{i \in I_t \cap D_1 : \varphi(i) = x\}| = 1$, then, by (38), I_t consists of a unique element i with $w(i) = \frac{1}{\ell}$. In this case, $f_t(x) = 1$; hence $f_t \in L_0(X) \subseteq L_{k-1}(X)$. Consider the other case where $|\{i \in I_t \cap D_1 : \varphi(i) = x\}| = 0$. The summation in (39) has at most length $k-1$ and is expressed as

$$f_t(x) = \sum_{m=1}^{k-1} \frac{|\{i \in I_t \cap D_{m+1} : \varphi(i) = x\}|}{\ell^m} \text{ for each } x \in X.$$

Since $|\{i \in I_t \cap D_{m+1} : \varphi(i) = x\}| \leq |\{i \in D_{m+1} : \varphi(i) = x\}| = v_{m+1}(x) < \ell$ for all $m \leq k-1$ by (36), the sum is expressed as the form of a number in Π_{k-1}. Hence, $f_t(x) \in \Pi_{k-1}$.

Now, we have, by (38),

$$\begin{aligned}\sum_{x\in X} f_t(x) &= \sum_{x\in X}\sum_{m=1}^{k} \frac{|\{i \in I_t \cap D_m : \varphi(i) = x\}|}{\ell^{m-1}} = \sum_{m=1}^{k} \frac{|I_t \cap D_m|}{\ell^{m-1}} \\ &= \sum_{m=1}^{k}\sum_{i\in I_t\cap D_m} \frac{1}{\ell^{m-1}} = \ell \sum_{m=1}^{k}\sum_{i\in I_t\cap D_m} \frac{1}{\ell^{m}} = \ell \times \sum_{i\in I_t} w(i) = 1.\end{aligned}$$

Thus, $f_t \in L_{k-1}(X)$ for all $t = 1, ..., \ell$. Finally, for each $x \in X$, $\sum_{t=1}^{\ell} \frac{1}{\ell} \times f_t(x)$ is calculated as

$$\sum_{t=1}^{\ell}\sum_{m=1}^{k} \frac{|\{i \in I_t \cap D_m : \varphi(i) = x\}|}{\ell^m} = \sum_{m=1}^{k}\sum_{t=1}^{\ell} \frac{|\{i \in I_t \cap D_m : \varphi(i) = x\}|}{\ell^m} = \sum_{m=1}^{k} \frac{v_m(x)}{\ell^m} = f(x).$$

■

Proof of Theorem 3.1. First, we sketch the proof. Let $\langle \succsim_k \rangle_{k=0}^{\rho}$ be the resultant relations from the DP, provided that $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ is given. To prove that the well-definedness of each $\succsim_k$ in $L_k(X)$, we consider a stronger relation $\succsim_{V,k}$ but show that this $\succsim_{V,k}$ is a binary relation and for any $f, g \in L_k(X)$, and $f \succsim_k g$ implies $f \succsim_{V,k} g$. Hence, $\succsim_k$ is also a binary relation and is well-defined. For this, we can assume $\rho = \infty$. Indeed, when $\rho < \infty$, we extend $\langle \succsim_{B,k} \rangle_{k=0}^{\rho}$ to $\langle \succsim_{B,k} \rangle_{k=0}^{\infty}$ by assuming that $\succsim_{B,t} = \succsim_{B,\rho}$ for all $t \geq \rho$.

Let $\langle \succsim_{B,k} \rangle_{k=0}^{\infty}$ satisfying Axioms B0 to B3 be given. We define $V : X \to R$ by: for any $x \in X$,

$$V(x) = \inf\{\lambda \in \Pi_\infty : [\overline{y}, \lambda; \underline{y}] \succsim_{B,k} x \text{ for some } k < \infty\}. \tag{40}$$

Because of Axiom B0, this is well-defined. Also, $V(\overline{y}) = 1$ and $V(\underline{y}) = 0$ by Axiom B1. We show that for any $x \in X$ and $\lambda \in \Pi_\infty$,

$$[\overline{y}, \lambda; \underline{y}] \succsim_{B,\infty} x \implies \lambda \geq V(x); \text{ and } x \succsim_{B,\infty} [\overline{y}, \lambda; \underline{y}] \implies V(x) \geq \lambda. \tag{41}$$

Recall $\succsim_{B,\infty} = \cup_{k=0}^{\infty} \succsim_{B,k}$. Let $\lambda \in \Pi_k$ and $[\overline{y}, \lambda; \underline{y}] \succsim_{B,k} x$ for some k. By (40), $\lambda \geq V(x)$. Next, let $x \succsim_{B,k} [\overline{y}, \lambda; \underline{y}]$ for some k. By (40), there is a sequence $\{\lambda_\nu\}$ in Π_∞ with $[\overline{y}, \lambda_\nu; \underline{y}] \succsim_{B,\infty} x$ for all $\nu \geq 0$ and $\lambda_\nu \to V(x)$ as $\nu \to \infty$. Now, $[\overline{y}, \lambda_\nu; \underline{y}] \succsim_{B,\infty} x$ implies $[\overline{y}, \lambda_\nu; \underline{y}] \succsim_{B,k'} x$ for some k'. Taking $k'' \geq \max\{k, k'\}$, we have $[\overline{y}, \lambda_\nu; \underline{y}] \succsim_{B,k''} x$ and $x \succsim_{B,k''} [\overline{y}, \lambda; \underline{y}]$ by B3. Thus, $\lambda_\nu \geq \lambda$ by Axiom B2. Since this holds for all $\nu \geq 0$, we have $V(x) \geq \lambda$.

