Expected Utility Theory with Probability Grids and Preferential Incomparabilities

Mamoru Kaneko

Waseda INstitute of Political EConomy
Waseda University
Tokyo, Japan
Expected Utility Theory with Probability Grids and Preferential Incomparabilities

Mamoru Kaneko†

May 31, 2017

Abstract

We reformulate expected utility theory by separating the measurement of utility from pure alternatives and its extension to lotteries involving risks. At the same time, we introduce cognitive bounds on the depths of permissible probability values; for example, probabilities have only decimal (or binary) fractions of finite depths that are not greater than a given bound. We allow the preference relation in question to be incomplete. When no depth restrictions are given, the axioms determine uniquely a complete preference relation, which can be considered classical expected utility theory. When a finite cognitive bound is given, the axioms allow multiple preference relations including incomparabilities on some lotteries. We give a complete characterization of preferential incomparabilities, and a representation theorem in terms of a 2-dimensional vector-valued utility function. We exemplify the Allais paradox within our theory, and argue that the prediction of our theory is well compatible with the experimental results reported by Kahneman-Tversky when we adopt a very shallow cognitive bound.

JEL Classification Numbers: C72, C79, C91

Key words: Expected Utility, Measurement of Utility, Probability Grids, Cognitive Bounds, Preferential Incomparabilities

1 Introduction

Although expected utility (EU) theory plays an important role in economics and game theory, various paradoxes have been reported since Allais [2]. To study such paradoxes, here we restrict available probabilities in EU theory to decimal (ℓ-ary, in general) fractions up to a cognitive bound ρ; if ρ is a finite natural number k, it is given as Π_ρ = Π_k = {0, \frac{1}{10^k}, ..., \frac{10^k}{10^k}}; and if ρ = ∞, it is given as Π_ρ = \bigcup_{k=0}^{∞} Π_k. We allow a preference relation to be incomplete. The limit case with ρ = ∞ can be considered classical EU theory in the sense of von Neumann-Morgenstern [22] (cf., Herstein-Milnor [10] and Hammond [7] for later developments). Our theory indicates where the Allais paradox emerges and shows how it may be avoided.

For a finite k, the set of grids Π_k = \{0, \frac{1}{10^k}, \frac{1}{10^k}, ..., \frac{10^k}{10^k}\} may be interpreted as a partition of

*The author thanks P. Wakker, J. J. Kline, M. Lewandowski, L. Tang, A. Dominik, and S. Shiba for helpful comments on earlier versions of this paper. The author is supported by Grant-in-Aids for Scientific Research No. 26245026, Ministry of Education, Science and Culture.

†Waseda University, Tokyo, Japan (mkanekoepi@waseda.jp)
the continuum $[0, 1]$ so that the grids in $\Pi_k$ represent the components in the partition. We take, however, the opposite view that a set of grids is, cognitively and theoretically, more primitive than $[0, 1]$. Our development goes from shallower to deeper cases; the decision maker thinks about his evaluation with a small $k$ to larger $k$’s up to a given bound $\rho$. Even when $\rho = \infty$, $\Pi_\rho = \cup_{k=0}^{\infty} \Pi_k$ does not reach the continuum $[0, 1]$ and it is still a subset of the set of rational numbers. We are particularly interested in cases of finite cognitive bounds, and the limit case is the reference point. In this sense, our theory deals with “bounded rationality” due to Simon [19], [20]; Simon [20] criticized EU theory as a description of “super rationality.” The limit case in our theory is interpreted as corresponding to Simon’s “super rationality”.

A hint for our development is found in von Neumann-Morgenstern [22]. They divided the motivating argument into

(i) measurement of utility in terms of probability; and

(ii) extension to lotteries involving risks such as plans for future events.

This separation, however, is not reflected in their mathematical development. We separate and formulate these steps as mathematical induction, that is, from simple to more complex cases. In particular, we restrict Step (i) to the measurement of utility from a pure (no risk) alternative in terms of two benchmarks and available probabilities. Step (ii) extends this measurement to lotteries with probabilities of larger depths.

As in Figure 1, these steps are described by Axioms B0 to B4. The lotteries with probability grids and a sequence of preference relations $\succsim_k^{\rho}$ prepare the language of our theory; $\rho = 2$ in Figure 1. These axioms describe the meaning of benchmarks, how the decision maker evaluates a given pure alternative, and how he extends preferences $\succsim_k$ to $\succsim^{k+1}$ ($k = 1, \ldots, \rho - 1$) over the lotteries with finer probability grids.

The benchmarks and permissible probabilities form a benchmark scale, depicted in Figure 2; an analogy is a thermometer with the 0 and 100 degree points with the other 99 grids. With the benchmark scale, the decision maker measures utility from each pure alternative; in Figure 2, $y$ corresponds to a grid in the benchmark scale, but neither $x$ nor $z$ has no corresponding grids.\footnote{This method is dual to the measurement method in terms of certainty equivalent of a lottery (cf., Kontek-Lewandowski [14] and its references). In our method, the set of benchmarks up to some cognitive bound is a basic}

Figure 1: From Simple to Complex

(iii) transitivity of preference relations; and

(iv) extension of preferences from simple to complex cases.

These steps provide the foundation for our theory, and the measured utilities

\[\begin{align*}
\text{Base Facet} & \rightarrow \text{lotteries of depth 1} \\
\text{lotteries of depth 2} & \rightarrow \text{lotteries of depth 3} \\
\text{lotteries of depth 4} & \rightarrow \text{lotteries of depth 5} \\
& \rightarrow \text{lotteries of depth n}
\end{align*}\]
are extended to more complex lotteries up to those within a given cognitive bound $\rho$ ($\rho$ is finite or infinite). The extension is formulated in terms of mathematical induction from the lotteries of depth $k$ to those of depth $k + 1$ up to $\rho$. Specifically, the extension is obtained by applying Axioms B4 and B0 (transitivity).

It is an important feature that our theory is constructive. In the literature, imposed axioms are typically requirements for a preference relation but do not directly describe a process of decision making. Here, the axioms describe decision making; the weakest relation $\succeq_1^{\mathsf{c}}$, which we call the central (preference) relation, is defined in Section 4.1 from this point of view. We will study the behavior of this relation.

To illustrate the above discussion, we use an example in Kahneman-Tversky [12], which we call the KT example and discuss in detail in Section 6. Consider three alternatives $y$, $y$, $y$ with strict preferences $y \succ y \succ y$ for the decision maker. The reward from $y$, $y$, or $y$ is 4,000$, 3,000$, or 0$, respectively. Taking $y$ and $y$ as benchmarks, the decision maker evaluates $y$ in terms of $y$, $\lambda$; $y$ and permissible probabilities; the decision maker engages in a thought experiment to find (choose) some probability $\lambda$ so that $y$ is indifferent to the lottery consisting of $y$ with probability $\lambda$ and $y$ with the remaining probability $1 - \lambda$. This indifference is given by the first preference relation $\sim_1$ in $\mathcal{R}_k$, that is,

$$y \sim_1 [y, \lambda; y]$$

We call $[y, \lambda; y]$ a benchmark lottery (standard gamble in the literature). In Figure 2, the decision maker succeeds in finding the indifferent $[y, \lambda; y]$ to $y$ but he does not succeed for $x$ or $z$. This depends upon the choice of benchmarks $y$ and $y$; in this paper, the benchmarks are fixed, but we remark on the different choices of benchmarks in Section 7.2.

The thought experiment starts to find $\lambda$ in $\Pi_1 = \{0, 1 \frac{1}{10}, ..., 10 \frac{1}{10}\}$ for (1). If the decision maker finds no satisfactory $\lambda$ in $\Pi_1$, he moves to the next $\Pi_2 = \{0, 1 \frac{1}{10}, ..., 10 \frac{2}{10}\}$. Once he finds some

\footnote{Bloome et al. [6] gave a constructive framework for expected utility theory from the viewpoint of propositional logic. A hierarchy of decision making is described by a compound conditional statement; the theory is constructed based on induction over the nested structure of conditional statements. Ours is based on the induction over the depths of probability grids.}

\footnote{In the experiments reported in [12], the money amount was measured in the Israel pounds at that time (the median net monthly income for a family was approximately 3,000 Israel pounds).}

Figure 2: Measurement step
λ_\rho in \Pi_\rho for (1), λ_\rho is considered as the “utility value” from y. The decision maker may stop at depth l(1) ≤ ρ even though he finds no satisfactory λ, perhaps because it is too complex. Now, the decision maker has the set Y of such pure alternatives y with some probability λ_y. These form a base facet \( F = \langle y, y; Y; \{λ_y\}_{y \in Y} \rangle. \) We impose Axioms B1 to B3 on F.

The decision maker consciously chooses a λ_y for (1). On the other hand, some lotteries are given to him, for example, \( d = [y, \frac{25}{100}; y], \) which is not a benchmark lottery. Still, the decision maker looks for \( λ \) so that
\[
[y, \frac{25}{100}; y] \sim^k [y, λ; y].
\] (2)

Here, \( \sim^k \) is the indifference relation for some extension step \( k. \) This search differs from the above thought experiment because \( y \) is already evaluated in (1). Suppose that \( λ_y = \frac{85}{100} \) is the probability given by (1). In classical EU theory, \([y, λ_y; y]\) is substituted for \( y \) in the left lottery because the decision maker is indifferent between them;
\[
[y, \frac{25}{100}; y] \text{ and } [y, \frac{85}{100}; y, \frac{25}{100}; y] = [y, \frac{2125}{1000}; y].
\] (3)

The right-hand lottery includes probability values of the 4th decimal place, while the lottery in (1) involves only probabilities of the 2nd decimal place. Probability values including the 4th decimal places are, perhaps, too precise for ordinary decision making. 4

We treat the lotteries with probabilities of shallow depths in the following manner. First, we consider the sets of lotteries \( L_0(Y), L_1(Y), \ldots, \) where \( L_k(Y) \) is the set of lotteries over Y with probability values of at most the k-th decimal place. Lottery \([y, \frac{25}{100}; y]\) belongs to \( L_0(Y), \) and \([y, \frac{2125}{1000}; y]\) belongs to \( L_1(Y) \) but not to \( L_2(Y). \) We formulate Step (ii) as a composition of lotteries from \( L_k(Y) \) to a lottery in \( L_{k+1}(Y), \) which forms mathematical induction. In order to describe these steps, we have prepared a sequence of preference relations \( \gtrless \) for either \( ρ < \infty \) or \( ρ = \infty. \)

In case \( ρ = \infty, \) there are no restrictions on Step (ii). In this case, the resulting preference relation \( \gtrless^\infty = \bigcup_{k=0}^\infty \gtrless^k \) is expressed by expected utility; here, the resulting relation \( \gtrless^\infty \) enjoys completeness and is uniquely determined by Axioms B0 to B4. In this sense, our theory becomes classical EU theory, corresponding to Simon’s “super rationality.” This is described in Table 1.2.

| \( ρ < \infty \) | \( M(F; ρ) \subseteq L_0(Y) \) |
| \begin{align*}
\text{multiple and incomplete } & \gtrless^\rho \\
\text{ weakest } & \gtrless_c
\end{align*} | \( \implies \) |

| \( ρ = \infty \) | \( M(F; ρ) = L_0(Y) \) |
| \begin{align*}
\text{unique and complete } & \gtrless^\infty = \gtrless_c
\end{align*} |

In case \( ρ < \infty, \) the axioms allow multiple and incomplete resulting relation \( \gtrless^\rho \) on \( L_0(Y), \) which is the last component of \( \bigcup_{k=0}^\infty \gtrless^k \) \( /k=0. \) Here, we consider the weakest preference sequence \( \bigcup_{k=0}^\infty \gtrless_c \) which is uniquely determined by Axioms B0-B4; others contain some superfluous preferences. We call the resulting relation \( \gtrless^\rho \) (last component of \( \bigcup_{k=0}^\infty \gtrless_c \)) the central preference relation, and this gives a a certain subset \( M(F; ρ) \) of \( L_0(Y), \) called the measurable domain, so that each \( f \in M(F; ρ) \) is measured by the benchmark scale. The central relation \( \gtrless^\rho \) is complete over \( M(F; ρ) \) but exhibits incomparabilities outside \( M(F; ρ). \) These are summarized in Tables 1.1.

When \( f \) is outside \( M(F; ρ), \) a lottery \( f \) is not exactly measured by the benchmark scale, but it has still lower and upper bounds. The least upper bound (LUB) and greatest lower bound (GLB) of \( f \) are well defined. Using these concepts, we can fully characterize the incomparabilities involved in \( \gtrless_c. \) These characterizations are summarized as a representation theorem: \( \gtrless_c \) is

---

4Recall that the significance level for hypothesis testing in statistics is typically 5% or 1%.  

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>Table 1.2</th>
</tr>
</thead>
</table>
represented by the 2-dimensional vector-valued function consisting of the LUB and GLB. This is coherent with the suggestion given by von Neumann-Morgenstern [22], p.29. These are given in Section 5.

We apply our analysis to the Allais paradox in the above mentioned example due to Kahneman-Tversky [12] in Section 6. In this example, we avoid the Allais paradox if and only if the cognitive bound \( \rho \) is 2. Here, incomparability plays an important role; we assume that when an experimental subject is asked to choose one from two alternatives, he would choose one even though the alternatives are incomparable for him. The concepts of LUB and GLB are crucial for this analysis. In the end of Section 6, we give a numerical test of the prediction of our theory in the example, and it is exactly the same as the result reported in [12].

There are various studies of the Allais paradox suggesting some possible resolutions. A typical one is to allow the utility function to be nonlinear with respect to probabilities (cf., Prelec [16], Kontek-Lewandowski [14]); and the other is to allow similarities so that small differences in probabilities are indistinguishable (cf. Rubinstein [17]). Our theory differs from the former in that we still keep linearity. The latter assumes the continuum \([0, 1]\) and similarity is introduced as a superstructure of it, while our theory is based on the view that probability grids are primitives and go to a richer world but do never reach the continuum \([0, 1]\). In the example due Kahneman-Tversky, we need to go only to \( \rho = 2 \) and \( \Pi_2 = \{ \frac{0}{10^\ell}, \frac{1}{10^\ell}, \ldots, \frac{10^\ell}{10^\ell} \} \).

This paper is organized as follows: Section 2 explains the concept of probability grids. Section 3 gives a formulation of a basic facet together with the basic axioms for measurement, and another axiom for extension. Section 4 discusses the central preference relation \( \succ^\rho \) for \( \rho < \infty \) and \( \rho = \infty \), and Section 5 studies incomparabilities involved in \( \succ^\rho \) for \( \rho < \infty \). Section 6 examines a paradoxical example from Kahneman-Tversky [12]. Section 7 remarks on a derivation of a basic facet and on multiple base facets. Section 8 concludes this paper and comments on further possible studies.

## 2 Lotteries with Probability Grids and Preferences

First, we define the concept of probability grids and the sets of lotteries with probability grids. Then, Lemma 2.2 shows that these sets are connected from coarser to finer ones, which facilitates mathematical induction. Finally, a preference relation is defined.

Let \( \ell \) be an integer with \( \ell \geq 2 \). This \( \ell \) is the base for describing probability values; we take \( \ell = 10 \) in all examples in this paper. The *set of probability grids* \( \Pi_k \) is defined as

\[
\Pi_k = \left\{ \frac{\nu}{\ell^k} : \nu = 0, \ldots, \ell^k \right\} \text{ for any finite } k \geq 0. \tag{4}
\]

Here, \( \Pi_1 = \left\{ \frac{\nu}{\ell} : \nu = 0, \ldots, \ell \right\} \) is the basic set of probability grids, but \( \Pi_0 = \{0, 1\} \) is used for convenience’s sake. We use the standard arithmetic rules over \( \Pi_\infty := \bigcup_{k=0}^\infty \Pi_k \). This includes reduction by eliminating common factors; for example, when \( \ell = 10 \), \( \frac{20}{10^2} \) is the same as \( \frac{2}{10} \). Hence, \( \Pi_k \subseteq \Pi_{k+1} \) for \( k = 0, 1, \ldots \) The parameter \( k \) is precision of probabilities the decision maker can use. We define the *depth* of each \( \lambda \in \Pi_\infty \) by: \( \delta(\lambda) = k \) iff \( \lambda \in \Pi_k - \Pi_{k-1} \).

We use the standard equality = and strict inequality > over \( \Pi_\infty \). Then, trichotomy holds: for any \( \lambda, \lambda' \in \Pi_\infty \),

\[
\text{either } \lambda > \lambda', \lambda = \lambda', \text{ or } \lambda < \lambda'. \tag{5}
\]

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

Now, we show that each \( \Pi_k \) is obtained from \( \Pi_{k-1} \) by taking weighted sums of elements in \( \Pi_k \) with equal weights. This lemma is basic for our induction method.
Lemma 2.1 (Decomposition of probabilities). \( \Pi_k = \{ \sum_{t=1}^{\ell} \frac{1}{r_i} \lambda_t : \lambda_1, \ldots, \lambda_\ell \in \Pi_{k-1} \} \) for any \( k (1 \leq k < \infty) \).

**Proof.** It is easy to see that the right-hand set is included in \( \Pi_k \). Consider the converse. Each \( \lambda \in \Pi_k \) is expressed as \( \lambda = 1 \) or \( \lambda = \sum_{t=1}^{\ell} \frac{1}{r_i} \) where \( 0 \leq \nu_t < \ell \) for \( t = 1, \ldots, k \). If \( \lambda = 1 \), let \( \lambda_1 = \ldots = \lambda_\ell = 1 \in \Pi_{k-1} \), and \( \lambda = \sum_{t=1}^{\ell} \frac{1}{r_i} \lambda_t \). Consider the second case. Then, let \( \lambda_1 = \ldots = \lambda_{\nu_t-1} = 1, \lambda_{\nu_t+1} = \frac{\nu_t}{\gamma_t} + \ldots + \frac{\nu_{\ell}}{\gamma_{\ell}}, \) and \( \lambda_t = 0 \) for \( t = \nu_t + 2, \ldots, \ell \). This definition is applied even when \( \nu_t = 0 \). These \( \lambda_1, \ldots, \lambda_\ell \) belong to \( \Pi_{k-1} \) and \( \lambda = \sum_{t=1}^{\ell} \frac{1}{r_i} \lambda_t \). \( \square \)

In general, \( \Pi_\infty = \cup_{k=0}^{\infty} \Pi_k \) is a countable and proper subset of the set of all rational numbers. For example, when \( \ell = 10 \), \( \Pi_\infty \) has no recurring decimals, which are also rationals. It is important to note that \( \Pi_\rho \) depends upon the base \( \ell \); for example, \( \Pi_1 \) with \( \ell = 3 \) has \( \frac{1}{3} \), but \( \Pi_\infty \) with \( \ell = 10 \) has no element corresponding to \( \frac{1}{3} \).

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

\[
L_k(Y) = \{ f : f \) is a function from \( Y \) to \( \Pi_k \) with \( \sum_{y \in Y} f(y) = 1 \}. \tag{6}
\]

Since \( \Pi_k \subseteq \Pi_{k+1} \) for all \( k \), it holds that \( L_k(Y) \subseteq L_{k+1}(Y) \). We define \( L_\infty(Y) = \cup_{k=0}^{\infty} L_k(Y) \).

We denote the **comparative bound** by \( \rho \), which is a natural number or infinity. If \( \rho = k < \infty \), then \( L_\rho(Y) = L_k(Y) \), and if \( \rho = \infty \), then \( L_\rho(Y) = L_\infty(Y) \). For \( f, g \in L_\rho(Y) \), we define \( f = g \iff f(y) = g(y) \) for all \( y \in Y \). Also, we define the **depth of a lottery** \( f \) in \( L_\rho(Y) \), denoted by \( \delta(f) = k \), if \( f \in L_k(Y) \) and \( \delta(f) = \infty \), if \( f \in L_{k+1}(Y) \). We use the same symbol \( \delta \) for the depth of a lottery and the depth of a probability.

Now, we formulate a connection from \( L_{k-1}(Y) \) to \( L_k(Y) \). Specifically, let \( f = (f_1, \ldots, f_\ell) \) be an \( \ell \) vector of lotteries \( \mathbf{f} = (f_1, \ldots, f_\ell) \) in \( L_{k-1}(Y)^\ell = L_{k-1}(Y) \times \cdots \times L_{k-1}(Y) \). We say that \( \mathbf{f} = (f_1, \ldots, f_\ell) \) is a decomposition of \( f \in L_k(Y) \) iff

\[
f(y) = \sum_{i=1}^{\ell} \frac{1}{r_i} f_i(y) \quad \text{for all} \quad y \in Y. \tag{7}
\]

Let \( \mathbf{e} = (\frac{1}{\ell}, \ldots, \frac{1}{\ell}) \), and \( f \) is denoted by \( \mathbf{e} \ast f \) or \( \sum_{i=1}^{\ell} \frac{1}{\ell} \ast f_i \).

Lemma 2.2 states that each lottery in \( L_k(Y) \) is expressed as a weighted sum of some \( (f_1, \ldots, f_\ell) \) in \( L_{k-1}(Y)^\ell \) with the equal weights. This connects \( L_{k-1}(Y) \) to \( L_k(Y) \), which facilitates our induction method described in Figure 1. A proof is given in the Appendix; it is not so simple as Lemma 2.1. The inclusion \( \supseteq \) of the right-hand set in \( L_k(Y) \) is simple, but the converse is complicated since its proof involves precise construction of a decomposition \( \mathbf{f} = (f_1, \ldots, f_\ell) \).

**Lemma 2.2 (Decomposition of lotteries):** \( L_k(Y) = \{ \mathbf{e} \ast f : f \in L_{k-1}(Y)^\ell \} \) for any \( k \) \( (1 \leq k < \infty) \).

The lottery \( d = [y, \frac{25}{10}; y] \) in Section 1 has two types of decompositions:

\[
\frac{5}{10} \ast [y, \frac{5}{10}; y] + \frac{5}{10} \ast y \quad \text{and} \quad \frac{2}{10} \ast y + \frac{1}{10} \ast [y, \frac{5}{10}; y] + \frac{7}{10} \ast y. \tag{8}
\]

In the first, a decomposition \( \mathbf{f} = (f_1, \ldots, f_{10}) \) is given as \( f_1 = \ldots = f_5 = [y, \frac{5}{10}; y] \) and \( f_6 = \ldots = f_{10} = y \). In the second, \( f \) is given as \( f_1 = f_2 = y, f_3 = [y, \frac{5}{10}; y] \) and \( f_4 = \ldots = f_{10} = y \). The proof of Lemma 2.2 constructs the second type of a decomposition in a general manner.

Finally, let \( \succcurlyeq \) be a binary relation over \( L_k(Y) \). The expression \( f \succcurlyeq g \) means that \( f \) is strictly preferred to or is indifferent to \( g \). We define the **strict (preference) relation** \( \succ \), **indifference relation** \( \sim \), and **incomparability relation** \( \nvdash \) by

\[
f \succ g \iff f \succcurlyeq g \text{ and not } g \succcurlyeq f; \tag{9}
f \sim g \iff f \succcurlyeq g \text{ and } g \succcurlyeq f;\]

\[
f \nvdash g \iff \text{neither } f \succcurlyeq g \text{ nor } g \succcurlyeq f.
\]
The incomparability relation \( \succcurlyeq \) is new and is studied in the subsequent sections. Nevertheless, all the axioms we assume are about the relation \( \succcurlyeq \), \( \succ\) and \( \sim \). The relation \( \succcurlyeq \) is characterized as the residual part of these relations.

The relation \( \succcurlyeq \) is a subset of \( L_\rho(Y) \times L_\rho(Y) \); we call a pair \( \langle f, g \rangle \in \succcurlyeq \) a preference instance, and \( \{ \langle f, g \rangle, \langle g, f \rangle \} \subseteq \succcurlyeq \) an indifference instance. For example, if \( f \succ g \), then \( \langle f, g \rangle \in \succcurlyeq \) but \( \langle g, f \rangle \notin \succcurlyeq \), and if \( f \succcurlyeq g \), there are no preference instances between \( f \) and \( g \). We sometimes omit “instance”.

3 EU Theory with Probability Grids

In this section, we give an axiomatic system describing decision making. We show that when \( \rho = \infty \), our theory is regarded as classical expected utility theory.

3.1 A sequence of preference relations up to a cognitive bound

We prepare a sequence of preference relations \( \langle \succcurlyeq^0, \ldots, \succcurlyeq^\rho \rangle \) for \( \rho < \infty \), and \( \langle \succcurlyeq^0, \succcurlyeq^1, \ldots \rangle \) for \( \rho = \infty \); either sequence is denoted by \( \langle \succcurlyeq^k \rangle_{k=0}^\rho \). Each \( \succcurlyeq^k \) is associated with a cognitive depth \( l(k) \) so that \( \succcurlyeq^k \) is a binary relation over \( L_{l(k)}(Y) \). We define the depth sequence \( \langle l(k) \rangle_{k=0}^\rho \) of \( \langle \succcurlyeq^k \rangle_{k=0}^\rho \) as follows:

\[
\begin{align*}
l(0) &= 0, \quad 1 \leq l(1) \leq \rho; \quad \text{and} \\
l(k + 1) &= \min\{l(k) + 1, \rho\} \quad \text{for all } k \ (1 \leq k < \rho).
\end{align*}
\]

Once \( l(k) \) reaches the cognitive bound \( \rho \), \( l(k) \) becomes constant; so, \( L_{l(k)}(Y) = L_\rho(Y) = L_{l(\rho)}(Y) \).

Then, \( \succcurlyeq^0 \) claims the existence of some preferences, and \( \succcurlyeq^k \) includes also some existences but some part of it is generated from \( \succcurlyeq^{k-1} \) for \( k \geq 1 \). The intended meanings are expressed in terms of the corresponding axioms.

We require the following uniformly over \( \langle \succcurlyeq^k \rangle_{k=0}^\rho \) for each \( \succcurlyeq^k \) in \( \langle \succcurlyeq^k \rangle_{k=0}^\rho \) with \( k < 1 + \rho \).

**Axiom B0. (Transitivity):** for any \( f, g, h \in L_{l(k)}(Y) \), if \( f \succcurlyeq^k g \) and \( g \succcurlyeq^k h \), then \( f \succcurlyeq^k h \).

Transitivity makes no existential claim on preferences, but it is a conditional statement. In this respect, this axiom differs from the others except Axiom 4.

We assume neither completeness: for any \( f, g \in L_{l(k)}(X) \), \( f \succcurlyeq^k g \) or \( g \succcurlyeq^k f \); nor reflexivity: for any \( f \in L_{l(k)}(X) \), \( f \succcurlyeq^k f \). These will be assumed partially, while they will be derived from our axioms when \( \rho = \infty \). Our main concern is to study incomparability \( f \not\succcurlyeq^k g \) when \( \rho < \infty \).