We define the *eu-function* V_e and the *eu-preference relation* $\succsim_V$ over $L_\infty(X)$ by

$$V_e(f) = \textstyle\sum_{x\in X} f(x)V(x) \text{ for any } f \in L_\infty(X); \tag{42}$$

$$\text{for any } f, g \in L_\infty(X), \quad f \succsim_V g \iff V_e(f) \geq V_e(g). \tag{43}$$

Since each f has a finite support, the sum in (42) is well-defined. Thus, $\succsim_V$ is a complete and transitive binary relation over $L_\infty(X)$. Since $V(\overline{y}) = 1$ and $V(\underline{y}) = 0$, we have $V_e([\overline{y}, \lambda; \underline{y}]) = \lambda$. Thus, it follows from (41) and (43) that

$$[\overline{y}, \lambda; \underline{y}] \succsim_{B,\infty} x \implies [\overline{y}, \lambda; \underline{y}] \succsim_V x; \text{ and } x \succsim_{B,\infty} [\overline{y}, \lambda; \underline{y}] \Longrightarrow x \succsim_V [\overline{y}, \lambda; \underline{y}]. \tag{44}$$

The key step is that for any k $(0 \leq k < \infty)$ and any $f, g \in L_k(X)$,

$$f \succsim_k g \Longrightarrow \ f \ \succsim_V \ g; \ \text{ and } f \succ_k g \Longrightarrow \ f \ \succ_V \ g. \tag{45}$$

Using this, we prove that $f \succsim_k g$ and not $(f \succsim_k g)$ do not happen simultaneously. Suppose, on the contrary, that $f \succsim_k g$ and not $(f \succsim_k g)$ happen. The latter, not $(f \succsim_k g)$, happens as a part of $g \succ_k f$. However, by (45), we have $f \succsim_V g$ and $g \succ_V f$, which is impossible because $\succsim_V$ is a binary relation by (43). Hence, either $f \succsim_k g$ or not $(f \succsim_k g)$. This implies that $\succsim_k$ is a binary relation over $L_k(X)$.

To show (45), first, we note the following:

$$V_e(\mathbf{e} * \mathbf{f}) = \sum_{t=1}^{\ell} \tfrac{1}{\ell} V_e(f_t) \text{ for any } \mathbf{f} \in L_k(X)^{\ell} \text{ and } k \geq 0. \tag{46}$$

This follows from (42) that $V_e(\mathbf{e} * \mathbf{f}) = \sum_{x \in X} (\mathbf{e} * \mathbf{f})(x) V(x) = \sum_{x \in X} \sum_{t=1}^{\ell} \frac{1}{\ell} f_t(x) V(x) = \sum_{t=1}^{\ell} \frac{1}{\ell} \sum_{x \in X} f_t(x) V(x) = \sum_{t=1}^{\ell} \frac{1}{\ell} V_e(f_t)$. Here, interchangeability of $\sum_{x \in X}$ and $\sum_{t=1}^{\ell}$ follows from the fact that $f_1, ..., f_\ell$ have finite supports.

The last step is to show (45) along the inductive construction of $\succsim_k, k = 0, ...$ For $k = 0$, it follows from Axiom B0 and (44) that $\overline{y} \succsim_V x \succsim_V \underline{y}$ for any $x \in X$.

Now, consider $k0$. Then, $f \succsim_k g$ is $f \succsim_{B,k} g$. By (44), we have $f \succsim_V g$.

Consider $k1 : f \succsim_k g$ is derived by Axiom C1 with the decompositions $\mathbf{f} = (f_1, ..., f_\ell)$ of f and $\mathbf{g} = (g_1, ..., g_\ell)$ of g with $\mathbf{f} \succsim_{k-1} \mathbf{g}$. Here, the induction hypothesis is that (45) holds for $\succsim_{k-1}$ and $\succ_{k-1}$. Hence, $\mathbf{f} \succsim_{k-1} \mathbf{g}$ implies $f_t \succsim_V g_t$ for $t = 1, ..., \ell$. By (46), $V_e(f) = V_e(\mathbf{e} * \mathbf{f}) \geq V_e(\mathbf{e} * \mathbf{g}) = V_e(g)$, i.e., $f \succsim_V g$. The case of $f \succ_k g$ is similar.