Now, we introduce the concept of a base facet, which requires properties for the results in Step (i) of measurement in the mind of the decision maker. We impose three axioms on a base facet. A base facet is given as \( F = (\mathbf{y}, \mathbf{y}; Y; \{ \lambda_y \}_{y \in Y}) \); \( \mathbf{y} \) and \( y \) are called the upper and lower benchmarks, \( Y \) is a finite set of pure alternatives with \( \mathbf{y}, y \in \mathbf{y} \), and \( \{ \lambda_y \}_{y \in Y} \) is a family in \( \Pi_{\rho} \) with \( \lambda_{\mathbf{y}} = 1 \) and \( \lambda_{\mathbf{y}} = 0 \). To simplify the subsequent presentation, we assume

\[
0 < \lambda_y < 1 \quad \text{for all } y \in Y \setminus \{ \mathbf{y}, \mathbf{y} \}; \quad (11)
\]

\[
1 \leq l(1) = \max\{ \delta(\lambda_y) : y \in Y \} < 1 + \rho. \quad (12)
\]

\(^5\)We stipulate \( 1 + \infty = \infty \). Then, we can express the two statements “\( k \leq \rho \) if \( \rho < \infty \) and \( k < \rho \) if \( \rho = \infty \)” as “\( k < 1 + \rho \).”
It follows from (12) that $0 < \lambda_y < 1$ for some $y \in Y$.

**Example 3.1.** Let $Y = \{\varphi, y, y\}$, $\lambda_\varphi = 1$, $\lambda_y = \frac{85}{100}$, and $\lambda_{y'} = 0$. Then, the smallest $\rho$ is 2; the description of $\lambda_y = \frac{85}{100}$ needs $L_2(Y)$. Then, $l(0) = 0$, $l(1) = l(2) = 2$. Here, $\succsim^0$ is defined over $L_0(Y) = Y$, but $\succsim^1$ and $\succsim^2$ are relations over $L_2(Y)$, but they still differ. When $\rho = 3$, $l(0) = 0$, $l(1) = 2$, and $l(2) = l(3) = 3$; $\succsim^2$, $\succsim^3$ are relations over $L_3(Y)$.

We call a lottery $f$ in $L_k(Y)$ a **benchmark lottery** of depth at most $k$ iff $f(\varphi) = \lambda$ and $f(y) = 1 - \lambda$ for some $\lambda \in \Pi_k$, which we denote by $[\varphi, \lambda; y]$. The **benchmark scale** of depth at most $k$ is given as $B_k([\varphi; \lambda; y]) = \{(\varphi, \lambda; y) : \lambda \in \Pi_k\}$: the dots in Figure 2 are the benchmark lotteries of at most depth $\rho = 2$. When $\rho = \infty$, we define $B_\rho([\varphi; y]) = B_\infty([\varphi; y]) = \cup_{k=0}^\infty B_k([\varphi; y])$.

Suppose that a sequence $([\varphi; y])^\rho_{k=0}$ is given. We adopt the following three axioms:

**Axiom B1** (Between the benchmarks): $\varphi > 0$ $y > 0$ for all $y \in Y - \{\varphi, y\}$.

**Axiom B2** (Measurement with the benchmarks): $y \sim 1 [\varphi, \lambda; y]$ for each $y \in Y$.

Let $k$ be a natural number with $0 \leq k < 1 - \rho$.

**Axiom B3** (Benchmark scales): for $\lambda, \lambda' \in \Pi(k)$, $\lambda \geq \lambda'$ if and only if $[\varphi, \lambda; y] \succsim^k [\varphi, \lambda'; y]$.

Axiom B1 states that any pure alternative $y \in Y - \{\varphi, y\}$ is between the upper and lower benchmarks. Axiom B2 states that each $y$ is measured in terms of $B_{l(k)}([\varphi; y])$. Axiom B3 states that $B_{l(k)}([\varphi; y])$ is a scale of measurement. It holds that

$$
\lambda = \lambda' \iff [\varphi, \lambda; y] \sim^k [\varphi, \lambda'; y]; \quad \text{and} \quad \lambda > \lambda' \iff [\varphi, \lambda; y] \succsim^k [\varphi, \lambda'; y].
$$

(13)

Hence, $\succsim^k$ is a complete and reflexive relation over $B_{l(k)}([\varphi; y])$ by (5). In sum, these three axioms describe Step (i) of measurement. We may write B3 when $k$ is not important.

Lemma 3.1 states that $F = ([\varphi, y]; Y; \{\lambda_y\}_{y \in Y})$ fixes a complete preference relation over $Y$.

**Lemma 3.1 (Measurement lemma).** Let $F = ([\varphi, y]; Y; \{\lambda_y\}_{y \in Y})$ be a base facet with Axioms B0, B2, and B3.

1. **(Completeness over $Y$):** $y \succsim^1 z$ or $z \succsim^1 y$ for any $y, z \in Y$;
2. **(Uniqueness):** For each $y \in Y$, if $y \sim^1 [\varphi, \lambda; y]$, then $\lambda = \lambda_y$.

**Proof (1):** Let $y, z \in Y$. By B2, $y \sim^1 [\varphi, \lambda_y; y]$ and $z \sim^1 [\varphi, \lambda_z; y]$. By (5), either $\lambda_y > \lambda_z$, $\lambda_y = \lambda_z$, or $\lambda_y < \lambda_z$. In the first case, $[\varphi, \lambda_y; y] \succsim^1 [\varphi, \lambda_z; y]$ by (13); therefore by B0, we have $y \succsim^1 z$. The third case is symmetric. In the second case, $y \sim^1 [\varphi, \lambda_y; y] = [\varphi, \lambda_z; y] \sim^1 z$ by (13). Thus, $y \sim^1 z$ by B0.

2. **Consider the contrapositive.** Let $\lambda > \lambda_y$. Then, by (13), $[\varphi, \lambda; y] \succsim^1 [\varphi, \lambda_y; y]$. $\lambda < \lambda_y$ is symmetric. Hence, $\lambda_y$ is unique.

Lemma 3.1 implies the existence of a utility function $u_o : Y \rightarrow R$ defined by $u_o(y) = \lambda_y$ all $y \in Y$; $u_o$ represents the relation $\succsim^1$ over $Y$, that is, for any $y, y' \in Y$,

$$
u_o(y) \geq u_o(y') \text{ if and only if } y \succsim^1 y'.
$$

(14)

This $u_o(\cdot)$ does not yet capture the entire $\succsim^1$ over $L_{l(1)}(Y)$, but plays an important role in consideration of the preference sequence $([\varphi; y])^\rho_{k=0}$.

### 3.2 Step of extension

Now, we give the axiom for Step (ii) of extension. For $f = (f_1, \ldots, f_\ell)$ and $g = (g_1, \ldots, g_\ell)$, we write $f \succsim^k g$ when $f_t \succsim^k g_t$ for all $t = 1, \ldots, \ell$. Recall that $f$ is called a decomposition of $f$ when
\[ f = e \ast f. \]

**Axiom B4 (Extension):** Let \( k < \rho \), and \( f, g \) decompositions of \( f \in L_{l(k+1)}(Y) \) and \( g \in B_{l(k+1)}(\gamma; y) \). Then, (1): \( f \succ^k g \) implies \( f \succ^{k+1} g \); and (2): \( g \succ^k f \) implies \( g \succ^{k+1} f \).

Each has the strict part, e.g., in (1), if \( f \succ^k g \) and \( f_t \succ^k g_t \) for some \( t \), then \( f \succ^{k+1} g \).

This generates new preference instances from \( \succ^k \). For \( \rho = \infty \), B4 is applied for any finite \( k \), but for \( \rho < \infty \), it is applied at most \( \rho \) times.

**Lemma 3.2 (Preservation):** Let \( k < \rho \), \( f \in L_{l(k)}(Y) \), and \( g \in B_{l(k)}(\gamma; y) \). Then, \( f \succ^k g \) implies \( f \succ^{k+1} g \); and \( g \succ^k f \) implies \( g \succ^{k+1} f \). Each has the strict part.

**Proof.** Suppose \( f \succ^k g \). Then, \( f, g \in L_{l(k)}(Y) \leq L_{l(k+1)}(Y) \). Let \( f_1 = \ldots = f_{\ell} = f \) and \( g_1 = \ldots = g_{\ell} = g \). Then, \( f = \sum_{t=1}^{\ell} \frac{1}{t} * f_t \) and \( f = \sum_{t=1}^{\ell} \frac{1}{t} * g_t \). Hence, \( (f_1, \ldots, f_{\ell}) \) and \( (g_1, \ldots, g_{\ell}) \) are decompositions of \( f \) and \( g \). By B4, we have \( f \succ^{k+1} g \).

The decision maker constructs his preference relations \( \succ^1, \ldots, \succ^k \) step-by-step from the preference instances given by Axioms B1 to B3. Extensions take the form of mathematical induction. The preferences in \( \succ^0 \) belong purely to the induction base, and the preferences in \( \succ^1 \) asserted by Axioms B2 and B3 are also part of the induction base, but some preferences in \( \succ^1 \) are already derived by Axiom B4. Recall \([y, \frac{5}{10}; y] = \frac{5}{10} * y + \frac{5}{10} * y\). Since \( y \succ^0 y \) by B1 and \( y \succ^0 y \) by B3\(^0\), we have the following by B4:

\[
[y, \frac{5}{10}; y] = \frac{5}{10} * y + \frac{5}{10} * y \succ^1 \frac{5}{10} * y + \frac{5}{10} * y = y.
\] (15)

This preference is not included in B2 and B3\(^1\), since \([y, \frac{5}{10}; y]\) is not a benchmark lottery.

One question is the consistency of the axiomatic system \( T = \langle F, L_\rho(Y); B0 \to B4 \rangle \); consistency means that given a base facet \( F \), there is a preference sequence \( \langle \succ^k \rangle_{k=0}^\rho \) satisfying Axioms B0 to B4. In fact, classical EU theory gives an answer to this question. First, we define the EU function \( u_{eu} \) over \( L_\infty(Y) \) based on \( u_o(y) = \lambda_y \) for all \( y \in Y \) and the eu-preference relation \( \succ_{eu} \) over \( L_\infty(Y) \) by

\[
u_{eu}(f) = \sum_{y \in Y} f(y) u_o(y) \text{ for any } f \in L_\infty(Y); \tag{16}
\]

for any \( f, g \in L_\infty(Y) \), \( f \succ_{eu} g \iff \nu_{eu}(f) \geq \nu_{eu}(g) \). \tag{17}

Then, we restrict this relation \( \succ_{eu} \) to \( L_{l(k)}(Y) \), denoted by \( \succ_{eu}^k \); that is, \( \succ_{eu}^k = \succ_{eu} \cap (L_{l(k)}(Y) \times L_{l(k)}(Y)) \). Note that \( \succ_{eu}^\rho \) is complete over \( L_{l(k)}(Y) \). By this restriction, we have \( \langle \succ_{eu}^k \rangle_{k=0}^\rho \) in the either case \( \rho < \infty \) or \( \rho = \infty \). Note that \( \succ_{eu} = \bigcup_{k=0}^\infty \succ_{eu}^k \) is the same as \( \succ_{eu}^\infty \). Then, we have the following two lemmas.

**Lemma 3.3.** It holds that:

\[
u_{eu}(e \ast f) = \sum_{t=1}^{\ell} \frac{1}{t} \nu_{eu}(f_t) \text{ for any } f \in L_k(Y)^F \text{ and } k \geq 0. \tag{18}
\]

**Proof.** It follows from (16) that \( \nu_{eu}(e \ast f) = \sum_{y \in Y} (e \ast f)(y) u_o(y) = \sum_{y \in Y} \sum_{t=1}^{\ell} \frac{1}{t} f_t(y) u_o(y) = \sum_{t=1}^{\ell} \frac{1}{t} \sum_{y \in Y} f_t(y) u_o(y) = \sum_{t=1}^{\ell} \frac{1}{t} \nu_{eu}(f_t). \)

The following lemma implies that the consistency of \( T = \langle F, L_\rho(Y); B0 \to B4 \rangle \).

**Lemma 3.4 (Consistency).** Let \( F = \langle \gamma, y; Y; \{\lambda_y\}_{y \in Y} \rangle \) be a basic facet. In either case \( \rho < \infty \) or \( \rho = \infty \), \( \langle \succ_{eu}^k \rangle_{k=0}^\rho \) satisfies Axioms B0 to B4.
Proof. \( \preceq_{\text{eu}}^{k, \rho} _{k=0} \) satisfies B0. Since \( u_{\text{eu}}(f) \) is based on \( u \), \( \preceq_{\text{eu}}^0 \) and \( \preceq_{\text{eu}}^1 \) satisfies B1 to B3. Using (18), we can verify B4. ■

Let \( \preceq_{\text{ku}}^{k, \rho} _{k=0} \) be any preference sequence satisfying Axioms B0 to B4. We focus on the resulting preference relation rather than intermediate preference relations in \( \preceq_{\text{k}u}^{k, \rho} _{k=0} \). The resulting (preference) relation of \( \preceq_{\text{k}u}^{k, \rho} _{k=0} \) is given as the last \( \preceq_{\text{ku}}^{\rho} \) if \( \rho < \infty \) and as the union \( \preceq_{\text{ku}}^{\infty} = \cup_{k=1}^{\infty} \preceq_{\text{ku}}^{k} \) if \( \rho = \infty \). The relation \( \preceq_{\text{ku}}^{\infty} \) is regarded as classical EU theory.

**Theorem 3.1 (EU Theorem).** Consider any \( \preceq_{\text{ku}}^{k, \infty} _{k=0} \) with B0 to B4. Then,

\[
\text{for any } f, g \in L_{\infty}(Y), \quad f \preceq_{\text{ku}}^{\infty} g \iff f \preceq_{\text{eu}} g.
\]

**Proof:** We show by induction over \( k = 0, \ldots \) that

\[
f \sim_{k+1} f, u_{\text{eu}}(f); y \]

for all \( f \in L_{\text{lu}}(Y) \).

Let \( k = 0 \). Then, \( L_{0}(Y) = Y \). By Lemma 3.1, for any \( y \in Y \), there is a unique \( \lambda \) such that \( y \sim_{0} f, u_{\text{eu}}(f) \). Moreover, for any \( y \in Y \), there is a unique \( \lambda \) such that \( y \sim_{0} f, u_{\text{eu}}(f) \).

Suppose the induction hypothesis that (20) holds for \( k \). Take any \( f \in L_{\text{lu}(k+1)}(Y) \). By Lemma 2.2, we have a decomposition \( f \in L_{\text{lu}(k+1)}(Y) \) such that \( \mathbf{e} \star f = f \). By the induction hypothesis, there is a \( g = (g_1, \ldots, g_{\ell}) \) such that \( f \sim_{k} g \) and \( g_t = [f, u_{\text{eu}}(f); y] \) for \( t = 1, \ldots, \ell \). Applying B4, we have \( f = \mathbf{e} \star f \sim_{k+1} \mathbf{e} \star g \), and this becomes \( f \sim_{k+1} \sum_{t=1}^{\ell} \mathbf{e} \star g_t = \sum_{t=1}^{\ell} \mathbf{e} \star [f, u_{\text{eu}}(f); y] \)

Let \( f, g \in L_{\infty}(Y) \). By (20), for some \( k, f \sim_{k} [f, u_{\text{eu}}(f); y] \) and \( g \sim_{k} [f, u_{\text{eu}}(f); y] \) for all \( k \geq k_0 \). Since \( \preceq_{\infty} = \cup_{k=0}^{\infty} \preceq_{\text{ku}}^{k} \), we can take a \( k \geq k_0 \) so that \( f \preceq_{\infty} g \iff f \preceq_{k} g \). By B0, (17), and B3, \( f \preceq_{k} g \iff [f, u_{\text{eu}}(f); y] \preceq_{k} [f, u_{\text{eu}}(f); y] \iff u_{\text{eu}}(f) \geq u_{\text{eu}}(f) \). In sum, \( f \preceq_{\infty} g \iff f \preceq_{\text{ku}}^{\infty} g \) ■

Since each \( \preceq_{\text{ku}}^{k} \) in \( \preceq_{\text{ku}}^{k, \infty} _{k=0} \) is complete, \( \preceq_{\text{ku}}^{k, \infty} _{k=0} \) is the strongest, i.e., for any \( \preceq_{\text{ku}}^{k, \infty} _{k=0} \) with B0 to B4,

\[
\text{for any } f, g \in L_{\text{lu}(k)}(Y) \text{ and } k \geq 0, \quad f \preceq_{k} g \implies f \preceq_{\text{ku}}^{k} g.
\]

The relation \( \preceq_{\text{ku}}^{k} \) includes a lot of preference instances not derived from B0 to B4. This is the polar opposite to our target relation that is derived purely from B0 to B4. This will be discussed in Section 4.1.

Our theory \( T = \{F, L_{\rho}(Y); B0 \text{ to } B4\} \) with \( \rho = \infty \) is regarded as classical EU theory in that the resulting \( \preceq_{\text{ku}}^{\infty} = \cup_{k=0}^{\infty} \preceq_{\text{ku}}^{k} \) is represented by the expected utility function in the sense of (16) and (17). Allowing \( \ell \)-ary compound lotteries, we can have a direct axiomatization for \( \preceq_{\text{ku}}^{\infty} \) not through \( \preceq_{\text{ku}}^{k, \infty} _{k=0} \).

## 4 Constructed Relations \( \preceq_{\text{ku}}^{k, \rho} _{k=0} \) and Measurable Domain \( M(F; \rho) \)

In the literature of utility theory, axioms are typically regarded as “natural requirements” for a preference relation. In contrast, our axioms describe the process of decision making from the base preferences to complex ones. Indeed, we can construct \( \preceq_{\text{ku}}^{k, \rho} _{k=0} \) purely from B0 to B4, which is the logically weakest among \( \preceq_{\text{ku}}^{k, \rho} _{k=0} \)'s satisfying B0 to B4. When \( \rho = \infty \), the resulting relation \( \preceq_{\text{ku}}^{\infty} \) is still the same as \( \preceq_{\text{ku}}^{\infty} \) . Then, we consider the measurable domain \( M(F; \rho) \) in terms of \( \preceq_{\text{ku}}^{\rho} \), which coincides with \( L_{\rho}(Y) \) when \( \rho = \infty \). Our aim is to study the behavior of \( \preceq_{\text{ku}}^{\rho} \) in \( L_{\rho}(Y) \) when \( \rho < \infty \).
4.1 Constructed relations $\langle \succeq^k \rangle_{k=0}^\rho$

First, we look at the nature of each axiom. Axiom B0 (transitivity) itself asserts no preference instances; instead, it is conditional that if some preferences are given, new preferences are constructed. On the contrary, B1 to B3 assert some preference instances. Axiom B4 is conditional similar to B0.

Let a base facet $F$ be given. We define $\succeq_0^c$ and $\succeq_1^c$ by

$$\succeq_0^c = B1 \cup B3^0$$
$$\succeq_1^c = ([\succeq_0^c]^4 \cup B2 \cup B3^1)^{tr}.$$  \hspace{1cm} (22) 

Here, $B1 \cup B3^0$ is the set of preferences asserted by B1 and B3$^0$. That is, $\succeq_0^c = \{(y, y), (y, y) : y \neq y, \{y, y\}\} \cup \{(y, y), \{y, y\}, \{y, y\}\}$. This set satisfies transitivity. In the second, $([\succeq_0^c]^4$ is the set of preferences derived from $\succeq_0^c$ by B4, and $B2 \cup B3^1$ is the set of preferences stated in Axioms B2 and B3$^1$, and then its transitive closure is taken; the set with the outer $[\cdot]^tr$ is the transitive closure of the set included in $\cdot$.$^6$

Suppose that $\succeq_k^c$ is already defined for $k \geq 1$. Now, we define $\succeq_{k+1}^c$ as follows:

$$\succeq_{k+1}^c = ([\succeq_k^c]^4 \cup B3^{k+1})^{tr},$$  \hspace{1cm} (23) 

where $([\succeq_k^c]^4$ is the set of preferences obtained from $\succeq_k^c$ by B4, the set $B3^{k+1}$ consists of preferences asserted by Axiom B3$^{k+1}$, and then we take its transitive closure.

In (22) and (23), the transitive closure is taken after adding the preferences asserted by the correspond axioms. It is important to note that each preference asserted by B1 to B4 contains a benchmark lottery. We can write this observation as the following lemma.

**Lemma 4.1.** If $f \succeq_{k+1}^c g$, there are $f = h_0, h_1, ..., h_m = g$ in $L(l(k+1))Y$ such that for $n = 0, ..., m - 1$,

$$\langle h_n, h_{n+1} \rangle \in \left\{ \begin{array}{ll}
(\succeq_k^c)^4 \cup B2 \cup B3^1 & \text{if } k = 0 \\
(\succeq_k^c)^4 \cup B3^{k+1} & \text{if } k \geq 1
\end{array} \right.$$  \hspace{1cm} (24) 

and each preference instance $\langle h_n, h_{n+1} \rangle$ contains at least one lottery from $B(l(k+1))Y$.

The sequence $\langle \succeq_k^c \rangle_{k=0}^\rho$ is uniquely defined either for $\rho < \infty$ or $\rho = \infty$. We call this $\langle \succeq_k^c \rangle_{k=0}^\rho$ the **constructed (preference) sequence**. Then, we have the following theorem.

**Theorem 4.1 (Weakest).** The constructed sequence $\langle \succeq_k^c \rangle_{k=0}^\rho$ is the weakest among any sequences $\langle \succeq_k^c \rangle_{k=0}^\rho$'s with B0 to B4, that is; for any $f, g \in L(\rho)Y$ and any $k (0 \leq k < 1 + \rho)$,

$$f \succeq_k^c g \implies f \succeq_k^c g.$$  \hspace{1cm} (25) 

**Proof.** We can check that $\langle \succeq_k^c \rangle_{k=0}^\rho$ satisfies Axioms B0 to B4. Each relation $\succeq_k^c$ satisfies B0, since it is the transitive closure of the inside of $\cdot$. Axioms B1 to B3 hold by construction. Axiom B4 holds since $([\succeq_k^c]^4$ is assumed in the definition of $\succeq_{k+1}^c$ for $k \geq 0$.

We prove (25) by induction over $k < 1 + \rho$. For $k = 0, 1$, we have (25) by (22). Suppose that (25) holds for $k$, that is, $\succeq_k^c$ is a subset of $\succeq_k^c$. Hence, $([\succeq_k^c]^4 \subseteq \langle \succeq_k^c \rangle^4$, so, $\succeq_k^c = ([\succeq_k^c]^4 \cup B3^{k+1})^{tr} \subseteq \succeq_{k+1}^c$. The last inclusion is by B4, B3$^{k+1}$, and B0 for $\succeq_{k+1}^c$.$\blacksquare$

$^6$The transitive closure of an $S \subseteq L(\rho)Y \times L(\rho)Y$ is the set of all pairs in $L(\rho)Y \times L(\rho)Y$ that are connected by a finite chain of pairs in $S$. 

11
When \( \rho = \infty \), it follows from (21) and (25) that \( \preceq_{k+1}^\infty \) is the strongest sequence and \( \preceq_c^k \) is the weakest. Nevertheless, Theorem 3.1 implies that \( \preceq_c^\infty = \cup_{k=0}^\infty \preceq_c^k \) coincides with the eu-relation \( \preceq_{eu} \).

It follows from Lemmas 3.2 (preservation) and 4.1 that the constructed sequence \( \preceq_c^k \) satisfies the monotonicity that for \( k < k' < 1 + \rho \), \( \preceq_c^k \) is a subset of \( \preceq_c^{k'} \). This may not be satisfied by an arbitrary \( \preceq_c^k \).

Each \( \preceq_c^k \) in \( \preceq_c^k \) depends upon \( \rho \) when \( \rho < \infty \), because \( L(k) \) may reach \( \rho \) before \( k = \rho \), i.e., \( \preceq_c^k \) and \( \preceq_c^{k+1} \) are defined over the same \( L(l(k)) \), but if \( \rho \) increases, then \( \preceq_c^{k+1} \) is defined over the larger set \( L(l(k+1)) \). In this paper, we do not discuss this dependence, we use the expression \( \preceq_c^k \).

We call the resulting relation \( \preceq_c^\rho \) of \( \preceq_c^k \) the central (preference) relation. When \( \rho < \infty \), this relation \( \preceq_c^\rho \) involves incomparabilities, which will be studied in Section 5.

### 4.2 Measurable domain \( M(F; \rho) \)

Let \( \preceq_c^k \) be the sequence of constructed relations. We define \( M^k(F; \rho) \) by: for \( k < 1 + \rho \),

\[
M^k(F; \rho) = \{ f \in L(l(k))(Y) : f \sim^k_c [\gamma, \lambda; \mu] \text{ for some } \lambda \in \Pi_l(k) \}.
\]

(26)

By (22), \( M^0(F; \rho) = \{ \gamma, \lambda; \mu \} \) and \( M^1(F; \rho) = Y \cup B_{l(1)}(\gamma, \lambda) \). It holds by Lemma 3.2 that \( M^k(F; \rho) \subseteq M^{k+1}(F; \rho) \) for \( k < \rho \). We define \( M(F; \rho) = M^0(F; \rho) \) if \( \rho < \infty \) and \( M(F; \rho) = \cup_{k=0}^{\infty} M^k(F; \rho) \) if \( \rho = \infty \). We call \( M(F; \rho) \) the measurable domain, and each \( f \in M(F; \rho) \) a measurable lottery (with respect to the benchmark scale \( B_{l(0)}(\gamma, \lambda) \)). We study the behavior of \( \preceq_c^\rho \) in \( L(\rho)(Y) - M(F; \rho) \) as well as in \( M(F; \rho) \).

It follows from Axiom B3 that the probability weight \( \lambda \) in the right-hand side of (26) is unique for each \( f \in M(F; \rho) \); thus we can denote this unique \( \lambda \) by \( \lambda_f \). Then, it holds that \( f \sim^\rho_c g \) and \( g \in B_{l(0)}(\gamma, \lambda) \iff g = [\gamma, \lambda_f; \mu] \).

A measurable lottery in \( M(F; \rho) \) can be reduced to a vector of measurable lotteries.

**Theorem 4.2 (Reduction of a measurable lottery).** Let \( 1 \leq k < \rho \). Then, \( f \in M^{k+1}(F; \rho) \) if and only if

\[
f = e * f \quad \text{for some } f = (f_1, ..., f_\ell) \in M^k(F; \rho) \ell.
\]

(27)

**Proof.** The if part follows from Axiom B4. We prove the only-if part. For \( k = 1 \), we can modify the following proof by using (22) rather than (23). Suppose that \( 1 < k < \rho \). Take any \( f \in M^{k+1}(F; \rho) \). Then, there is a \( g \in B_{l(k+1)}(\gamma, \lambda) \) such that \( f \sim^k_c g \).

By Lemma 4.1, we can take a sequence \( f = h_0, h_1, ..., h_m = g \) in \( L(l(k+1))(Y) \) so that the each indifference \( h_n \sim^{k+1}_c h_{n+1} \) belongs to \( (\preceq_c^k)^4 \cup B_{l(k)}(\gamma, \lambda) \). We can assume that these lotteries are all distinct. This implies that \( h_n \sim^{k+1}_c h_{n+1} \) belongs to \( (\preceq_c^k)^4 \). Since this relation is obtained by B4, \( h_n \) and \( h_{n+1} \) have decompositions \( h_n = (h_{n,1}, ..., h_{n,\ell}) \) and \( h_{n+1} = (h_{n+1,1}, ..., h_{n+1,\ell}) \) so that \( h_n \sim^k_c h_{n+1} \) and one of \( h_n \), \( h_{n+1} \) is in \( B_{l(k)}(\gamma, \lambda)^\ell \).