Consider $k2 :$ (11) states that this step is to extend the preferences obtained from $k0$ and $k1$ by transitivity. We can prove (45) by induction along this extension. Let $f \succsim_k h$ and $h \succsim_k g$. We assume that (45) holds for these. By (45), we have $f \succsim_V h$ and $h \succsim_V g$. By transitivity of $\succsim_V$, we have $f \succsim_V g$.■

Calculations of the *glb* and *lub* in Table 6.1: Here, we calculate the results in Table 6.1 for Case $(B) : y \sim_B [\overline{y}, \frac{83}{10^2}; \underline{y}]$.

Case $\rho = 2$: C1 is not applied to substituting $[\overline{y}, \frac{83}{10^2}; \underline{y}]$ for y in $d = \frac{25}{100} y * \frac{75}{100} \underline{y}$, but may be possible if we sacrifice accuracy; that is, we substitute $\overline{y}$ and $\underline{y}$ for y in d, and can obtain the following

$$\tfrac{25}{100} \overline{y} * \tfrac{75}{100} \underline{y} \succ_2 d \succ_2 \underline{y}. \tag{47}$$

Thus, the *lub* and *glb* of d are given as $\overline{\lambda}_d = \frac{25}{10^2}$ and $\underline{\lambda}_d = 0$. We verify only $\overline{\lambda}_d = \frac{25}{10^2}$: first, $\frac{5}{10} \overline{y} * \frac{5}{10} \underline{y} \succ_1 \frac{5}{10} y * \frac{5}{10} \underline{y}$ holds by C1 since $\overline{y} \succ_0 y$ and $\underline{y} \sim_0 \underline{y}$ by B1 and Step 0 of the DP. Hence $\frac{25}{100} \overline{y} * \frac{75}{100} \underline{y} = \frac{5}{10} (\frac{5}{10} \overline{y} * \frac{5}{10} \underline{y}) * \frac{5}{10} \underline{y} \succ_2 \frac{5}{10} (\frac{5}{10} y * \frac{5}{10} \underline{y}) * \frac{5}{10} \underline{y} = \frac{25}{100} y * \frac{75}{100} \underline{y}$. This is the best upper bound $\mathbf{B}_2(\overline{y}; \underline{y})$. It is assumed by Example 3.1.(1) that $[\overline{y}, \frac{9}{10}; \underline{y}] \succ_1 y$. But $\rho = 2$ prevents an application of Rule C1 from improving the evaluation in (47).

Case $\rho = 3$: Consider the second line in Table 6.1, i.e., $\overline{\lambda}_d = \frac{225}{10^3}$ and $\underline{\lambda}_d = \frac{175}{10^3}$. Here, we verify only $\underline{\lambda}_d = \frac{175}{10^3}$. First, $y \succ_1 [\overline{y}, \frac{7}{10}; \underline{y}]$ by Example 3.1.(1). Then, by C1, $[y, \frac{5}{10}; \underline{y}] = \frac{5}{10}y * \frac{5}{10}\underline{y} \succ_2 \frac{5}{10}[\overline{y}, \frac{7}{10}; \underline{y}] * \frac{5}{10}\underline{y} = [\overline{y}, \frac{35}{10^2}; \underline{y}]$. Again, by C1, $d = \frac{25}{10^2}y * \frac{7}{10^2}\underline{y} = \frac{5}{10}[y, \frac{5}{10}; \underline{y}] * \frac{5}{10}\underline{y} \succ_3 \frac{5}{10}[\overline{y}, \frac{35}{10^2}; \underline{y}] * \frac{5}{10}\underline{y} = [\overline{y}, \frac{175}{10^3}; \underline{y}]$. Hence, $d \succ_3 [\overline{y}, \frac{175}{10^3}; \underline{y}]$. By watching this derivation carefully, we see that this is the best lower bound of d in $\mathbf{B}_3(\overline{y}; \underline{y})$. Thus, $\underline{\lambda}_d = \frac{175}{10^3}$.

Case $\rho \geq 4$: We have no constraint on substitution of $[\overline{y}, \frac{83}{10^2}; \underline{y}]$ for y; we have the third line in in Table 6.1, i.e., $\underline{\lambda}_d = \overline{\lambda}_d = \frac{2075}{10^4}$. We verify this following our derivation process; since $\frac{5}{10}y * \frac{5}{10}\underline{y} \sim_3 \frac{425}{10^3}\overline{y} * \frac{575}{10}\underline{y}$ by C1 and $y \sim_{B,2} [\overline{y}, \frac{83}{10^2}; \underline{y}]$, we have $d = \frac{5}{10}(\frac{5}{10}y * \frac{5}{10}\underline{y}) * \frac{5}{10}\underline{y} \sim_4 \frac{2075}{10^4}\overline{y} * \frac{7825}{10}\underline{y}$.■

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

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

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

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

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

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

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

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

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

[12] Kaneko, M., and A. Matsui, (1999), Inductive game theory: discrimination and prejudices, *Journal of Public Economic Theory* 1, 101-137.

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

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

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

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

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

[18] 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.

[19] Rubinstein, A., (1998), *Modeling Bounded Rationality,* MIT Press, Cambridge.

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

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

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

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

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

[25] Tserenjigmid, G., (2015), Theory of Decisions by Intra-dimensional Comparisons, *Journal of Economic Theory* 159, 326-338.

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