The above holds for any \( n = 0, ..., m-1 \). Hence, by B0, we have \( h_{0,t} \sim^k_c h_{m,t} \) for all \( t = 1, ..., \ell \). Thus, \( f = h_0 \) and \( h_m = g \) have decompositions \( h_0 \), \( h_m \) with \( h_{0,t} \sim^k_c h_{m,t} \) for \( t = 1, ..., \ell \). Since \( h_m = g \in B_{l(k+1)}(\gamma, \lambda) \), we have \( h_m \in B_{l(k)}(\gamma, \lambda)^\ell \). Thus, since \( h_{0,t} \sim^k_c h_{m,t} \) for each \( t = 1, ..., \ell \), each \( f_t = h_{0,t} \) belongs to \( M^{k}(F; \rho) \). This means (27).

The following theorem states that the central relation \( \preceq_c^\rho \) coincides with the eu-preference relation \( \preceq_{eu} \) over \( M(F; \rho) \). For \( \rho = \infty \), this theorem corresponds to Theorem 3.1 (EU-Theorem).
Theorem 4.3 (Uniqueness of $\succeq_c^\rho$ over $M(\mathbb{F}; \rho)$).
(1): $\lambda_f = u_{\text{eu}}(f)$ for any $f \in M(\mathbb{F}; \rho)$.
(2): For any $f, g \in M(\mathbb{F}; \rho)$, $f \succeq_{\text{eu}} g \iff f \succeq_c^\rho g$.

Proof. (1): We prove by induction over $k = 1, \ldots$ that $\lambda_f = u_{\text{eu}}(f)$ for any $f \in M_k(\mathbb{F}; \rho)$. Let $k = 1$. Then, $f \in M_1(\mathbb{F}; \rho) = Y \cup B_{(1)}(\mathbb{F}; y)$. Then, if $f \in Y$, we have $\lambda_f = u_{\text{eu}}(f)$ by Lemma 3.1(2); and if $f \in B_{(1)}(\mathbb{F}; y)$, then $u_{\text{eu}}(f) = \lambda_f \cdot 1 + (1 - \lambda_f) \cdot 0 = \lambda_f$. Now, suppose that $\lambda_f = u_{\text{eu}}(f)$ for any $f \in M^k(\mathbb{F}; \rho)$ and $k < \rho$. Let $f \in M^{k+1}(\mathbb{F}; \rho)$. Then $f \sim^{k+1}_c [\mathbb{F}; \lambda_f; y]$. By Theorem 4.2, there is an $f = (f_1, \ldots, f_t) \in M^k(\mathbb{F}; \rho)^t$ such that $f = e * f$. For each $t = 1, \ldots, t$, we have $\lambda_{f_t} = u_{\text{eu}}(f_t) \in \Pi_{l(k)}$ since $f_t \in M^k(\mathbb{F}; \rho)$. Hence, by B4, $[\mathbb{F}; \lambda_f; y] \sim^{k+1}_c f = e * f \sim^{k+1}_c \sum_{t=1}^t \frac{1}{t} [\mathbb{F}; \lambda_{f_t}; y] = [\mathbb{F}; \sum_{t=1}^t \frac{1}{t} \lambda_{f_t}, y]$. By B3 and (18), we have $\lambda_f = \sum_{t=1}^t \frac{1}{t} \lambda_{f_t} = \sum_{t=1}^t u_{\text{eu}}(f_t) = u_{\text{eu}}(\sum_{t=1}^t \frac{1}{t} f_t) = u_{\text{eu}}(f)$.

(2): Let $f, g \in M(\mathbb{F}; \rho)$. By (26), $f \sim^{\rho}_c [\mathbb{F}; \lambda_f; y]$ and $g \sim^{\rho}_c [\mathbb{F}; \lambda_g; y]$. By B0, B3, and (1) of this theorem, $f \succeq_{\text{eu}} g \iff [\mathbb{F}; \lambda_f; y] \succeq^{\rho}_c [\mathbb{F}; \lambda_g; y] \iff u_{\text{eu}}(f) = \lambda_f \geq \lambda_g = u_{\text{eu}}(g) \iff f \succeq_{\text{eu}} g$.\]

The following theorem gives a necessary condition for $f \in M(\mathbb{F}; \rho)$ and its implication.

Theorem 4.4 (Criterion for $M(\mathbb{F}; \rho)$).

(1): $f \in M(\mathbb{F}; \rho)$ implies $\delta(u_{\text{eu}}(f)) \leq \rho$.
(2): $M(\mathbb{F}; \rho) = L_\rho(Y)$ if and only if $\rho = \infty$.

Proof. (1): Let $f \in M(\mathbb{F}; \rho)$. Then, $\lambda_f \in \Pi(\rho)$. Since $u_{\text{eu}}(f) = \lambda_f$ by Theorem 4.3(1), we have $\delta(u_{\text{eu}}(f)) = \delta(\lambda_f) \leq l(\rho) = \rho$.

(2): By definition, $M(\mathbb{F}; \rho) \subset L_\rho(Y)$. We consider the converse. Let $\rho = \infty$. Take an $f \in L_\infty(Y)$. Then, by Theorem 3.1 (in particular, (20)), we have $f \sim^{\rho}_c [\mathbb{F}; \lambda_f; y]$ where $\lambda_f = u_{\text{eu}}(f)$. Hence, $f \in M(\mathbb{F}; \infty)$.

Let $\rho < \infty$. By (12), there is some $y \in Y$ with $\delta(\lambda_y) > 0$. This means that $\lambda_y$ is expressed as $\frac{1}{\nu}$ for some $k \geq 1$ and $\nu$, where $\nu$ has no factor $\ell$. Consider $f = [y, \frac{1}{\nu}; y] \in L_\rho(Y)$. Then, $u_{\text{eu}}(f) = \frac{1}{\nu} \cdot \lambda_y$; so, $\delta(u_{\text{eu}}(f)) = \rho + \delta(\lambda_y) > \rho$. Hence, $f \notin M(\mathbb{F}; \rho)$ by (1) of this theorem. Thus, $M(\mathbb{F}; \rho) \subset L_\rho(Y)$.

The converse of (1) does not hold: for example, let $Y = \{\mathbb{F}, y, y\}$, $\lambda_y = \frac{2}{10}$, and $\rho = 1$. The lottery $f = [y, \frac{5}{10}; y]$ belongs to $L_1(Y)$. Here, $u_{\text{eu}}(f) = \frac{5}{10} \cdot \frac{1}{10} = \frac{1}{10}$ and $\delta(u_{\text{eu}}(f)) = 1$. However, since $y \sim^{\rho}_c [\mathbb{F}, \frac{1}{2}; y]$, we cannot apply B4 to obtain $f \in M(\mathbb{F}; \rho)$.

The first assertion of the following theorem states that there are no indifferences in $L_\rho(Y) - M(\mathbb{F}; \rho)$, though lotteries in $L_\rho(Y) - M(\mathbb{F}; \rho)$ show strict preferences with some lotteries. The second states that reflexivity holds exactly in $M(\mathbb{F}; \rho)$.

Theorem 4.5. Let $f, g \in L_\rho(Y)$.

(1) (No indifferences outside $M(\mathbb{F}; \rho)$): If $f \notin M(\mathbb{F}; \rho)$, then $f \not\sim^{\rho}_c g$.
(2) (Reflexivity): $f \sim^{\rho}_c f$ if and only if $f \in M(\mathbb{F}; \rho)$.

Proof (1): Suppose that $f \notin M(\mathbb{F}; \rho)$ and $g \in M(\mathbb{F}; \rho)$. If $f \not\sim^{\rho}_c g$, then $f \in M(\mathbb{F}; \rho)$ by B0, a contradiction. Hence, $f \sim^{\rho}_c g$.

Now, let $f, g \notin M(\mathbb{F}; \rho)$. Suppose $f \sim^{\rho}_c g$. There is a smallest $k \leq \rho$ such that $f \sim^{k}_c g$. By Lemma 4.1, there are $f = h_0, h_1, \ldots, h_m = g \in L_{(k)}(Y)$ such that for $n = 0, \ldots, m - 1$, each pair $h_n, h_{n+1}$ satisfies (24), i.e., $h_n \not\sim^{k}_c h_{n+1}$ or $h_{n+1} \not\sim^{k}_c h_n$ for $n = 1, \ldots, m - 1$. Since $f \sim^{k}_c g$, we have $h_n \sim^{k}_c h_{n+1}$ for $n = 1, \ldots, m - 1$. Lemma 4.1 states that for each $n = 0, \ldots, m - 1$, at least one of $h_n$ and $h_{n+1}$ belongs to $B_{(k)}(\mathbb{F}; y)$. Since $h_0 = f \notin M(\mathbb{F}; \rho)$ and $h_1 \in B_{(k)}(\mathbb{F}; y) \subset M(\mathbb{F}; \rho)$, we
have, by the above paragraph, $h_0 \succsim^k_c h_1$, a contradiction. Hence, $f \not\succsim_c g$.

(2): The if part is by (26) and B0, and the only-if part (contrapositive) follows from (1) of this theorem.

Our basic principle is that the benchmarks are used as a scale of measurement; thus Axiom $B3^k$ assumes that $\succsim^k$ satisfies completeness (including reflexivity) and transitivity over $\mathcal{B}_{l(k)}(\mathcal{F}, \mathcal{Y})$. Only lotteries in $M(F; \rho)$ are measured by this principle; and reflexivity holds only for $f \in \mathcal{M}(F; \rho)$. Reflexivity is almost innocuous; but if we adopt it as an axiom, we need to restate the results in Section 5.

5 Incomparabilities in $\succsim^\rho$ and Representation

The central relation $\succsim^\rho$, which is the resulting relation of $(\succsim^k)^\rho\ k=0$, is complete over the measurable domain $M(F; \rho)$, but show some incomparabilities outside $M(F; \rho)$. It follows from Theorems 4.3 and 4.4 that this occurs only when $\rho < \infty$. In this section, we assume $\rho < \infty$ and study incomparabilities in $\succsim^\rho$. We define the concepts, LUB (lowest upper bound) and GLB (greatest lower bound), of $f \in L_\rho(Y)$ with respect to $\succsim^\rho$. Using these, we characterize incomparabilities in $\succsim^\rho$, and summarize these characterizations as a representation of $\succsim^\rho$ in terms of a 2-dimensional vector-valued function taking the LUB and GLB of each $f \in L_\rho(Y)$. In the following, we abbreviate $\succsim^\rho$ as $\succsim_c$ when no confusions are expected.

5.1 LUB and GLB

We have the following three mutually exclusive and exhaustive cases:

- **C**: $f, g \in M(F; \rho)$;
- **IA**: $f \in L_\rho(Y) - M(F; \rho)$ and $g \in M(F; \rho)$ (and the symmetric case);
- **IB**: $f, g \in L_\rho(Y) - M(F; \rho)$.

Theorem 4.3 states that in case **C**: $\succsim_c = \succsim^\rho$ coincides with $\succsim^\rho$; $f$ and $g$ are always comparable. To study incomparabilities involved in the central relation $\succsim_c$ in cases **IA** and **IB**, we introduce the LUB and GLB of each $f \in L_\rho(Y)$. First, we need to show the following lemma.

**Lemma 5.1.** $\mathcal{Y} \succsim_c f \succsim_c \mathcal{Y}$ for any $f \in L_\rho(Y)$.

**Proof.** We show by induction over $k = 0, ..., \rho$ that $\mathcal{Y} \succsim^k_c f \succsim^k_c \mathcal{Y}$ for any $f \in L_k(Y)$. Note that the assertion is not for “$f \in L_{\ell(k)}(Y)$”. Let $f \in L_0(Y) = Y$. Then, by B1 and B3$^0$, we have the assertion. Now, let $f \in L_1(Y)$. Then, $f$ can be expressed as $f = \sum_{t=1}^\ell \frac{1}{t} * y_t$ for some $\{y_1, ..., y_\ell\} \subseteq Y$. Since $\mathcal{Y} \succsim^0_c y_t$ for all $y_t \in Y$ by B1, we have, by B4, $\mathcal{Y} = \sum_{t=1}^\ell \frac{1}{t} * \mathcal{Y} \succsim^1_c \mathcal{Y}$.

Suppose the induction hypothesis that $\mathcal{Y} \succsim^k_c f \succsim^k_c \mathcal{Y}$ for any $f \in L_k(Y)$ with $k < \rho$. Since $L_k(Y) \subseteq L_{\ell(k)}(Y)$, this hypothesis makes sense. Consider $f \in L_{k+1}(Y)$. Then, by Lemma 2.2, there is a vector $f = (f_1, ..., f_\ell) \in L_k(Y)$ such that $f = e * f$. By the induction hypothesis, $\mathcal{Y} \succsim^k_c f_t \succsim^k_c \mathcal{Y}$ for any $t = 1, ..., \ell$. By Axiom B4, we have $\mathcal{Y} = e * \mathcal{Y} \succsim^k+1_c f = e * f \succsim^k+1_c e * \mathcal{Y} = \mathcal{Y}$, that is, $\mathcal{Y} \succsim^k+1_c f \succsim^k+1_c \mathcal{Y}$.

Since $\mathcal{Y}$ and $\mathcal{Y}$ are in $\mathcal{B}_\rho(\mathcal{F}, \mathcal{Y})$, Lemma 5.1 guarantees that every $f \in L_\rho(Y)$ has upper and
lower bounds in $\mathcal{B}_\rho(\bar{\gamma}; y)$. We define the LUB $\overline{x}_f$ and GLB $\underline{\lambda}_f$ of $f \in L_\rho(Y)$ by

$$
\overline{x}_f = \min \{ \lambda_g : g \in \mathcal{B}_\rho(\bar{\gamma}; y) \text{ with } g \succ_c f \};
\underline{\lambda}_f = \max \{ \lambda_g : g \in \mathcal{B}_\rho(\bar{\gamma}; y) \text{ with } f \succ_c g \}.
$$

(28)

The following observations are useful: for $f \in L_\rho(Y)$,

$$
\overline{x}_f = \underline{\lambda}_f = \lambda_f \text{ if } f \in M(F; \rho) \quad \text{and} \quad \overline{x}_f > \underline{\lambda}_f \text{ if } f \notin M(F; \rho).
$$

(29)

The first holds by Theorem 4.3. Consider the second. Let $f \notin M(F; \rho)$. Then, the preferences inside in (28) are strict by Theorem 4.5. Hence, $[\bar{\gamma}, \overline{x}_f; y] = g \succ_c f \succ_c h = [\underline{\gamma}, \underline{\lambda}_f; y]$ for some $g, h \in \mathcal{B}_\rho(\bar{\gamma}; y)$. By B0, we have $[\bar{\gamma}, \overline{x}_f; y] \succ_c [\underline{\gamma}, \underline{\lambda}_f; y]$; so $\overline{x}_f > \underline{\lambda}_f$ by B3.

Theorem 5.1 characterizes incomparability between $f \notin M(F; \rho)$ and $g \in M(F; \rho)$; recall that $f \succ_c g$ means that $f$ and $g$ are incomparable with respect to $\succ_c$. This theorem will be used for consideration of the Allais paradox in Section 6.

**Theorem 5.1 (Incomparability in case IA):** Let $f \notin M(F; \rho)$ and $g \in M(F; \rho)$. Then,

$$
f \succ_c g \iff \underline{\lambda}_f \geq \lambda_g; \quad \text{and} \quad g \succ_c f \iff \lambda_g \geq \overline{x}_f; \quad \text{(30)}$$

$$
f \succ_c g \iff \overline{x}_f > \lambda_g > \underline{\lambda}_f. \quad \text{(31)}$$

**Proof.** Since $f \notin M(F; \rho)$ and $g \in M(F; \rho)$, we have $f \prec_c g$ by Theorem 4.5. Consider the first equivalence of (30). If $f \succ_c g$, then $f \succ_c g \sim_c [\bar{\gamma}, \lambda_g; y]$, which implies $\underline{\lambda}_f \geq \lambda_g$ by (28) and B3. Conversely, if $\underline{\lambda}_f \geq \lambda_g$, then $f \succ_c [\bar{\gamma}, \underline{\lambda}_f; y] \succ_c [\underline{\gamma}, \lambda_g; y] \succ_c g$ by (28) and B3, which implies $f \succ_c g$ by B0 and $f \prec_c g$. The other equivalence of (30) is similar.

By (30), $f$ and $g$ are comparable if and only if $\lambda_g \geq \overline{x}_f$ or $\underline{\lambda}_f \geq \lambda_g$. This is equivalent to (31).

We have a similar characterization of incomparability in case IB.

**Theorem 5.2 (Incomparability in case IB)** Let $f, g \notin M(F; \rho)$. We have the following characterizations of $\succ_c$ and $\succ_c$:

$$
f \succ_c g \iff \underline{\lambda}_f \geq \overline{x}_f; \quad \text{(32)}$$

$$
f \succ_c g \iff \overline{x}_f > \underline{\lambda}_f \text{ and } \overline{x}_f > \lambda_g. \quad \text{(33)}$$

When one of $f$ and $g$ are in $M(F; \rho)$, (32) are (33) are reduced to (30) and (31). In sum, Theorems 5.1 and 5.2 provide complete characterization of incomparabilities involved in the central relation $\succ_c = \succ_c^\rho$.

**Proof of Theorem 5.2.** Since $f \prec_c g$ by Theorem 4.5, (33) follows from (32).

We prove that for $f, g \notin M(F; \rho)$,

(1): $f \succ_c g \iff f \succ_c h \succ_c g$ for some $h \in \mathcal{B}_\rho(\bar{\gamma}; y)$;

(2): $f \succ_c h \succ_c g$ for some $h \in \mathcal{B}_\rho(\bar{\gamma}; y) \iff \underline{\lambda}_f \geq \overline{x}_f$.

These imply (32). Let us see (2). Suppose that $f \succ_c h \succ_c g$ for some $h \in \mathcal{B}_\rho(\bar{\gamma}; y)$. Since $\underline{\lambda}_f$ is the LUB of $f$, by (28), $\underline{\lambda}_f \geq \lambda_h$. Similarly, $\lambda_h \geq \overline{x}_f$. Conversely, let $\underline{\lambda}_f \geq \overline{x}_f$. Then, by (28), $f \succ_c [\bar{\gamma}, \overline{x}_f; y] \succ_c [\underline{\gamma}, \underline{\lambda}_f; y] \succ_c g$. Hence, we can adopt $[\bar{\gamma}, \overline{x}_f; y]$ for $h$.

Consider (1): The direction $\iff$ is by B0. Now, suppose $f \succ_c g$, and we derive a contradiction from:

there is no $h \in \mathcal{B}_\rho(\bar{\gamma}; y)$ such that $f \succ_c h \succ_c g$. \quad \text{(34)}
Since $L_p(Y)$ is finite, there is a finite sequence of distinct $g_0, ..., g_m$ such that $f = g_m \succ_c ... \succ_c g_0 = g$. By (34), all of the lotteries are in $L_p(Y) - M(F; \rho)$, because, otherwise, there would be a $g_s \in M(F; \rho)$ such that $f \succ g_s \succ g$; so $f \succ g_s \succ [\gamma, \lambda_{g_s} ; y] \succ g$, a contradiction to (34). Since they are distinct in $L_p(Y) - M(F; \rho)$, it holds by Theorem 4.5 that $g_{n+1} \succ g_n$ for all $n$. We take one pair $g_{n+1}$ and $g_n$. For simplicity, we let $f$ and $g$ are such a pair assuming that they are immediately next.

Now, we define $(\succ^0_k, \succ^1_k, ..., \succ^\rho_k)$ by

$$\succ^k \equiv \succ^k_c - \{(f, g)\} \text{ for all } k = 0, ..., \rho.$$  \hfill (35)

It suffices to show that $(\succ^0_k, \succ^1_k, ..., \succ^\rho_k)$ satisfies B0 to B4, which is a contradiction to the fact that $\succ^c$ is the weakest relation satisfying B0 to B4; thus, $f \succ g$ implies that $f \succ h \succ g$ for some $h \in B(p; y)$. Now, we prove $(\succ^0_k, \succ^1_k, ..., \succ^\rho_k)$ satisfies each axiom.

B0: Recall that $f$ and $g$ are immediately next by definition. The preference $f \succ^k_g$ is a conclusion of transitivity only if one of the premises of transitivity is $f \succ^k_g$ itself. Hence, elimination of $(f, g)$ means that this transitivity holds for $\succ^k_c$ in the trivial sense.

B1, B2, B3: Since these are about lotteries in $Y \cup B(k; y) \subseteq M(F; \rho)$ and $f, g \notin M(F; \rho)$, these axioms are not affected by (35) for $\succ^k_c$.

B4: This is about a pair of lotteries where at least one is a benchmark lottery. However, neither $f$ nor $g$ is a benchmark lottery. Hence, B4 is not affected by (35).$\blacksquare$

5.2 Representation theorem

We consider the vector-valued function $\Lambda(f) = (\lambda_f, \lambda_f)$ for any $f \in L_p(Y)$. By (29), $\Lambda(f)$ consists of a vector of identical components if and only if $f \in M(F; \rho)$. We define the binary relation $\geq$ over $\Pi_p \times \Pi_p$ by

$$(\xi_1, \xi_2) \geq (\eta_1, \eta_2) \iff \xi_2 \geq \eta_1.$$  \hfill (36)

Using (29), this relation is transitive over $\{\Lambda(f) : f \in L_p(Y)\}$ but does not necessarily satisfy reflexivity since $\lambda_f > \lambda_g$ if $f \notin M(F; \rho)$. Hence, $\geq$ is not a partial ordering on $\{\Lambda(f) : f \in L_p(Y)\}$.

Using this function $\Lambda(\cdot)$ with $\geq$, we can summarize Theorems 4.3, 5.1, and 5.2.

**Theorem 5.3 (Representation).** For any $f, g \in L_p(Y)$, $f \succ_c g \iff \Lambda(f) \geq \Lambda(g)$.

**Proof.** Consider case C: $f, g \in M(F; \rho)$. Since $\Lambda(f) = (\lambda_f, \lambda_f)$ and $\Lambda(g) = (\lambda_g, \lambda_g)$, the right-hand side of (36) is $\lambda_f \geq \lambda_g$. Thus, the assertion follows from Theorem 4.3.

When at least one of $f, g$ does not belong to $M(F; \rho)$, we have $f \sim_c g$ by Theorem 4.5.(1).

This is applied to two cases: IA and IB.

Now, consider case IA: $f \notin M(F; \rho)$ and $g \in M(F; \rho)$. Hence, the assertions are stated as $f \succ g \iff \lambda_f \geq \lambda_g$ and $g \succ f \iff \lambda_g \geq \lambda_f$. These are (30) of Theorem 5.1. Hence, $f \succ g \iff \Lambda(f) \geq \Lambda(g)$ and $g \succ f \iff \Lambda(g) \geq \Lambda(f)$.

Consider case IB: $f, g \notin M(F; \rho)$. The assertion is stated as $f \succ g \iff \lambda_f \geq \lambda_g$. This equivalence holds by (32) of Theorem 5.2.$\blacksquare$

In the context of classical expected utility theory, von Neumann-Morgenstern [22], p.29, indicated a possibility of a representation of a preference relation involving incompatibilities in terms of a many-dimensional vector function. Theorem 5.3 takes a special type of their indication, though ours needs a 2-dimensional vector function.
Debura et al. [8] and Baucells-Shapley [5] obtained the representation result in the form that an incomplete preference relation is represented by a class of expected utility functions. In this literature, available probabilities are given as arbitrary real numbers in the interval $[0, 1]$. In our approach, this may be interpreted as corresponding to the case of $\rho = \infty$, though $\Pi_\infty = \bigcup_{k=0}^{\infty} \Pi_k$ is a countable subset of the set of rational numbers, where completeness is derived. In Section 7, we will briefly mention that when there are multiple base facets, incompleteness would be natural even when $\rho = \infty$: perhaps, this corresponds to the approach by Debura et al. [8] and Baucells-Shapley [5].

### 6 Allais Paradox

We examine the Allais paradox in our EU theory. First, we need to have a connection from the theory to experimental outcomes. Incomparability plays an important role in this connection.

We use the experimental result reported in Kahneman-Tversky [12]. In the KT 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 $d$.

$$a = [4000, \frac{80}{100}; 0]; \ (20\%); \ b = 3000 \ \text{with probability} \ 1; \ (80\%)$$

$$c = [4000, \frac{20}{100}; 0]; \ (65\%); \ d = [3000, \frac{25}{100}; 0]; \ (35\%).$$

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

$$0.80u(4000) + 0.20u(0) < u(3000) \quad (37)$$

$$0.20u(4000) + 0.80u(0) > 0.25u(3000) + 0.75u(0).$$

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 $a \land d$. EU theory itself predicts the outcomes $a \land c$ and $b \land d$. This is a variant of many experiments reported since Allais [2].

In [12], no more information is mentioned about the choices other than the percentages. Consider three possible distributions of the answers in terms of percentages over the four cases, described in Table 6.1: the first, second, or third entry in each cell is the percentage derived by assuming 65%, 52%, or 45% for $b \land c$. The first 65% is the maximum percentage for $b \land c$, which implies 0% for $a \land c$, and these determine the 20% in $a \land d$ and 15% in $b \land d$. The second entries are calculated 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 $b \land c$. In the third entries, 45% is the minimum possibility for $b \land c$. We interpret this table as meaning that each cell was observed as non-negligible at a significant level.

<table>
<thead>
<tr>
<th></th>
<th>$c: 65%$</th>
<th>$d: 35%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a: 20%$</td>
<td>$a \land c: EU$: 0 // 13 // 20</td>
<td>$a \land d: paradox: 20 // 7 // 0$</td>
</tr>
<tr>
<td>$b: 80%$</td>
<td>$b \land c: paradox: 65 // 52 // 45$</td>
<td>$b \land d: EU$: 15 // 28 // 35</td>
</tr>
</tbody>
</table>

*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 [16] and its references).*
We return to our theory. A set of choice problems $E$ for an experiment consists of unordered pairs $\{f, g\}$ of distinct lotteries in $L_{\rho}(Y)$. An outcome $\psi$ is a function on $E$ with $\psi(\{f, g\}) = f$ or $g$ for each $\{f, g\} \in E$. The set of outcomes significantly observed is denoted by $O(E)$. In the KT example, the set of choice problems is $E_{KT} = \{\{a, b\}, \{c, d\}\}$, and the outcome $b \wedge c$ is expressed as $\psi_{bc}(\{a, b\}) = b$ and $\psi_{bc}(\{c, d\}) = c$. As Table 6.1 was interpreted as meaning that all the four cases occurred at non-negligible levels, we set $O(E_{KT}) = \{\psi_{ac}, \psi_{ad}, \psi_{bc}, \psi_{bd}\}$.

We ask the question of whether some theory $T = \langle L_{\rho}(Y), F; B0$ to B4 $\rangle$ “explains” a given observed outcome $\psi \in O(E_{KT})$. To make the term “explain” meaningful, we specify a domain where $T$ varies. Here, the domain is given in the context of the KT example; a theory $T$ has two parameters $\lambda_y$ and $\rho$. We need $\rho \geq 2$ to have lottery $d$. Given $\rho$, $T(\rho)$ denotes the set of all theories with possible values $\lambda_y \in \Pi(l(1))$ with $2 \leq l(1) \leq \rho$. Each $T \in T(\rho)$ is written as $T(\lambda_y)$.

We say that $T(\rho)$ is non-trivially consistent with $\psi \in O(E_{KT})$ iff there is a $T(\lambda_y) \in T(\rho)$ such that

$$f \succ_c g \text{ for some } \{f, g\} \in E_{KT},$$

for any $\{f, g\} \in E_{KT}$, $f \succ_c g \Rightarrow \psi(\{f, g\}) = f$,

where $\succ_c$ is the central relation determined by $T(\lambda_y) \in T(\rho)$. Thus, each observed outcome $\psi \in O(E_{KT})$ is explained in a non-trivial manner by a specification of $\lambda_y$.

We have the following result.

**Theorem 6.1 (KT experiment) (1):** $T(\rho)$ is non-trivially consistent with each $\psi$ in $O(E_{KT})$ if and only if $\rho = 2$.

**2:** Let $\rho \geq 3$ and $\psi \in O(E_{KT})$. Then, $T(\rho)$ is non-trivially consistent with $\psi$ if and only if $\psi = \psi_{ac}$ or $\psi = \psi_{bd}$.

Thus, our theory explains all the fourth experimental outcomes in Table 6.1 for $\rho = 2$, and it returns to the Allais paradox for $\rho \geq 3$. For $\rho = 2$, by an appropriate choice of $\lambda_y$, (38) holds avoiding an indifferences between $a$ and $b$, and (39) holds for $\{c, d\}$ in the trivial sense since $c$ and $d$ are incomparable. Here, it is our presumption that when people are asked to choose one from $c$ and $d$, they would choose one even though those are incomparable for them. This presumption will be discussed more after the proof of Theorem 6.1. In the following proof, we choose particular $\lambda_y$, but the argument works more generally and quite accurately; which is discussed also after the proof.

First, we calculate the LUB $\bar{\lambda}_d$ and GLB $\underline{\lambda}_d$ of $d$ for $\rho = 2$:

$$\text{if } \lambda_y = \frac{85}{110}, \text{ then } \bar{\lambda}_d = \frac{23}{110} \text{ and } \underline{\lambda}_d = \frac{16}{110};$$

(40)

$$\text{if } \lambda_y = \frac{25}{110}, \text{ then } \bar{\lambda}_d = \frac{21}{110} \text{ and } \underline{\lambda}_d = \frac{14}{110}.$$  

Consider case $\lambda_y = \frac{85}{110}$. The GLB $\underline{\lambda}_d = \frac{16}{110}$ is calculated as follow: first, $l(1) = 2$, and we have $y \sim_{1} \frac{[7, 85/110; y]}{\frac{85}{110}} \succ_{1} \frac{[7, 5/110; y]}{\frac{5}{110}}$ by B2 and B3. By B0, B3, and B4, $[y, \frac{5}{110}; y] = \frac{5}{110} \ast y + \frac{5}{110} \ast y \succ_{1} y$. Using these, $d = [y, \frac{25}{110}; y]$ is reduced, by B4, to

$$\frac{2}{110} \ast y + \frac{5}{110} \ast [y, \frac{5}{110}; y] + \frac{5}{110} \ast y \succ_{2} \frac{2}{110} \ast [y, \frac{8}{110}; y] + \frac{1}{110} \ast y + \frac{7}{110} \ast y = [y, \frac{16}{110}; y].$$

Thus, $d = [y, \frac{25}{110}; y] \succ_{2} [y, \frac{16}{110}; y]$. Lottery $d = [y, \frac{25}{110}; y]$ can be decomposed in different manners, but this is the best evaluation of a lower bound of $d$ in $B_2(Y, y)$.

The LUB $\bar{\lambda}_d = \frac{23}{110}$ is obtained as follows: Similar to the above, $[y, \frac{9}{110}; y] \succ_{1} y$, and by B4, $[y, \frac{5}{110}; y] = \frac{5}{110} \ast y + \frac{5}{110} \ast y \succ_{1} \frac{5}{110} \ast y + \frac{5}{110} \ast y = [y, \frac{5}{110}; y]$. Using these, we have

$$[y, \frac{23}{110}; y] = \frac{2}{110} \ast [y, \frac{9}{110}; y] + \frac{1}{110} \ast [y, \frac{5}{110}; y] + \frac{7}{110} \ast y \succ_{2} \frac{2}{110} \ast y + \frac{1}{110} \ast [y, \frac{5}{110}; y] + \frac{7}{110} \ast y.$$  

18
The last one is $d$, and this is the best upper evaluation of $d$.

When $\lambda_y = \frac{75}{10^2}$, the LUB and GLB are calculated similarly.

**Proof of Theorem 6.1.(1):** We show that $T(\lambda_y)$ with $\lambda_y = \frac{85}{10^2}$ is non-trivially consistent with $\psi_{bc}$ and $\psi_{bd}$; and $T(\lambda_y)$ with $\lambda_y = \frac{75}{10^2}$ is non-trivially consistent with $\psi_{ac}$ and $\psi_{ad}$.

Let $\lambda_y = \frac{85}{10^2}$. Here, $b \succ^c a$; so, (38) holds. Consider the choice between $c$ and $d$. We already obtained (40): $\lambda_d = \frac{23}{10^2} > \lambda_c = \frac{20}{10^2} > \lambda_d = \frac{16}{10^2}$. Hence, by Theorem 5.1, $c$ and $d$ are incomparable. Hence, (39) holds for either $\psi_{bc}$ and $\psi_{bd}$ in a trivial sense. Thus, $T(\lambda_y)$ with $\lambda_y = \frac{85}{10^2}$ is non-trivially consistent with $\psi_{bc}$ and $\psi_{bd}$.

Let $\lambda_y = \frac{75}{10^2}$. Here, $a \succ^c b$, i.e., (38) holds. Since $\lambda_d > \lambda_c = \frac{20}{10^2} > \lambda_d$, by Theorem 5.1, $c$ and $d$ are incomparable. Hence, (39) holds for either $\psi_{ac}$ and $\psi_{ad}$ in a trivial sense.

(2): Let $\rho \geq 3$. We have the following:

\[ \lambda_y = \frac{8}{10} \implies c = \{\gamma, \frac{3}{10}, y\} \sim^3 d = \{y, \frac{25}{10}, \gamma\}; \]
\[ \lambda_y > \frac{8}{10} \implies d \succ^c c; \text{ and } \lambda_y < \frac{8}{10} \implies c \succ^3 d. \]

In case $\lambda_y = \frac{8}{10}$, $a \sim^c b$, and $c \succ^3 d$ is obtained from $y \sim^1 \{\gamma, \frac{8}{10}, y\}$ and $d = \{y, \frac{25}{10}, \gamma\}$; since $\{y, \frac{8}{10}, y\} = \frac{1}{10} + y + \frac{9}{10} * y \sim^2 \{y, \frac{8}{10}, y\}$ and thus $d = \frac{2}{10} * y + \frac{5}{10} * \{y, \frac{8}{10}, y\} + \frac{3}{10} * y \sim^3 \{\gamma, \frac{20}{10}, y\}$. Thus, (38) does not hold. In case $\lambda_y > \frac{8}{10}$, modifying the argument for $\lambda_y = \frac{8}{10}$, we obtain $b \succ^c a$ and $d \succ^c c$; thus, for $\rho = 3$, $T(\lambda_y)$ is non-trivially consistent only with $\psi_{bd}$. In case $\lambda_y < \frac{8}{10}$, we have $a \succ^c b$ and $c \succ^3 d$; $T(\lambda_y)$ is non-trivially consistent only with $\psi_{ac}$. For $\rho \geq 4$, the above conclusion remains by Lemma 3.2.

Let us look at case $\rho = 2$ from the numerical point of view. In fact, (40) holds more generally. Consider $(A): \lambda_y \in \{\frac{81}{10^2}, \ldots, \frac{29}{10^2}\}$ and $(B): \lambda_y \in \{\frac{71}{10^2}, \ldots, \frac{79}{10^2}\}$. In each case, the LUB $\lambda_d$ and GLB $\lambda_d$ of $d = \{y, \frac{25}{10}, \gamma\}$ are uniformly given as $(A): \lambda_d = \frac{21}{10^2}$ and $\lambda_d = \frac{16}{10^2}$; and $(B): \lambda_d = \frac{21}{10^2}$ and $\lambda_d = \frac{14}{10^2}$.

Consider case $(A)$. Since $c = \{\gamma, \frac{2}{10}, y\}$ and $d$ are incomparable for a subject with $\lambda_y \in \{\frac{81}{10^2}, \ldots, \frac{29}{10^2}\}$, he could not make a decision. However, he was asked to choose either $c$ or $d$, and would make a choice. Here, we give a postulate: each subject interprets $d$ as if $d$ was $\{\gamma, \lambda; y\}$ with one $\lambda \in \{\frac{17}{10^2}, \ldots, \frac{22}{10^2}\}$; if $\lambda \in \{\frac{17}{10^2}, \frac{18}{10^2}, \frac{19}{10^2}\}$, he would choose $c$; if $\lambda \in \{\frac{21}{10^2}, \frac{22}{10^2}\}$, he would choose $d$; and if $\lambda = 10^2$, his propensity of choice $c$ or $d$ is equal. Also, we postulate that the people are equally distributed over $\{\frac{17}{10^2}, \ldots, \frac{22}{10^2}\}$, then the average ratio of choices is $3.5 : 2.5 = 7 : 5$. Similarly, in case $(B)$, the ratio is: $5.5 : 0.5 = 11 : 1$.

Taking 20% for $a$ and 80% for $b$, the over all percentage of $c$ is given as:

\[ (\frac{20}{10^2} \times \frac{11}{10^2} + \frac{80}{10^2} \times \frac{7}{10^2}) \times 100 = 65\% \]

This is exactly the same as the overall percentage of $c$ reported in [12]. We took 20% for $a$ and 80% for $b$ from in [12], but the ratios for $c$ and $d$ are from our theory with the assumption of uniform distribution over $\{\frac{17}{10^2}, \ldots, \frac{22}{10^2}\}$ in case $(A)$ and $\{\frac{16}{10^2}, \ldots, \frac{20}{10^2}\}$ in case $(B)$; thus (42) is a conjecture of our theory. The exact equivalence itself is not more than coincidence. Nevertheless, our development serves some good way to reconcile expected utility theory with the Allais paradox.

7 Two Remarks on Further Theoretical Developments

We give two remarks on our theory. The first is on the base facet $F = (\gamma, y; \{\lambda_y\}_{y \in Y})$; specifically, it is about a derivation of value $\lambda_y$ as well as $Y$ itself relative to given benchmarks.
Following Simon’s [19] idea of “satisficing/aspiration”. The second is about vertical and horizontal extensions of a base facet with different benchmarks.

(1): Derivation of a base facet by Simon’s satisficing/aspiration: Our theory has some affinity to Simon’s [19] notion of satisficing/aspiration. Let $\mathcal{F}, y$ be given benchmark lotteries, let $x$ be a pure alternative from a given set $X$. The decision maker evaluates the proposition:

$$x \sim (\mathcal{F}, \lambda; y)$$

based on his satisficing quintuple $(\mathcal{F}, y; \varphi; \pi)$, where $\varphi: X \times \Pi_{(1)} \to R_+$ is a satisficing function, and $\pi$ is an aspiration level in $R_+$. We assume that for each $x \in X$, $\varphi(x, \lambda)$ is a single-peaked (in the weak sense) function over each $\Pi_k$ $(1 \leq k \leq l(1))$. The value $\varphi(x, \lambda)$ ($\lambda \in \Pi_k$) is the satisficing degree of proposition (43). The value $\varphi(x, \lambda)$ is compared with the aspiration level $\pi$, and the assessment of (43) is accepted if and only if $\varphi(x, \lambda) \geq \pi$.

The decision maker starts with $\Pi_1$ and continues evaluating a given $x \in X$ using $\Pi_t$ until he finds some $\lambda_x$ with $\varphi_t(\lambda, x) \geq \pi$ or if he meets the upper limit $l(1)$ without finding, he quits his search. Let $Y_t$ be the set of pure alternatives $x$ having (43) in $\Pi_t$, and $Y = \cup_{t \leq l(1)} Y_t$. Thus, we have a base facet $F = (\mathcal{F}, y; Y; \{\lambda_y\}_{y \in Y})$.

The choice of $\lambda_x$ has still some arbitrariness (imprecision); we evaluate the size of this arbitrariness. For each $y \in Y_t$ $(t \leq l(1))$, we denote, by $\lambda_{sy}$ and $\lambda^{*}_y$, the smallest and largest $\lambda$ satisfying $\varphi(y, \lambda) \geq \pi$. Since $\varphi(y, \lambda)$ is single-peaked with respect to $\lambda \in \Pi_t$,

$$\lambda_{sy} \leq \lambda \leq \lambda^{*}_y \text{ if and only if } \varphi(y, \lambda) \geq \pi.$$  

(44)

Since such a $\lambda_y$ is not found at round $t - 1$ and since $\varphi_t(\lambda, x)$ is single-peaked, the difference between $\lambda^{*}_y$ and $\lambda_{sy}$ is $\lambda^{*}_y - \lambda_{sy} \leq \left(\frac{1}{t} - \frac{1}{2}\right) - \left(\frac{1}{t} + \frac{1}{2}\right) = \frac{t - 2}{2t}$. For $t = 2$, this difference is $\frac{4}{8}$, if $\ell = 10$, then $\frac{t - 2}{2t} = \frac{8}{10}$. In the example of Section 6, when $\lambda_{sy}$ and $\lambda^{*}_y$ are given as $\frac{83}{100}$ and $\frac{87}{100}$, the probability $\lambda_y = \frac{85}{100}$ adopted in Section 6 is approximated by the lower $\frac{83}{100}$ and upper $\frac{87}{100}$.

Thus, the value $\lambda_y$ for a pure alternative $y$ is also subject to imprecision in addition to the cognitive bound $\rho$ on permissible probabilities. This is related to the literature of imprecise probability (cf., Augustin et al. [1]) and of similarity (cf., Rubinstein [17], Tserenjigmid [21]). In this literature, imprecision/similarity is given in the mind of the decision maker under the assumption that expression/similarity is defined over all real number probabilities. In our treatment, imprecision is involved in the decision maker’s bounded thought process of finding probabilities in probability grids. Following Simon [19], after this thought process, probability values are treated as fixed, but imprecision remains from the outside point of view.

(2): Vertical and horizontal extensions of a base facet: We have assumed that the benchmarks $\mathcal{F}$ and $y$ are given. The choice of the lower $y$ could be natural, for example, the status quo. The choice of $\mathcal{F}$ may be more temporary in nature. In general, benchmarks $\mathcal{F}$ and $y$ are not really fixed; there are different benchmarks than the present ones. We consider some extensions of the benchmark choices.

One possibility is a vertical extension: we take another pair of benchmarks $\mathcal{F}$ and $y$ with some relation such as $\mathcal{F} \succ 0 y \succ 0 y$. The new set of pure alternatives is given as $Y(\mathcal{F}; y)$. Then, the set $Y(\mathcal{F}; y)$ may be included in $Y(\mathcal{F}; y)$. This extension may be straightforward where there is no cognitive bound $\rho$, but with a cognitive bound $\rho$, 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. In the case of preferences, we may have multiple preference systems even for similar target problems, which
may not be unified.

Another possible extension is a horizontal extension. For example, $y$ is the present status quo for a student who faces a choice problem between the alternative $\overline{y}$ of going to work for a large company and the alternative $\overline{y}$ of going to graduate school. The student may not be able to make a comparison between $\overline{y}$ and $\overline{y}$, which he can make a comparison between detailed choices after the choice of $\overline{y}$ or $\overline{y}$. This involves incomparabilities different from those in this paper.

In case $\rho = \infty$, each base facet leads to completeness within itself, but, across the facets, we may still have incomparabilities. This incompleteness may correspond to the literature on EU theory without the completeness axiom since Aumann [4], Dubra, et al. [8] and Baucells-Shapley [5].

To have an extension both in the vertical and horizontal manners, it may be crucial to consider a required cognitive depths as well as incomparabilities. This is an open problem of importance.

8 Conclusions and Some Remaining Problems

We developed EU utility theory with probability grids and preferential incomparabilities. The set of available probabilities is restricted to the form of $\ell$-ary fractions up to a given cognitive bound $\rho$. The theory has two steps: measurement and extension. Axiom B0 (transitivity) is uniformly assumed. Then, the measurement step is formulated in terms of a base facet $F = (\overline{y}, y; \{\lambda_{y}\}_{y \in Y})$ with Axioms B1 to B3. The extension step is formulated as Axiom B4. When $\rho = \infty$, Axioms B0 to B4 determine uniquely the complete preference relation $\succ_{\text{eu}}^\infty$; this corresponds to classical EU theory. Our main concern was the bounded case $\rho < \infty$.

When $\rho < \infty$, the resulting preference relation $\succ_{\rho}$ is neither unique nor complete. To study this case, we provided the measurable domain $M(F; \rho)$; the resulting $\succ_{\rho}$ shows completeness inside $M(F; \rho)$, while it involves incomparabilities outside $M(F; \rho)$. In Section 5, we gave a complete characterization of incomparabilities involved in $\succ_{\rho}$, and also the representation theorem of $\succ_{\rho}$ in terms of the 2-dimensional vector-valued function taking the values LUB and GLB of each lottery $f$.

We applied our theory to the Allais paradox in Section 6. We showed that the prediction of our theory is compatible with the experimental result in Kahneman-Tversky [12] for $\rho = 2$, and that for $\rho > 2$, the theory leads to the Allais paradox. Incomparability is crucial for the result for $\rho = 2$.

Our theory with Remark (1) in Section 7 may synthesize the decision maker’s past experiences, behavioral criteria, and beliefs/knowledge. This is related to the problem of how “probability” is interpreted; in particular, the frequentist viewpoint (cf., Hu [11]). Perhaps, this is also related to the foundation of inductive game theory (cf., Kaneko-Kline [13]), which remains open.

Another problem is to connect our theory to the literature of subjective utility/probability from Savage [18] and Anscombe-Aumann [3] and also to the recent literature on subjective utility/probability without the completeness axiom (cf., Nau [15], Galaabaatar-Karni [9]). This is remaining.

Appendix: Proof of Lemma 2.2.

Let $k \geq 1$. We show that if $f \in L_k(Y)$, then $f = e * f$ for some $f \in L_{k-1}(Y)^\ell$. We can assume that $\delta(f) = k$; indeed, if $\delta(f) \leq k - 1$, then we take $f = (f, \ldots, f) \in L_{k-1}(Y)^\ell$ and $e = e * f$. By $\delta(f) = k \geq 1$, for each $y \in Y$, $f(y)$ is expressed as $\sum_{m=1}^{k} \frac{v_m(y)}{\ell^m}$, where $0 \leq v_m(y) < \ell$ for all
$D_1 = \{1, \ldots, \tau_1\}$ \ldots \ldots $D_\ell = \{\tau_{k-1} + 1, \ldots, \tau_\ell\}$

$\forall (i, j) \in I_1 \times I_\ell$

\[
\sum_{y \in Y} \frac{v(y)}{\ell^t} \quad \ldots \quad \sum_{y \in Y} \frac{v(y)}{\ell^t}
\]

\begin{align*}
D_1 &= \{1, \ldots, \tau_1\} & \ldots & & D_\ell &= \{\tau_{k-1} + 1, \ldots, \tau_\ell\} \\
\mathcal{W}(i) &= \frac{1}{\ell^t} & \ldots & \frac{1}{\ell^k}
\end{align*}

Figure 3: Construction of $I_1, \ldots, I_\ell$

$m \leq k$.

We would like to partition $\sum_{y \in Y} \sum_{m=1}^{k} \frac{v_m(y)}{\ell^m} = \frac{1}{\ell^t}$ into $\ell$ portions so that each has the same sum $\frac{1}{\ell^t}$. However, this may not be directly possible; for example, if $\ell = 10$ and $v_1(y) = 2$, then $\frac{v_1(y)}{\ell^t} = \frac{2}{10}$ exceeds $\frac{1}{10}$. We avoid this by dividing $\frac{v_m(y)}{\ell^m}$ into $\frac{1}{\ell^m} + \ldots + \frac{1}{\ell^t}$. In Figure 3, $\sum_{y \in Y} \frac{v_1(y)}{\ell^t}$ is represented by $1, \ldots, \tau_1 := \sum_{y \in Y} v_1(y)$ with weight $\frac{1}{\ell}$ for each $i = 1, \ldots, \tau_1$. In general, $\sum_{y \in Y} \frac{v_m(y)}{\ell^m}$ is represented by a set natural numbers $D_m$ with weight $\frac{1}{\ell^m}$ for each element in $D_m$. Thus, the set $I = \{1, \ldots, \tau_k\}$ with $\tau_k = \sum_{m=1}^{k} \sum_{y \in Y} v_m(y)$ is partitioned in $D_1, \ldots, D_\ell$ with respect to the depths of associated weights. We partition this set $I$ once more into $I_1, \ldots, I_\ell$ so that the summation of the weights over each $I_t$ is $\frac{1}{\ell^t}$ for $t = 1, \ldots, \ell$. Using these two partitions, we construct $\mathbf{f} = (f_1, \ldots, f_\ell)$.

Formally, let $\tau_0 = 0$ and $\tau_m = \sum_{m=1}^{M} \sum_{y \in Y} v_d(y)$ for $m = 1, \ldots, k$. Then, let $I = \{1, \ldots, \tau_k\}$, and $D_m = \{\tau_{m-1} + 1, \ldots, \tau_m\}$ for $m = 1, \ldots, k$. We associate the weight $w(i)$ to each $i \in I$ by

$$w(i) = \frac{1}{\ell^m} \text{ if } i \in D_m.$$  \hspace{1cm} (45)

Each $\frac{v_m(y)}{\ell^m}$ is represented by the $v_m(y)$ number of elements in $D_m$ associated with weights $\frac{1}{\ell^m}$. Then, $\sum_{i \in D_m} w(i) = \sum_{y \in Y} \frac{v_m(y)}{\ell^m}$, and $\sum_{m=1}^{k} \sum_{i \in D_m} w(i) = \sum_{m=1}^{k} \sum_{y \in Y} \frac{v_m(y)}{\ell^m} = \sum_{y \in Y} f(y) = 1$.

Now, we define the function $W$ over $I$ by: $W(j) = \sum_{i \leq j} w(i)$ for any $j \in I$. Then, $W(\tau_k) = \sum_{m=1}^{k} \sum_{i \in D_m} w(i) = 1$. It holds that for any $j \in I$ and $t = 1, \ldots, \ell$,

$$\frac{1}{\ell^t} \leq W(j) < \frac{1}{\ell} \implies W(j + 1) \leq \frac{1}{\ell^t}.$$  \hspace{1cm} (46)

Indeed, if $j \in D_m$, then $W(j)$ is expressed as $\frac{1}{\ell^m}$ for some positive integer $s < \ell^m$, and $W(j + 1) = W(j) + \frac{1}{\ell^m}$ for some $m' \geq m$. Thus, we have $W(j + 1) \leq \frac{1}{\ell^t}$. Using this fact and $W(\tau_k) = 1$, we find another partition $I_1, \ldots, I_\ell$ of $I$ so that $I_t = \{\sigma_{t-1}, \ldots, \sigma_t\}$ and $\sigma_0 = 0 < \sigma_1 < \ldots < \sigma_\ell = \tau_k$ such that

$$\sum_{i \in I_t} w(i) = \frac{1}{\ell^t} \text{ for } t = 1, \ldots, \ell.$$  \hspace{1cm} (47)

These $I_1, \ldots, I_\ell$ are nonempty, while some of $D_1, \ldots, D_\ell$ may be empty. Each $I_t$ consists of successive elements in $I$. If $D_1 \neq \emptyset$, then $I_1, I_{t_1}$ are singleton sets, but the others may include some of $D_1, \ldots, D_\ell$. In Figure 3, the last $I_\ell$ may include $D_\ell$ and others.

For each $m = 1, \ldots, k$, we label each element $i$ in $D_m = \{\tau_{m-1} + 1, \ldots, \tau_m\}$ by $\varphi(i) \in Y$. Since
Thus, \( D_m \) represents \( \sum_{y \in Y} v_m(y) \), we can find a function \( \varphi : D_m \to Y \) so that it partition \( D_m \) with \( v_m(y) = \{ i \in D_m : \varphi(i) = y \} \) for each \( y \in Y \).

We define lotteries \( f_1, \ldots, f_\ell \in L_{k-1}(Y) \) by: for \( t = 1, \ldots, \ell \),

\[
f_t(y) = \sum_{m=1}^{k} \frac{\{ i \in I_t \cap D_m : \varphi(i) = y \}}{\ell^{m-1}} \quad \text{for each } y \in Y. \tag{48}
\]

If \( \{ i \in I_t \cap D_1 : \varphi(i) = y \} \) = 1, then, by (47), \( I_t \) consists of a unique element \( i \) with \( w(i) = \frac{1}{\ell} \), in which case \( \{ i \in I_t \cap D_m : \varphi(i) = y \} = 0 \) for \( m > 1 \). The other case is \( \{ i \in I_t \cap D_1 : \varphi(i) = y \} \leq \{ i \in D_m : \varphi(i) = y \} = v_m(y) < \ell \) for all \( m \leq k \), the value \( f_t(y) \) is expressed as \( \frac{\nu}{\ell^{k-1}} \) for some \( \nu \leq \ell^{k-1} \). Also, we have, by (47),

\[
\sum_{y \in Y} f_t(y) = \sum_{y \in Y} \sum_{m=1}^{k} \frac{\{ i \in I_t \cap D_m : \varphi(i) = y \}}{\ell^{m-1}} = \sum_{m=1}^{k} \frac{|I_t \cap D_m|}{\ell^{m-1}} = \ell \times \sum_{i \in I_t} w(i) = 1.
\]

Thus, \( f_t \in L_{k-1}(Y) \) for all \( t = 1, \ldots, \ell \). Finally, \( \sum_{t=1}^{\ell} \frac{1}{\ell} f_t(y) \) is calculated as

\[
\sum_{t=1}^{\ell} \frac{1}{\ell} f_t(y) = \sum_{m=1}^{k} \sum_{t=1}^{\ell} \frac{\{ i \in I_t \cap D_m : \varphi(i) = y \}}{\ell^{m}} = \sum_{m=1}^{k} \frac{v_m(y)}{\ell^{m}} = f(y).
\]

References


